Macroeconomic Effects on the Stockholm Stock Exchange

An Application of the Arbitrage Pricing Theory

Robin Pettersson

Robin Pettersson
Spring 2013
Bachelor Thesis, 15 credits
Economics C-level 30 ECTS
Supervisor: Kurt Brännäs
Abstract

In my thesis I estimated the impact of macroeconomic factors, belonging to the United States of America and Eurozone, on the return of the Swedish stock market index OMXS30. The theoretical perspective of this thesis was the Arbitrage Pricing Theory, which states that the return of a stock market index is a linear function of relevant macroeconomic factors. Since the macroeconomic factors I used in the estimation originated from other economic regions than the stock market index, economic interdependence is also an interesting part of this thesis. To estimate the effect of the chosen macroeconomic factors on the return of the OMXS30, two multiple time series models were created and estimated with different results. Using ordinary least squares I draw the conclusion that it is plausible that the chosen macroeconomic factors affect the return of the OMXS30.
Table of Contents

1. Introduction........................................................................................................................................... 1

2. Theory, data and statistical models........................................................................................................ 2
   2.1 The Arbitrage Pricing Theory............................................................................................................... 2
   2.2 Prior research ....................................................................................................................................... 3
      2.2.1 Economic Forces and the Stock market ......................................................................................... 3
      2.2.2 Arbitrage Pricing Theory in a small open economy .................................................................. 4
      2.2.3 Multivariate regression tests of the Arbitrage Pricing Theory .................................................. 5
   2.3 Choice of stock market index .............................................................................................................. 5
   2.4 Choice of region ................................................................................................................................. 6
   2.5 Choice of macroeconomic factors ..................................................................................................... 7
   2.6 Descriptive statistics ......................................................................................................................... 9
   2.7 Statistical models, significance levels and hypotheses .................................................................... 10

3. Method.................................................................................................................................................. 12
   3.1 Data collection ................................................................................................................................... 12
   3.2 Correction and data processing ....................................................................................................... 12
   3.3 OLS and VAR assumptions ............................................................................................................. 13

4. Result................................................................................................................................................... 15
   4.1 Model 1 .............................................................................................................................................. 15
   4.2 Model 2 .............................................................................................................................................. 16
   4.3 Data reliability ................................................................................................................................... 17
   4.4 Granger-causality test ..................................................................................................................... 17

5. Discussion ............................................................................................................................................ 18

6. Conclusion ............................................................................................................................................ 20

6. Source reference .................................................................................................................................. 21

Appendix 1.............................................................................................................................................. 22

Appendix 2.............................................................................................................................................. 29
1. Introduction

Arguably the most conventional and common method of valuing stocks is to discount the stocks expected future profits. A higher expected future profit corresponds to a higher stock price vice versa. It is thus profits, a microeconomic variable that is central to the most established ways of valuating stocks.

Simply taking expected future profits into consideration when valuing stocks may seem to a method that is too simple to explain movements in stock prices. A firm is not an isolated entity in an economy; rather a firm could be dependent on factors like consumption, inflation and the availability of labor. Therefore it could be possible that market agents, to some degree, take news and predictions about macroeconomic factors like industrial production and unemployment into consideration when valuing a stock.

In 1976 economist Stephen A. Ross published *The Arbitrage Theory of Capital Asset Pricing*. The theory had another angle of attack compared to previous theories regarding valuation of stocks. The theory, which is a general theory of asset pricing, holds that the expected return of a stock can be modeled as a linear function of various macroeconomic factors.

Many studies and applications of the Arbitrage Pricing Theory have been carried out with varying results. The purpose of this study was to examine whether macroeconomic factors from the USA and Eurozone affect the return of the stock market index OMXS30. The knowledge gained from this study could be of interest to those who are financially exposed to a stock market index or individual stocks.

A frequent way of testing the Arbitrage Pricing Theory is to develop and test a statistical model that examine the correlation between a nation's macroeconomic factors and stock market indices, that belong to that same nation. But is it reasonable to think that a nation's stock market index can be influenced by macroeconomic factors from other regions? The economies of the 21st century are, after all very interdependent and often affect each other.

To examine whether that theory was reasonable, the Swedish stock market index OMXS30 and macroeconomic factors from the Eurozone and USA was chosen as test subjects. To test the general hypothesis that macroeconomic factors from the Eurozone and USA affect the return of the OMXS30, two multivariate time series models were developed and tested. Thus the thesis question was; Does macroeconomic factors from the USA and Eurozone affect the return of the stock market index OMXS30?
2. Theory, data and models

The first part of this chapter is a short briefing of the Arbitrage Pricing Theory (APT), which is the theoretical perspective of this thesis. The second part is a review of selected portions of relevant previous research and the third part is a description of the data and statistical models that were used.

2.1 The Arbitrage Pricing Theory

The Arbitrage Pricing Theory is the theoretical standpoint and perspective of this thesis and will be summarized here and more thoroughly in Appendix 1. The theoretical groundwork for APT was provided by Ross (1976) and many have empirically tested the theory since then.

As the name of the theory suggests, APT attempts to explain assets pricing. Assets expected return can, according to the APT, be modeled as linear functions of various macroeconomic factors like industrial productivity and inflation. These variables are captured by betas, risk premiums, which are different in size depending on the asset. The APT holds three crucial assumptions, capital markets are assumed to be perfectly competitive, investors prefer more to less and the stochastic return generating process is a k-factor model;

\[ r_{it} = E_{it} + b_{i1}\delta_{1t} + b_{i2}\delta_{2t} + \ldots + b_{ik}\delta_{kt} + e_{it} \]

For \( i = 1, \ldots, N \)
For \( t = 1, \ldots, T \)

\( N \) = number of assets, \( T \) = number of time periods, \( r_{it} \) = random return of asset \( i \) in time \( t \), \( E_{it} \) = expected return for asset \( i \) in time \( t \), \( b_{ik} \) is the sensitivity of asset \( i \)'s return to movements in common factor \( \delta_{kt} \), and \( e_{it} \) = the mean one idiosyncratic component of the return of asset \( i \) in time \( t \) (Löflund, 1992).

The mean zero random common factors represent the unanticipated changes in macroeconomic factors and the sensitivity coefficients, measure the magnitude and direction of the reaction in asset returns. Investors are viewed as having well diversified portfolios and there exists no arbitrage possibilities. This means that no additional returns can be made by changing the weights of assets in the portfolio. When there are no more arbitrage possibilities, expected return will be a function of the various macroeconomic variables. There will be a
linear relationship between asset expected return and the relevant factors and the betas will capture the influence of the factors.

Equation 2.1: \( r_i = \lambda_0 + \lambda_1 \beta_{i1} + \lambda_2 \beta_{i2} + \ldots + \lambda_k \beta_{ik} \)

As can be seen in Equation 2.1, \( \lambda_0 \) does not have a coefficient. That is because \( \lambda_0 \) is independent of factor movement because \( \lambda_0 \) is seen as the risk free rate \( r_f \). \( \lambda_k \) can be thought of as the risk premium, sensitivity coefficient of factor \( k \). If for example \( \lambda_2 \) is not equal to zero, the stock market rewards risk that is associated with factor 2.

Assets that are positively related to unexpected increases in for example real production, have exposure to that specific macroeconomic factor and thus have higher expected return. Assets with a higher expected return are more valuable since investors want more rather than less. But what if a risk premium is negative? A negative risk premium, \( \lambda_i \) can also be a good thing since they hedge against the adverse influence of an increase in the uncertainty of a factor (Löflund, 1992). It is these \( \lambda_k \), risk premiums that were estimated for the stock market index. To reconnect to the purpose of the essay, it is these risk premiums that empirically was estimated.

2.2 Prior research

2.2.1 Economic Forces and the Stock Market

Chen, Roll and Ross (1986) modeled equity returns as a function of macroeconomic factors on non-equity asset returns (Chen, 1986, 386). The underlying thought behind the paper was that comovements of stocks suggest macroeconomic factors were affecting them. Efficient market theory and rational expectations and intertemporal asset-pricing theory were also assumed.

The macroeconomic factors that was assessed were Industrial production, change in Industrial production, unanticipated inflation, real return off treasury us bill, change in expected inflation, unanticipated changes in risk premium, unanticipated return of long duration bonds, \( EWNY(t) = \text{return of the equally weighted NYSE index} \), \( VWNY(t) = \text{return of the value-weighted NYSE index} \). The time periods analyzed were 1958 to 1984, 1958 to 1967, 1968 to 1977 and 1978 to 1984 (Chen, 1986, 391). To assert whether the macroeconomic factors could explain the asset return, a version of the Fama-Macbeth(1973) approach was used. The assets exposure to the macroeconomic factors was estimated by
regressing asset returns against the unanticipated changes in the variables. The estimated betas were used as explanatory variables in 12 cross-section regressions. The dependent variable in the cross-section regression was the return of the assets. The sum of the risk premiums was estimated.

The conclusion was that “stock returns are exposed to systematic economic news, they are priced according to their exposures.” (Chen, 1986, 402). The most significant variables were Industrial production, changes in risk premium and twists of the yield curve. A striking observation was that the stock market index NYSE explained a significant part of the time-series variability of the stocks, but it has an insignificant influence of the pricing of stocks in comparison with the macroeconomic variables (Ibid, 1986).

2.2.2 Arbitrage Pricing Theory in a small open economy – Empirical evidence from the Swedish stock

The purpose with the research report written by Löflund (1992) was to investigate the fundamental long term causes of stock movement in small open economies. The theoretical perspective was the APT which suggested that asset returns are dependent on different systemic macroeconomic factors. The report provides empirical tests of the APT with the Stockholm stock exchange as test subject (Löflund, 1992).

The macroeconomic variables were monthly inflation in Sweden and the OECD average, monthly growth rate of the industrial production in Sweden and the OECD average, the monthly growth of the USD, DEM and GBP, monthly growth rate of the SEK index, monthly long term government bonds holding period return and the monthly short run money market holding period return (Löflund, 1992).

Two methods were used, the Fama-Macbeth testing method and a simple multivariate test. The periods analyzed were 1977-82, 1978-83, 1979-84, 1980-85, 1981-86, 1982-87, 1983-88. The stock market data consisted of the monthly returns of the Stockholm stock exchange. Prices are adjusted for splits and dividends are assumed to be reinvested. The stocks used had been continuously listed during the period 1977-88 (Löflund, 1992).

“The main conclusion from the APT-pricing tests are that the risk-return relationship are overall weak and that there is considerable lack of robustness in the results. Some of the macroeconomic variables did turn out significant occasionally, but the significance was generally very marginal and very much dependent on the particular portfolio grouping employed” (Löflund, 1992)
As we see the test did not generate such good results. The macroeconomic factors displayed varying signs and sizes, maybe due to low significance of betas factors. The author believes that the choice of macroeconomic factors was unsuccessful.

During the period analyzed, 1977-88, Sweden underwent fundamental structural changes and the economy became more open. This could have had negative effects on the test. The author says that he has not provided a definite answer to the question of whether systematic macroeconomic factors affect the movement of stock returns (Löflund, 1992).

2.2.3 Multivariate Regression Tests of the Arbitrage Pricing Theory

The article published by Wei, Lee and Chen in the year of 1991 expands the theoretical basis of linear estimation of APT with regression analysis. The Regression type used was linear OLS and the null hypothesis was two-sided. The dependent variable was the monthly return “r” of all shares traded on the NYSE during January 1961 through December 1985. The stocks were also divided into small, medium, and large size. The regression coefficients would prove to be highly correlated with each other, which resulted in insignificant estimates of the coefficients.

2.3 Choice of stock market index

OMXS30 is a stock market index, denominated in Swedish Kronor, for the Stockholm Stock exchange(SSE) and is a market value-weighted index consisting of the 30 most traded firms on the SSE. The return of the OMXS30 was the dependent variable of my two multiple time series regression models. Since the nominal value of index only shows the value-weighted current price of the 30 firms, I transformed the data to show return instead. Se section 3.2.

OMXS30 is interesting since it is an index of the most traded stocks on the SSE. And since the value of a stock is the present value of future earnings, OMXS30 says something about the economy as a whole.

OMXS30 has a couple of interesting features. The data spans from the 30th of September 1986 till today which is a long time. That is of course good when preforming a statistical analysis. OMXS30 is fairly diversified, consisting of 30 stocks. The industry breakdown in 2013 is as follows; Basic Materials 1.32%, Consumer goods 6.87%, Consumer Services
11.38%, Financial 29.58%, Health Care 3.34%, industrial 28.95%, Oil & Gas 1.66%, Technology 8.37% and Telecommunications 8.07% (OMXNORDIC - OMXS30, 2013)

Another good thing with the OMXS30 is that there are no small firms in it. The stock prices of small firms have a tendency to be more volatile than larger firms. The accessibility of the data is very good and it is available online.

2.4 Choice of region

The two regions, whom macroeconomic variables are to be analyzed, are the Eurozone and the USA. The Eurozone differs from the US since it is a construction of nations that have a higher degree of independence compared to the individual states in the US.

There are some interesting similarities between the USA and the Eurozone. The two regions have their own currency, the Dollar in the US and the Euro in the Eurozone. The monetary policy in the US is administered by the Federal Reserve System and by the European central bank in the Eurozone. The two regions are large, open developed economies.

The Eurozone is a political, geographical and economic region, today consisting of 17 nation members of the European union, that are full members of the European monetary union and have the Euro as there national currency. The current 17 member nations did not join the Eurozone at the same time. The nations that joined the Eurozone in 1999 were Austria, Belgium, Finland, France, Germany, Ireland, Italy, Luxembourg, Portugal, Netherlands and Spain. Greece entered 2001, Slovenia 2007, Cyprus and Malta 2008, Slovakia 2009 and Estonia 2011 (Eurozone Portal, 2013).

The GDP of the Eurozone was 12.460 billion US dollars(USD) in 2009 according to the World Bank and it is the third largest economic entity behind the European union and the USA (World Bank - GDP, 2013).

The trade between Sweden and the Eurozone is considerable. During the time period from January to February 2013 Sweden exported goods of a value of 69.8 billion SEK(Swedish kronor) to the Eurozone. That is equivalent to 38.8 % of Sweden’s export. Sweden imported goods equivalent to 79.5 billion SEK which is 47.5% of the goods that Sweden import (SCB, 2013). Regions that use the Euro, but are not part of the European Monetary Union like the Vatican state and Kosovo will not be taken into consideration.

The United States of America has the largest economy in the world, a GDP equivalent to 14.219 billion USD in 2009 (“World Bank - GDP,” 2013). Sweden exported good equivalent to a value of 12.895 billion SEK, 7.2% of total export and imported good of a value of 5.085
billion which is equivalent to 3% of Sweden’s total imports. (“SCB Export och import,” 2013) Regions that use the dollar, but are not directly part of the USA will not be taken into consideration.

Due to the size of the economies of the Eurozone and the USA, the trade between the two regions and Sweden, it is reasonable that macroeconomic factors of these two regions may affect the Swedish economy. It is also reasonable that foreign and domestic agents acting on the Stockholm stock exchange take into account news about macroeconomic factors from the USA and the Eurozone.

2.5 Choice of macroeconomic factors

When I chose the macroeconomic factors, whose data was used in the statistical models, I assumed certain criteria. The first criterion in the choice of macroeconomic factors was that they must have been successful in the previous researches which I reviewed. The second criterion was that agents on the Stockholm Stock Exchange would consider that the factors were relevant.

The following ten macroeconomic factors were chosen to be used in the statistical estimation; Industrial production, CPI, 10 year bond yield and unemployment in the Eurozone and the USA. Also industrial confidence in the Eurozone and consumer confidence in the USA was chosen.

Industrial production is a measure of real, inflation adjusted, production and is expressed as a percentage of real output of a base year. It is often used as a measure of the economy’s wellbeing. In the Eurozone, industrial production is measured by Eurostat and by the Federal Reserve in the US.

As mentioned above (see section 2.2.1), in Economic Forces and the Stock Market by Chen, Industrial production was used as a macroeconomic factor. According to Chen, changes in expectations of industrial production could affect firms profits and hence the assets value. There are some differences between the US and the Eurozone Industrial production index. Both are seasonally adjusted but the US has its index in 2007=100 and the Eurozone has its index in 2010.

The macroeconomic factor Industrial confidence is an industrial survey carried out by the European commission. The survey is seasonally adjusted and measures confidence in the industry by answering questions about production expectations, order books and stocks of finished.
Industrial confidence is interesting since investors may interpret changes in industrial confidence as changes in future profit. An increase in industrial confidence may be seen as a good sign and an indication that production, consumption and hence profit will be higher.

The macroeconomic factor consumer confidence is interesting because it can reflect consumption and therefore possibly profits. The consumer confidence in the USA is measured by the Conference board which is an independent organization and the index is based on the survey answered by 5 000 US households. Each month, the households answer survey questions, for example how they feel about current business conditions, business conditions for the next six months, current employment conditions, employment conditions for the next six months and total family income for the next six months. Each question is answered as positive, negative or neutral.

The macroeconomic factor Consumer Price Index, CPI was used as a measure of inflation. The CPI is an index that reflects changes in the prices of specific goods and services that an average household might buy. The “basket” of goods might consist of good like clothing, shelter, fuel, food, transportation costs and it is the average change of these prices that constitutes the CPI. A change in the index form year to year is viewed as a change of the general price level. For example, an increase in the index between the years 2001 and 2002 is interpreted as inflation.

The CPI of the Eurozone is measured by Eurostat and the CPI of the USA is measured by the US department of labor.

Rising unemployment is often associated with a deteriorating economic climate, there for unemployment in the USA and Eurozone was seen as an interesting macroeconomic factor. A rise in unemployment may cause investors to become more pessimistic about the general economy and profits since a higher rate of unemployment could be interpreted as fewer consumers, less turnover and lower profits. This may decrease market agents’ demand for assets on the stock market, causing asset prices to go down resulting in lower return.

Unemployment in the Eurozone is measured by the ECB whilst unemployment in the US is measured by the Bureau of Labor Statistics.

Bonds are often viewed as a low-risk alternative to stocks. Generally stocks are more risky than government bond and when investors think that stocks are too risky, they often turn to the bond market. An increase in the demand for bonds lowers the yield vice versa. The bond yield chosen as macroeconomic factor is the 10 year government bond yield for the USA and 10 year government bond yield of Eurozone governments.
2.6 Descriptive statistics

Section 2.6 includes three graphs and tables that describe the data that was used when the models were estimated.

**Plot 2.6a Time-series**

Plot 2.6a displays time-series graphs of the 51 observations of the ten macroeconomic factors and the dependent variable. Due to lack of space in the graphs, the following abbreviations were used; OMXS30 return (r), Eurozone Industrial production (IndProdEZ), U.S Industrial production (IndProdUS), Eurozone industrial confidence (IndConfEZ) U.S Consumer confidence (ConConfUS), Eurozone unemployment (UnEmpEZ), the U.S. unemployment (UnEmpUS), Eurozone CPI (CPIez), U.S CPI (CPIus), Eurozone 10 year bond (Bond10ez) and U.S. 10 year bond (Bond10us).

**Table 2.6b Descriptive statistics**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>IndProdEZ</td>
<td>51</td>
<td>103.7</td>
<td>3.6</td>
<td>98.4</td>
<td>111.1</td>
</tr>
<tr>
<td>IndProdUS</td>
<td>51</td>
<td>94.4</td>
<td>3.1</td>
<td>89.8</td>
<td>100.0</td>
</tr>
<tr>
<td>IndConfEZ</td>
<td>51</td>
<td>-3.1</td>
<td>5.9</td>
<td>-13.2</td>
<td>7.8</td>
</tr>
<tr>
<td>ConConfUS</td>
<td>51</td>
<td>96.5</td>
<td>11.9</td>
<td>61.4</td>
<td>111.2</td>
</tr>
<tr>
<td>UnEmpEZ</td>
<td>51</td>
<td>8.9</td>
<td>0.5</td>
<td>7.6</td>
<td>9.4</td>
</tr>
<tr>
<td>UnEmpUS</td>
<td>51</td>
<td>5.3</td>
<td>0.6</td>
<td>4.4</td>
<td>6.3</td>
</tr>
<tr>
<td>CPIez</td>
<td>51</td>
<td>99.3</td>
<td>2.7</td>
<td>94.7</td>
<td>104.2</td>
</tr>
<tr>
<td>CPIus</td>
<td>51</td>
<td>193.4</td>
<td>7.4</td>
<td>182.6</td>
<td>205.9</td>
</tr>
<tr>
<td>Bond10ez</td>
<td>51</td>
<td>99.7</td>
<td>3.1</td>
<td>95.1</td>
<td>106.3</td>
</tr>
<tr>
<td>Bond10us</td>
<td>51</td>
<td>121.8</td>
<td>4.3</td>
<td>113.9</td>
<td>134.5</td>
</tr>
<tr>
<td>r</td>
<td>51</td>
<td>1.8</td>
<td>4.0</td>
<td>-9.2</td>
<td>9.5</td>
</tr>
</tbody>
</table>
Table 2.6b describes the mean, standard deviation, minimum and maximum value for each variable. The figures are rounded to one decimal place.

Table **2.6c Correlation matrix**

<table>
<thead>
<tr>
<th>Variable</th>
<th>IndProd EZ</th>
<th>IndProd US</th>
<th>IndConf EZ</th>
<th>IndConf US</th>
<th>UnEmp EZ</th>
<th>UnEmp US</th>
<th>CPI EZ</th>
<th>CPI US</th>
<th>Bond10 EZ</th>
<th>Bond10 US</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>IndProd EZ</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndProd US</td>
<td>0.85</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndConf EZ</td>
<td>0.64</td>
<td>0.57</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndConf US</td>
<td>0.03</td>
<td>0.04</td>
<td>0.53</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UnEmp EZ</td>
<td>-0.5</td>
<td>-0.1</td>
<td>-0.3</td>
<td>-0.36</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UnEmp US</td>
<td>-0.3</td>
<td>-0.23</td>
<td>-0.46</td>
<td>-0.84</td>
<td>0.64</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPI EZ</td>
<td>0.26</td>
<td>0.41</td>
<td>0.17</td>
<td>-0.8</td>
<td>0.43</td>
<td>0.75</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPI US</td>
<td>0.29</td>
<td>0.42</td>
<td>-0.15</td>
<td>-0.8</td>
<td>0.43</td>
<td>0.73</td>
<td>0.99</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond10 EZ</td>
<td>-0.05</td>
<td>0.2851</td>
<td>-0.27</td>
<td>-0.64</td>
<td>0.79</td>
<td>0.70</td>
<td>0.84</td>
<td>0.82</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond10 US</td>
<td>-0.12</td>
<td>0.15</td>
<td>-0.41</td>
<td>-0.74</td>
<td>0.74</td>
<td>0.75</td>
<td>0.81</td>
<td>0.78</td>
<td>0.94</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>-0.2</td>
<td>-0.07</td>
<td>0.07</td>
<td>0.5</td>
<td>0.02</td>
<td>-0.5</td>
<td>-0.01</td>
<td>0.02</td>
<td>0.09</td>
<td>0.03</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2.6c is a correlation matrix of all the variables. Some explanatory variables, such as U.S. inflation and Eurozone inflation exhibits high correlation. High correlation between the explanatory variables may impair the statistical inference due to multicollinearity. The explanatory variables that showed the strongest correlation with the dependent variable was Eurozone industrial production and U.S. consumer confidence.

### 2.7 Statistical models, significance levels and hypotheses

Section 2.7 includes a description of the statistical models, significance levels and hypotheses that was used.
Model 1 – The full model

\[ \bar{r}_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \beta_6 x_{i6} + \beta_7 x_{i7} + \beta_8 x_{i8} + \ldots + \beta_{10} x_{i10} + \varepsilon_i \]

\( \bar{r}_i \) is the \( i \):th observation of the dependent variable \( r \) (return of the OMXS30). Model 1 is a multiple time series model and the purpose of the beta coefficients is to estimate the so-called risk premium \( \lambda_k \) for each macroeconomic factor and thus measure how much they affect returns. All macroeconomic factors (See section 2.6) are included as explanatory variables in Model 1.

Model 2 – The reduced model

\[ \bar{r}_i = \beta_0 + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_6 x_{i6} + \beta_9 x_{i9} + \beta_{10} x_{i10} + \varepsilon_i \]

Model 2 is also a multiple times series model and it is a reduced form of Model 1. \( \bar{r}_i \) is the \( i \):th observation of the dependent variable \( r \) (return). The reason why a model with fewer macroeconomic factors was developed was due to the high risk of multicollinearity. By picking the most relevant macroeconomic factors, dough fewer, the risk of insufficient statistical significance due to multicollinearity was reduced. At the same time this might cause problems due to omitted variables. The macroeconomic factors that were included in Model 2 were Eurozone Industrial confidence, U.S Consumer confidence, U.S CPI, Eurozone and U.S 10 year bonds.

Significance levels and hypotheses

Two types of hypothesis tests were performed for the two models, an F-test and T-tests for each explanatory variable. The minimum significance level that the models and coefficients needed to fulfill to be accepted was a level of 10%

A 10% level of significance is considered to be quite high, but it was used since large levels of multicollinearity were expected due to endogenous nature of macroeconomic factors which OLS is sensitive to (Sims, 1980). Even dough a 10% level of significance was the minimal level, I also determined if the variables were significant up to a 5% level of significance. Since it was difficult to determine whether the explanatory variables would adopt a positive or negative value, a two-sided hypothesis was used for the T-test.
Table 2.7a Hypothesis one F-test

Hypothesis for the F-test used for Model 1 and Model 2. This type of hypothesis tests whether all explanatory variables equal to zero or not.

H₀: All explanatory variables coefficients are equal to zero
Hₐ: At least one of the coefficient does not equal to zero

Table 2.7b Hypothesis two T-tests

The statistical significance of the coefficients of the explanatory variables were tested with a two-sided hypothesis tests using the T-distribution and the attained p-values was compared with the chosen minimum level of significance.

H₀: βᵢ = 0
Hₐ: βᵢ ≠ 0

i = 1…10

3. Method

3.1 Data collection

The data was collected and downloaded via Excel and Datastream. A Monthly data interval was selected since macroeconomic data is usually published once a month. The first observation of the dependent variable was the “closing price” value of OMXS30 on the 16th of January 2003 and last observation from 16th June 2007, a total of 51 observations. The first observation of the explanatory variables was from the 15th January 2003 and the last observation was from the 15th of June 2007. Note that OMXS30 observations are lagged one day. The reason for the lag was to let the market internalize the information about the macroeconomic factors.

3.2 Correction and data processing

The collected data was controlled to ensure that there were no obvious errors. No critical errors were found in the data. The dependent variable OMXS30 had to be processed since the APT model asserts that return is dependent on macroeconomic factors to some degree.

Equation 3.2 \( \bar{r}_i = a_i + b_{i1}F_1 + b_{i2}F_2 + ... + b_{ik}F_k + e_i \)
To achieve this, the following process of the dependent variable was executed.

\[
\frac{(OMXS30_t - OMXS30_{t-1})}{OMXS30_{t-1}} = \bar{r}_t
\]

The result is that \(\bar{r}_t\) is equal to the \(i:\text{th}\) return of the stock market index OMXS30. The explanatory variables were not processed since that would not correspond to the APT model.

### 3.3 OLS and VAR assumptions

APT implies that "When all arbitrage possibilities have been executed there will exist an Approximately linear relationship between expected return and sensitivity to the factor." To meet that criterion, a linear rather than a nonlinear regression was used. The regression method used for Model 1 and Model 2 was ordinary least squares, OLS. OLS is a "minimum variance estimator" if the following classical assumptions are fulfilled (Studenmund, 2010).

The reason why these assumptions are illustrated here in section 3.3 is because the two statistical models were tested against these assumptions in section 4.3 of the thesis, with the purpose of ensuring the models' quality and reliability.

#### Assumption 1) The regression model is linear, is correctly specified and has an additive error term. If assumption 1 were to be violated, which is not unlikely, the model would not reflect the population perfectly. To ensure that the model is as correct as possible, the explanatory variables were chosen only if they were reasonable and theoretically sound. To counteract the wrong variables from being chosen, scatterplots, residual plots and hypothesis tests was used.

#### Assumption 2) The error term has a zero population mean. To ensure that the error term is centered around zero, analysis of the residual plots was made.

#### Assumption 3) All explanatory variables are uncorrelated with the error term. Violation of assumption 3 can cause unbiased estimation of coefficients.

#### Assumption 4) Observations of the error term are uncorrelated with each other. Violation of assumption 4 can cause serial correlation and unbiased standard errors. Breusch-Godfrey test for autocorrelation was used to test for serial correlation.

#### Assumption 5) The error term has a constant variance. Violation of assumption 5 can cause heteroscedasticity and unbiased standard errors. Breusch-Godfrey test for heteroscedasticity was used to test for heteroscedasticity.
**Assumption 6)** No explanatory variable is a perfect linear function of any other explanatory variable. Violation of assumption 6 can lead to multicollinearity. Sever multicollinearity can cause an increase in the variance of the estimated regression coefficients. VIF test was used to test for multicollinearity.

**Assumption 7)** The error term is normally distributed. Violation of assumption 7 can cause unreliable hypothesis testing. A normal probability plot was used to test the error terms distribution.

**Vector autoregressive models** (VAR models), which was used when performing the Granger-causality test, are a bit more advanced than ordinary regression models and the main difference is that all the variables in a VAR model are endogenous. In the VAR models, the current values of each endogenous variable depend on different combinations of the previous $k$ values of all variables and error terms. The model was advocated by Sims(1980) and is theory-free for estimating economic relationships. VAR models contain endogenous variables and the variables have strong relationship to themselves and their lags.

\[
y_{1t} = \beta_{10} + \beta_{11}y_{1t-1} + \ldots + \beta_{1k}y_{1t-k} + \alpha_{11}y_{2t-1} + \ldots + \alpha_{1k}y_{2t-k} + u_{1t}
\]
\[
y_{2t} = \beta_{20} + \beta_{21}y_{2t-1} + \ldots + \beta_{2k}y_{2t-k} + \alpha_{21}y_{1t-1} + \ldots + \alpha_{1k}y_{1t-k} + u_{2t}
\]

Where $u_{it}$ is an error term with $E[u_{it}]=0 \ (i=1,2)$, $E[u_{1t}, u_{2t}]=0$

The VAR model could be written as following for $g$ variables

\[
y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \ldots + \beta_k y_{t-k} + u_t
\]

Where $y_t, y_{t-1} + \ldots + y_{t-k}, \beta_0$ and $u_t$ are vectors $g \times 1$

$\beta_1, \beta_2 + \ldots + \beta$ are $g \times g$ vectors.

In order to answer the thesis question I had to examine the correlation between the dependent variable and the explanatory variables, but also causation. Since ordinary regression(OLS) only shows mere correlation, a Granger-causality test was implemented to examine if it was the macroeconomic factors that were affecting the return or vice versa. This is known as an exogeneity issue. A granger-causality test is performed after estimating the VAR model.
4. Result

Chapter 4 Results contains the results of the estimation of the two statistical models. The estimation of Model 1 and 2 were made using Minitab and the coefficients are rounded to two decimal. The Granger causality test was executed using STATA. The reliability of the models is presented in section 4.3.

4.1 Model 1

Section 4.1 presents the results from the estimation of Model 1 using linear regression.

Table 4.1a

<table>
<thead>
<tr>
<th>Variable</th>
<th>Region</th>
<th>Coefficient (95% CI)</th>
<th>SE coefficient</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td></td>
<td>18.3 (337,3;373.3)</td>
<td>177.8</td>
<td>0.10</td>
<td>0.919</td>
</tr>
<tr>
<td>Industrial production**</td>
<td>Eurozone</td>
<td>-1.71 (-3.54;0.12)</td>
<td>0.9146</td>
<td>-1.87</td>
<td>0.069</td>
</tr>
<tr>
<td>Industrial production</td>
<td>USA</td>
<td>-0.80 (-2.91;1.29)</td>
<td>1.051</td>
<td>-0.77</td>
<td>0.446</td>
</tr>
<tr>
<td>Industrial confidence</td>
<td>Eurozone</td>
<td>-0.34 (-1.07;0.4)</td>
<td>0.3663</td>
<td>-0.93</td>
<td>0.360</td>
</tr>
<tr>
<td>Consumer confidence*</td>
<td>USA</td>
<td>0.22 (0.02;0.41)</td>
<td>0.09937</td>
<td>2.20</td>
<td>0.036</td>
</tr>
<tr>
<td>Unemployment</td>
<td>Eurozone</td>
<td>-4.2 (-10.83;2.30)</td>
<td>3.283</td>
<td>-1.30</td>
<td>0.201</td>
</tr>
<tr>
<td>Unemployment</td>
<td>USA</td>
<td>0.53 (-11.47;12.53)</td>
<td>6.070</td>
<td>0.09</td>
<td>0.931</td>
</tr>
<tr>
<td>CPI**</td>
<td>Eurozone</td>
<td>-0.67 (-4.27;2.92)</td>
<td>1.799</td>
<td>-0.37</td>
<td>0.710</td>
</tr>
<tr>
<td>10 year bond</td>
<td>USA</td>
<td>0.66 (-0.03;1.35)</td>
<td>0.3467</td>
<td>1.90</td>
<td>0.065</td>
</tr>
<tr>
<td>10 year bond**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What can be seen in Table 4.1a is the estimation of Model 1 using linear regression with standardized betas. The coefficients that are significant up to a 5% level are highlighted in green* and those that are significant at a 10% level are highlighted in yellow**. As can be seen, not all variables were significant at neither 5% nor a 10% level. Only four in ten null hypotheses (See Table 2.8 b) could be rejected at a 5% or 10% significance level. The four significant variables were Eurozone industrial production (p=0.069), U.S. consumer confidence (p=0.036), U.S. CPI (p=0.057) and U.S. 10 year government bonds (p=0.065).

Table 4.1b

ANOVA and R values

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>10</td>
<td>235.89</td>
<td>23.59</td>
<td>1.53</td>
<td>0.163</td>
</tr>
<tr>
<td>Residual Error</td>
<td>41</td>
<td>631.68</td>
<td>15.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>51</td>
<td>867.57</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Note that the null hypothesis, that all coefficients are equal to zero (see Table 2.8a), cannot be rejected at a 10% nor a 5% level because they exceeded $P = 0.10$.

27.2% of the variation of the dependent variable (return of the OMXS30) could be explained by the variation of the explanatory variables and 9.4% could be explained when adjusted for the number of explanatory variables.

### 4.2 Model 2

Section 4.2 presents the results from the estimation of Model 2 using linear regression.

#### Table 4.2a

<table>
<thead>
<tr>
<th>Variable</th>
<th>Region</th>
<th>Coefficient (95% CI)</th>
<th>SE coefficient</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant*</td>
<td></td>
<td>-159 (-274.6; -43.4)</td>
<td>59.79</td>
<td>-2.66</td>
<td>0.011</td>
</tr>
<tr>
<td>Industrial confidence**</td>
<td>Eurozone</td>
<td>-0.60 (-1.24; 0.05)</td>
<td>0.3192</td>
<td>-1.87</td>
<td>0.068</td>
</tr>
<tr>
<td>Consumer confidence**</td>
<td>USA</td>
<td>0.13 (-0.02; 0.28)</td>
<td>0.07671</td>
<td>1.65</td>
<td>0.098</td>
</tr>
<tr>
<td>CPI*</td>
<td>USA</td>
<td>0.83 (0.16; 1.50)</td>
<td>0.3339</td>
<td>2.49</td>
<td>0.018</td>
</tr>
<tr>
<td>10 year bond*</td>
<td>Eurozone</td>
<td>-1.04 (-1.88; -0.20)</td>
<td>0.4232</td>
<td>-2.46</td>
<td>0.018</td>
</tr>
<tr>
<td>10 year bond*</td>
<td>USA</td>
<td>0.74 (0.18; 1.29)</td>
<td>0.2796</td>
<td>2.65</td>
<td>0.011</td>
</tr>
</tbody>
</table>

As can be seen in Table 4.2a, all null hypotheses (See Table 2.8b) belonging to the explanatory variables of Model 2 could be rejected at a 10% significance level and 3 out of 5 could be rejected at a 5% level. Model 2 was significant up to a 10% level. Note that the Eurozone coefficients were all negative and the US coefficients were all positive. To put this in perspective and to relate to the APT, we can say that the estimated coefficients of Model 1 and Model 2 represent the risk premiums of that specific macroeconomic factor (see Appendix 1). For example if we look at an estimated coefficient in Model 2, an increase of one unite in the macroeconomic factor “10 year bond” corresponds to an increase of 0.7406% in the return of the OMXS30.

#### Table 4.2b

**ANOVA and R**

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>5</td>
<td>151.48</td>
<td>30.30</td>
<td>1.98</td>
<td>0.096**</td>
</tr>
<tr>
<td>Residual Error</td>
<td>46</td>
<td>716.10</td>
<td>14.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>51</td>
<td>867.57</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
As can be seen in Table 4.2b, 17.5 % of the variation in the dependent variable \( r \) could be explained by the variation in the explanatory variables and 8.5% could be explained when adjusted to the number of explanatory variables. The null hypothesis (See table 2.8a) that all explanatory variables equal to zero could also be rejected at a 10% level of significance.

### 4.3 Data reliability

<table>
<thead>
<tr>
<th>Data reliability</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serial correlation</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Severe multicollinearity</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Heteroskedasticity</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Non-normally distributed residual</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>

In Table 4.3 we see that both Model 1 and Model 2 suffered from severe multicollinearity (VIF>10). Especially Model 1 had a large amount of multicollinearity. Not so surprisingly because the variables may be affecting each other to a large degree i.e they may be endogenous. For greater detail see Appendix 2

### 4.4 Granger-causality test

The purpose of the thesis was to answer whether macroeconomic factors from the USA and Eurozone are affecting the return of the OMXS30. As can be seen in section 4.1 and 4.2 two multiple time series models was estimated, one of which shows a level of significance for all variables of at least 10%, and many variables showing an individual significance level under 5% indicating that there is at least significant correlation between the variables.

Although common sense would argue that the macroeconomic factors are affecting the stock market index and not the other way around, we would also like to perform a statistical analysis and test the causality of the models. The Granger-causality test tests for “predictive-causality”, to what degree the explanatory variables can predict the movements of the dependent variables.

Using vector autoregression, VAR(1), and the variables of Model 2, a Granger-causality test was made. The null hypothesis is no Granger-causation.
Assuming a 5% level of significance, the null hypothesis of no Granger-causation for all explanatory variables could not be rejected. Therefore I could not say that all explanatory variables of Model 2 are Granger-causal with the dependent variable. The null hypothesis of no Granger-causality could be rejected for the variables U.S. CPI (p=0.05) and U.S. Bond (p=0.045). See Appendix 2c for greater detail.

5. Discussion

The question of this thesis is, Does macroeconomic factors from the USA and Eurozone affect the return of the stock market index OMXS30? To answer the question I developed two multiple time series models which was estimated using OLS. Two types of null hypotheses was used for each model. The first null hypothesis (See table 2.8a) was that all coefficients are equal to zero and the second hypothesis, which was tested against all explanatory variables, was that the coefficient is equal to zero. I chose two levels of significance, 5% and 10% which was the minimum level of significance.

The null hypothesis that all explanatory variables are equal to zero could not be rejected for Model 1 at a 10% significance level, but it could be rejected for Model 2. The second null hypothesis that the coefficient of the explanatory variable is equal to zero (See Table 2.8b) could be rejected in four out of ten cases in Model 1 and in all cases in Model 2. In three out of five cases, the hypothesis that the explanatory variable is equal to zero in Model 2 could be rejected at a 5% level of significance.

I think that the reason why the first null hypothesis in Model 1 and the null hypothesis of several of the explanatory variables were not significant up to a 10% level was due to severe multicollinearity. As can be seen in Appendix 2b, seven out of 10 explanatory variables in Model 1 suffered from severe multicollinearity. High amounts of multicollinearity can, as
described in section 3.3, cause unreliable estimation of the coefficients which can cause unreliable statistical inference and hypothesis testing.

I also think that the potential sever multicollinearity in both Model 1 and 2, which may have caused unreliable estimates of the coefficients, may have caused low R-values, the degree of which the explanatory variables can explain the variance in the dependent variable r(return of the OMXS30). The R-value of Model 1 was 27.2 % and 9.4 % adjusted for the number of explanatory variables. The R-value of Model 2 was 17.5 % and 8.5% adjusted for the number of explanatory variables. The R-values are not large, but certainly larger than zero. Besides sever multicollinearity, all other assumptions(See section 3.3) were fulfilled in Model 1 and Model 2.

The results from Model 2 showed that there certainly is significant correlation between the dependent and explanatory variables. But is there causation? The question of causation is relevant in the thesis since my thesis question asks if the macroeconomic variables affect the return of the OMXS30 and not the other way around. Though statistics cannot prove causation, it can prove granger-causation. Using a vector autoregression version of Model 2, I used STATA to answer the question of whether the explanatory variables Granger-cause with the dependent variable. I was not able to reject the null hypothesis (See Appendix 2C) that there is no Granger-causality for three out of five variables. I suspect this is due to estimation error. Economic theory suggests that the macroeconomic variables that I have chosen affect the index rather than the outer way around.

To answer my thesis question, Does macroeconomic factors from the USA and Eurozone affect the return of the stock market index OMXS30? judging from the results I can say that it is plausible that macroeconomic factors from USA and Eurozone affect the return of the stock market index OMXS30. Although macroeconomic factors may affect return I do not think it is the only thing that affects return. In the short run, psychological factors may play a greater role in determining stock prices.

Chen (1991) used OLS to estimate statistical models of APT. This method of estimation is unfortunately sensitive to multicollinearity which can lead to unreliable estimations. This could be a problem since my variables often showed large amounts of multicollinearity and the fact that macroeconomic factors often are endogenous (Sims 1980)

My dataset consisted of 51 observations during a specific time period. Had I chosen another time period, the statistical estimation would most certainly be different. This is due to the fact that the macroeconomic factors and return would probably have different values in another time period. Especially since macroeconomic factors often are cyclical, a period can be
characterized by negative results during some period and positive results during others. That fact may cause problems when using OLS and can cause unreliable statistical inference.

Even though the macroeconomic factors that I have analyzed are sound from an economic stand point and most of them have been used in previous research, it does not guarantee that they are the best at explaining return. It should also be noted that I did not use indicator variables, as some researchers have done, since I do not think that corresponds to the Arbitrage Pricing Theory. It could also be possible that different stock market index react differently to macroeconomic factors. Researchers like Chen (1986) have created their own indexes, bundles of stock which the they have tested APT on.

The Arbitrage Pricing Theory assumes that there are no arbitrage possibilities. For that to be possible, full information about the stock market must be available. Although it may be possible in theory, I think that it is unrealistic and that there may be arbitrage possibilities.

Chen (1991) who, like me used OLS to estimate APT also got coefficients that were highly correlated, which also might cause unreliable statistical inference (See section 2.1.3). Multicollineairity seems to be a recurring problem when using OLS to estimate APT, so my recommendation for future research in the field of empirical estimation of APT is to use estimation methods which are less sensitive to multicollinearity.

6. Conclusion

I have developed two multiple time series models with the purpose of investigating whether or not macroeconomic factors from the USA and Eurozone affected the return of the stock market index OMXS30. Model 1 had all the explanatory variables that I considered relevant and Model 2 was a reduced form of the first model and hade what I deemed the most relevant factors. My theoretical perspective when I developed the statistical models was the Arbitrage Pricing Theory, which states that the return of assets is a linear function of relevant macroeconomic factors.

The first model, Model 1, was not found to be sufficiently significant up to a 10% level. This was probably due to high multicollinearity. The reduced model, Model 2, would prove to be significant at a 10 % level and lower (see Table 2.8a). In both models, many explanatory variables individually would prove to be significant at a 5 % level.

A Granger-causality test was also executed and the null hypothesis that the variables are not Granger-causal could not be rejected for every explanatory variable. In summary, it is
reasonable to say that macroeconomic factors from the U.S. and the Eurozone affected the return of OMXS30 during the selected time period.

7. Source reference


Eurozone Portal [WWW Document], 2013. . Euroe Area - Eurozone. URL eurozone.europa.eu


SCB Export och import [WWW Document], 2013. URL http://www.scb.se/Pages/TableAndChart____142265.aspx


Studenmund, A.., 2010. "Using Econometrics".

Christopher A. Sims, 1980, "Macroeconomics and Reality", Econometrica 48


APPENDIX

Appendix 1: Arbitrage Pricing Theory

The first assumption of APT is that security returns are related to an unknown number of factors. Let’s start by looking at a simple one-factor model. In this case the factor is aggregate consumption, a macroeconomic variable. In this situation the securities returns are related to the size of aggregate consumption. In this case the equation for the model looks like this:
$$r_i = a_i + b_i F_i + e_i$$

$r_i$ = rate of return of security $i$

$a_i$ = Zero factor

$F_i$ = The value of the factor. In this example its aggregate consumption

$e_i$ = random error term

$b_i$ is the sensitivity of security $i$ to the factor. (Sharpe, 1999)

Let’s continue with a one-factor model but now we have a portfolio of three different stocks. In this example we have an invested wealth of $w_0 = \$12\,000\,000$ and $\$4\,000\,000$ invested in each stock. Another assumption that must be taken into consideration is that all agents on the market agree on the expected returns and sensitivities. In this example it is the following:

<table>
<thead>
<tr>
<th>Stock</th>
<th>Expected return</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stock 1</td>
<td>15%</td>
<td>0.9</td>
</tr>
<tr>
<td>Stock 2</td>
<td>21%</td>
<td>3.0</td>
</tr>
<tr>
<td>Stock 3</td>
<td>12%</td>
<td>1.8</td>
</tr>
</tbody>
</table>

**Table 1 appendix**

1.1 Arbitrage

Before continuing on to arbitrage portfolios, an important part of APT, the principals of arbitrage has to be explained. As an example, imagine that you are at a marketplace, say the annual Nolia fair in Umeå. Walking through the fair you discover that a buyer is willing to buy one kg of reindeer meet for $20. On the other side of the fair you discover that a meet vendor is selling reindeer meet for $18/kg.

Now you realize that you could make a small profit by buying the meet for $18 and selling it for $20. The buyer agrees to buy the meet for $20, gives you the money and you go and buy the meet for $18/kg and then give the meet to the buyer. By doing this you have realized a profit. In theory arbitrage should be risk free and the product or asset traded should be the same. It is the process of taking advantage of difference in prices for the same asset. (Sharpe, 1999)
Arbitrage is an important activity on the modern, efficient market for securities and other assets. Since arbitrage is by definition risk free, investors have incentive to do this and by doing this they can eliminate price differences.

When discussing the difference in price of a specific asset, arbitrage is straightforward. Arbitrage opportunities involving similar securities or portfolios are called “almost arbitrage”. Instead of price differences, exposure to a factor could be a reason to use arbitrage. The idea with a factor model is that assets and portfolios with the same level of factor sensitivity should have the same expected return, except return related to the nonfactor risk.

But what if they do not? If assets or portfolios have the same sensitivity to a factor, but different expected return, there exists an “almost arbitrage” opportunity. If investors have a profit motive and have the right knowledge and information, they will take advantage of these opportunities, causing the elimination of the difference in expected return (Sharpe, 1999). This process is crucial for the APT.

1.2 Arbitrage portfolios

According to the Arbitrage pricing theory, in order to increase expected return, investors are looking for possibilities to form an arbitrage portfolio of their current portfolios without taking on more risk. What is an arbitrage portfolio? An arbitrage portfolio is a modification to the current portfolio, does not require any additional funds from the investor and the result should be a positive increase in the expected return of the portfolio.

As an example, let $X_i$ denote the change in the investors portfolio of security $i$ and hence the weight of security $i$ in the arbitrage portfolio. Remember the requirement that no additional fund should invested in the portfolio, only changes in the weights of the securities $X_1 + X_2 + X_3 = 0$.

Another important assumption is that the arbitrage portfolio has no sensitivity to any factor, zero factor exposures. Portfolios sensitivity to a factor is the weighted average of the sensitivities of the assets in the portfolio. Since the arbitrage portfolio should have zero factor exposure, the sensitivity of the arbitrage portfolio can be shown like this: $b_1X_1 + b_2X_2 + B_3X_2 = 0$ and in our example $0.9X_1 + 3.0X_2 + 1.8X_3 = 0$

As we see, the weights($X_i$) must be made so that the sum of the three values equals zero. Thus in this example, the arbitrage portfolio has no sensitivity to aggregate consumption. The arbitrage portfolio also should no nonfactor risk, however the Arbitrage Pricing Theory assumes that the nonfactor risk is so small that it can be ignored.
We are now going to construct the arbitrage portfolio and there are lots of combinations to choose from. The condition that has to be meet is that the sensitivity of the arbitrage portfolio should be zero and we arbitrarily choose $x_1 = 0.1$

$$0.1 + X_2 + X_3 = 0$$
$$0.09 + 3.0X_2 + 1.8X_3 = 0$$

If $X_1 = 0.1$ and $b_1 = 0.9$ the $X_1b_1=0.09$ and $X_2 = 0.075$ and $X_3 = -0.175$

We have now realized a potential arbitrage portfolio, with the weights $X_1=0.1$ $X_2=0.075$, $X_3= -0.175$ and zero(0) factor sensitivity. To see if this is a true arbitrage portfolio we have to see if the creation of the arbitrage portfolio has created a positive expected return.

$$X_1r_1 + X_2r_2 + X_3r_3 > 0 \text{ or } 15X_1 + 21X_2 + 12X_3 > 0$$

The expected return for the arbitrage portfolio is then $(15\% \cdot 0.1) + (21\% \cdot 0.075) + (12\% \cdot -0.175) = 0.975\%$

Since 0.975% is a positive number, we have identified an arbitrage portfolio that fits all assumptions and rules.

The practical measures of making the arbitrage portfolio involve buying $1 200 000$ of stock 1 and $900 000$ of stock 2 with the money we get from selling $2 100 000$ of stock 3. We get these figures by multiplying the arbitrage portfolio weights and wealth, $X_iW_0$. $X_1 = 0.1$, $X_2=0.075$, $X_3 = -0.175$, $W_0 = $12 000 000.

Now we see why investors that are not concerned with nonfactor risk find arbitrage portfolios attractive. By creating and arbitrage portfolio and shifting the weights, expected return has risen without any additional investment (Sharpe, 1999).

1.3 Investors position

After creating an arbitrage portfolio, an investor can ether choose to the New portfolio, which is the sum of the old and arbitrage portfolio. Or an investor can hold the old and the new portfolio separately. Let’s take an example. The weights of stock 2, in the old portfolio the weight of stock 2 was 0.333 and in the arbitrage portfolio it was 0.075. The sum of these two is 0.333+0.075 = 0.408 and is equivalent to the weight of stock 2 in the New portfolio.
Following the same logic, the expected return of the New portfolio is the same as the sum of the old and the arbitrage portfolio. 

\[(0.433 \cdot 15\%) + (0.405 \cdot 21\%) + (0.158 \cdot 12\%) = 16.975\%\]

The sensitivity of the New portfolio is 

\[1.9 = 1.9 + 0\]

which is the sum of the old and arbitrage portfolio.

The standard deviation of the Old portfolio was 11% but what about the standard deviation of the New portfolio? The only deviation that the arbitrage portfolio has is the nonfactor risk and it is small. There for the deviation of the New portfolio will be a little bigger than 11% (Sharpe, 1999).

### 1.4 Pricing effects

By creating the arbitrage portfolio we are actually buying and selling different securities and if that is done in a large enough volume, the price of the securities will change. Since everyone is looking for and arbitrage opportunity, in the example above some amount of stock 1 and 2 will be purchased and stock 2 will be sold. The result of the activity is that the price of stock 1 and 2 will rise and the price of stock 3 will fall.

What is the consequence of the changed prices? The rise in the price of stock 1 and 2 will cause expected return of those stocks to fall, the selling pressure on stock 3 will cause expected return to rise.

\[\bar{r} = \frac{P_1}{P_0} - 1\]

Where \(P_0\) is the stock current price and \(P_1\) is the end-of-period price. In the example above the prices of stock 1 and 2 increased and therefore the denominator gets larger and \(\bar{r}\), expected returns gets smaller. The buying and selling activity will continue until there’s no arbitrage opportunities left. Additional portfolio adjustments that have a zero factor exposure and require no additional funds does not result in a positive change in expected return of the portfolio.

When all arbitrage possibilities have been executed there will exist an approximately linear relationship between expected return and sensitivity to the factor.

\[\bar{r} = \lambda_0 + \lambda_1 b_i\]
The lambdas are constants and the equation is the asset pricing model of the Arbitrage Pricing Theory when returns are generated by one factor. As we can see the equation is a simple linear relationship between expected return and sensitivities with a constant intercept. As an example, let’s say that that $\lambda_0$ equals 8 and $\lambda_1$ equals 4.

$$\bar{r} = 8 + 4b_i$$

The stock in the example portfolio would therefore have the following expected returns.

$$\bar{r}_1 = 8 + 4 \cdot 0.9 = 11.9\%$$
$$\bar{r}_2 = 8 + 4 \cdot 3.0 = 20.0\%$$
$$\bar{r}_3 = 8 + 4 \cdot 1.8 = 15.2\%$$

If we compare these expected returns with the expected returns in figure 1, the expected returns have fallen for stock 1 and 2 and risen for stock 3. This is the result from the creation of the arbitrage portfolio and the increase in buying pressure of stock 1 and 2 and selling pressure of stock 3 (Sharpe, 1999).

### 1.5 Two-factor models

In reality securities are affected by more than one factor and first the two-factor model is going to be explained, followed by the general model with more than two factors.

In this example each security has two sensitivities, $b_{i1}$ and $b_{i2}$, thus the factor model is the following:

$$r_i = a_i + b_{i1} + b_{i2} + e_i$$

We also have four securities in a portfolio with the following expected returns and sensitivities.

**Table 2 appendix**

<table>
<thead>
<tr>
<th>Stock</th>
<th>Expected return</th>
<th>Sensitivity to industrial production</th>
<th>Sensitivity to inflation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stock 1</td>
<td>15%</td>
<td>0.9</td>
<td>2.0</td>
</tr>
<tr>
<td>Stock 2</td>
<td>21%</td>
<td>3.0</td>
<td>1.5</td>
</tr>
<tr>
<td>Stock 3</td>
<td>12%</td>
<td>1.8</td>
<td>0.7</td>
</tr>
<tr>
<td>Stock 4</td>
<td>8%</td>
<td>2.0</td>
<td>3.2</td>
</tr>
</tbody>
</table>
In each asset, $5,000,000 is invested which equals a total of $20,000,000 = W_0

1.6 Arbitrage portfolios with the two-factor model

Is the portfolio above in equilibrium according to the APT? How do we know that the assets have the right weights in the portfolio? As long as it is possible to create an arbitrage portfolio with positive expected returns, the portfolio is not in equilibrium.

Just as with the previous example of an arbitrage portfolio, no additional investment is allowed and the arbitrage portfolio should have zero sensitivity to each factor.

\[ X_1 + X_2 + X_3 + X_4 = 0 \]

The sum of the Changes equals zero

\[ 0.9 + 3X_2 + 1.8X_3 + 2X_4 = 0 \]

Sensitivities regarding factor 1

\[ 2X_1 + 1.5X_2 + 0.7X_3 + 3.2X_4 = 0 \]

Sensitivities regarding factor 2

There are many solutions to these equations. We arbitrarily set \( X_1 \) to 0.1 and then solve the remaining weights. The result is the potential arbitrage portfolio with the following weights \( X_1 = 0.1, X_2 = 0.088, X_3 = -0.108, X_4 = -0.8 \)

We also have to check if the potential arbitrage portfolio has a positive expected return, which is a requirement. \( (0.1 \cdot 15\%) + (0.088 \cdot 21\%) + (-0.108 \cdot 12\%) + (-0.08 \cdot 8\%) = 1.41\% > 0 \)

Since 1.41\% is larger than zero, an arbitrage portfolio has been realized.

When investors see this arbitrage possibility, they will act upon it and buy more of stock 1 and 2 and sell some of stock 3 and 4. This will drive the prices of stocks 1 and 2 up and the prices of stock 3 and 4 down. Expected return of stocks 1 and 2 will decrease and expected return of stock 3 and 4 will increase. This will happen until there are no more arbitrage possibilities left and then we have reached equilibrium. In equilibrium, assets will have the following linear relationship between expected return and the two sensitivities:

\[ \tilde{r}_i = \lambda_0 + \lambda_1 b_{i1} + \lambda_2 b_{i2} \]

The equation is still linear but now it has three variables and dimensions instead of two.

Let’s explore an example. Let \( \lambda = 8, \lambda_1 = 4 \) and \( \lambda_2 = -2 \)

\[ \tilde{r}_i = 8 + 4b_{i1} - 2b_{i2} \]
After all arbitrage possibilities have been utilized, the stocks and portfolio will have the following expected returns.

\[
\bar{r}_1 = 8 + 4 \cdot 0.9 - 2 \cdot 2.0 = 7.6\%
\]
\[
\bar{r}_2 = 8 + 4 \cdot 3.0 - 2 \cdot 1.5 = 17.0\%
\]
\[
\bar{r}_3 = 8 + 4 \cdot 1.8 - 2 \cdot 0.7 = 13.8\%
\]
\[
\bar{r}_4 = 8 + 4 \cdot 2.0 - 2 \cdot 3.2 = 9.6\%
\]

What is the aftermath of the equilibrium portfolio? The expected return of stock 3 and 4 has risen whilst is has fallen for stock 1 and 2. Observe that stock 2 has very large estimated return and a large sensitivity to the first factor.

**1.7 Pricing effects in the two-factor model**

The pricing equation for the two-factor model is very similar to the one-factor model.

\[
\bar{r}_i = r_f + (\delta_1 - r_f)\beta_{1,i} + (\delta_2 - r_f)\beta_{2,i}
\]

Where \(\delta_1\) is a well-diversified, unit sensitive portfolio to the first factor Industrial productivity and \(\delta_2\) is a well-diversified unit sensitive portfolio of the second factor Inflation. The expected returns from the portfolios \(\delta\) is then subtracted by the risk-free rate. Now we see that the lambdas are equal to the unit sensitive well-diversified portfolios minus the risk-free rate and they are risk premiums.

Investing in the asset portfolios is more risky than the risk-free rate and the lambdas show how much more an investor gets in return from taking on more risk. In the example above the first factor gave a risk premium of 4 and the second a risk premium of -2. Different stocks have different sensitivities to the factor, stock 1 for example had a beta to the first factor of 0.9 and 2 for the second factor (Sharpe, 1999).

**1.8 APT and multiple factor models**

When analyzing what factors affect a security, most likely a researcher will use more than two factors in trying to explain expected return. Fortunately APT can handle more than 2 factors.
If there are more than 2 factors, we can use a general model with $k$ ($F_1, F_2, ..., F_k$) factors and $k$ sensitivities ($b_{i1}, b_{i2}, ..., b_{ik}$) model looks like this:

$$\bar{r}_i = a_i + b_{i1}F_1 + b_{i2}F_2 + \ldots + b_{ik}F_k + e_i$$

Conversely the APT pricing equation will be the following:

$$\bar{r}_i = \lambda_0 + \lambda_1 b_{i1} + \lambda_2 b_{i2} + \ldots + \lambda_k b_{ik}$$

It is still a linear equation but with $k + 1$ dimensions. Now we can look at a stocks expected return with $k$ factors and a risk-free rate. According to the APT, this is how expected return of a security with $k$ factors is calculated (Sharpe, 1999).

$$\bar{r}_i = r_f + (\delta_1 - r_f)b_{i1} + (\delta_2 - r_f)b_{i2} + \ldots + (\delta_k - r_f)b_{ik}$$

**Appendix 2: Statistical inference**

**A) Statistical inference for Model 1**

<table>
<thead>
<tr>
<th>Durbin-Watson test for serial correlation in Model 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>. dwstat</td>
</tr>
<tr>
<td>Durbin-Watson d-statistic(11, 51) = 2.008885</td>
</tr>
</tbody>
</table>

A Durbin-Watson d-statistic value close to 0 indicates negative serial correlation and a value close to 4 indicates positive serial correlation (Studenmund 2010). Since the estimates Durbin-Watson statistic value of the data in Model 1 was 2.008885 I did not reject the null hypothesis of no serial correlation.

<table>
<thead>
<tr>
<th>VIF values for Model 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Industrial production</td>
</tr>
<tr>
<td>Industrial production</td>
</tr>
<tr>
<td>Industrial confidence</td>
</tr>
<tr>
<td>Industrial confidence</td>
</tr>
<tr>
<td>Consumer confidence</td>
</tr>
<tr>
<td>Consumer confidence</td>
</tr>
<tr>
<td>Unemployment</td>
</tr>
<tr>
<td>Unemployment</td>
</tr>
<tr>
<td>CPI</td>
</tr>
<tr>
<td>CPI</td>
</tr>
<tr>
<td>10 year bond</td>
</tr>
<tr>
<td>10 year bond</td>
</tr>
</tbody>
</table>
Under the assumption that a VIF value of 10 or more is considered high multicollinearity, we see that seven out of 10 variables have high multicollinearity (Studenmund 2010).

**Breusch-Pagan test for heteroscedasticity in Model 1**

. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
- Ho: Constant variance
- Variables: fitted values of r

<table>
<thead>
<tr>
<th>chi2(1)</th>
<th>Prob &gt; chi2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.51</td>
<td>0.4772</td>
</tr>
</tbody>
</table>

H₀ Constant variance vs Hₐ Heteroskedasticity
Since p = 0.4772 > 0.05 we cannot reject the null hypothesis of constant variance. We do not have heteroscedasticity.

**Residual plots for Model 1**

The residuals of Model 1 looks fairly normally distributed. The residuals also look independent when plotted against the order of the observations.
B) Statistical inference Model 2

Durbin-Watson test for serial correlation in Model 2

. dwstat

Durbin-Watson d-statistic( 6, 51) = 1.818303

A Durbin-Watson d-statistic value close to 0 indicates negative serial correlation and a value close to 4 indicates positive serial correlation. (Studenmund 2010) Since the estimates Durbin-Watson statistic value of the data in Model 2 is 1.818303, I did not reject the nullhypotesis of no serial correlation.

VIF values for Model 2

<table>
<thead>
<tr>
<th>Variable</th>
<th>Region</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industrial confidence</td>
<td>Eurozone</td>
<td>16.153</td>
</tr>
<tr>
<td>Consumer confidence</td>
<td>USA</td>
<td>2.625</td>
</tr>
<tr>
<td>CPI</td>
<td>USA</td>
<td>20.342</td>
</tr>
<tr>
<td>10 year bond</td>
<td>Eurozone</td>
<td>5.554</td>
</tr>
<tr>
<td>10 year bond</td>
<td>USA</td>
<td>4.701</td>
</tr>
</tbody>
</table>

Under the assumption that a VIF value of 10 or more is considered high multicollinearity, we see that two out of 5 variables have high multicollinearity.

Breusch-Pagan test for heteroscedasticity in Model 2

. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
   Ho: Constant variance
   Variables: fitted values of r

   chi2(1)    =   1.57
   Prob > chi2 = 0.2103

H0 Constant variance vs Ha Heteroskedasticity Since p = 0.2103 > 0.05 we cannot reject the nullhypotesis of constant variance. I did not have heteroscedasticity.

Residual plots for Model 2
The residuals of Model 1 looks fairly normally distributed. The residuals also look independent when plotted against the order of the observations.

C) VAR and Granger-causality test

Output from Model 1 using vector autoregression via STATA.
Output from the Granger-causality test of Model 1 using \texttt{var(1)} via \textsc{Stata}. The null hypothesis is that there exists no granger-causality.
Granger causality Wald tests

<table>
<thead>
<tr>
<th>Equation</th>
<th>Excluded</th>
<th>chi2</th>
<th>df</th>
<th>Prob &gt; chi2</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>CPIus</td>
<td>3.8408</td>
<td>1</td>
<td>0.050</td>
</tr>
<tr>
<td>r</td>
<td>ConConfUs</td>
<td>.01882</td>
<td>1</td>
<td>0.891</td>
</tr>
<tr>
<td>r</td>
<td>IndConfEZ</td>
<td>1.6847</td>
<td>1</td>
<td>0.194</td>
</tr>
<tr>
<td>r</td>
<td>Bond10us</td>
<td>4.0362</td>
<td>1</td>
<td>0.045</td>
</tr>
<tr>
<td>r</td>
<td>Bond10ez</td>
<td>1.9726</td>
<td>1</td>
<td>0.160</td>
</tr>
<tr>
<td>r</td>
<td>ALL</td>
<td>5.7571</td>
<td>5</td>
<td>0.331</td>
</tr>
<tr>
<td>CPIus</td>
<td>r</td>
<td>2.932</td>
<td>1</td>
<td>0.087</td>
</tr>
<tr>
<td>CPIus</td>
<td>ConConfUs</td>
<td>13.258</td>
<td>1</td>
<td>0.000</td>
</tr>
<tr>
<td>CPIus</td>
<td>IndConfEZ</td>
<td>.18593</td>
<td>1</td>
<td>0.666</td>
</tr>
<tr>
<td>CPIus</td>
<td>Bond10us</td>
<td>.01337</td>
<td>1</td>
<td>0.908</td>
</tr>
<tr>
<td>CPIus</td>
<td>Bond10ez</td>
<td>.00827</td>
<td>1</td>
<td>0.928</td>
</tr>
<tr>
<td>CPIus</td>
<td>ALL</td>
<td>16.054</td>
<td>5</td>
<td>0.007</td>
</tr>
<tr>
<td>ConConfUs</td>
<td>r</td>
<td>.00118</td>
<td>1</td>
<td>0.973</td>
</tr>
<tr>
<td>ConConfUs</td>
<td>CPIus</td>
<td>.01767</td>
<td>1</td>
<td>0.894</td>
</tr>
<tr>
<td>ConConfUs</td>
<td>IndConfEZ</td>
<td>.21552</td>
<td>1</td>
<td>0.642</td>
</tr>
<tr>
<td>ConConfUs</td>
<td>Bond10us</td>
<td>.3339</td>
<td>1</td>
<td>0.563</td>
</tr>
<tr>
<td>ConConfUs</td>
<td>Bond10ez</td>
<td>.14014</td>
<td>1</td>
<td>0.708</td>
</tr>
<tr>
<td>ConConfUs</td>
<td>ALL</td>
<td>5.8884</td>
<td>5</td>
<td>0.317</td>
</tr>
<tr>
<td>IndConfEZ</td>
<td>r</td>
<td>.58454</td>
<td>1</td>
<td>0.445</td>
</tr>
<tr>
<td>IndConfEZ</td>
<td>CPIus</td>
<td>2.9939</td>
<td>1</td>
<td>0.084</td>
</tr>
<tr>
<td>IndConfEZ</td>
<td>ConConfUs</td>
<td>.03236</td>
<td>1</td>
<td>0.857</td>
</tr>
<tr>
<td>IndConfEZ</td>
<td>Bond10us</td>
<td>.14051</td>
<td>1</td>
<td>0.708</td>
</tr>
<tr>
<td>IndConfEZ</td>
<td>Bond10ez</td>
<td>1.8415</td>
<td>1</td>
<td>0.175</td>
</tr>
<tr>
<td>IndConfEZ</td>
<td>ALL</td>
<td>9.8765</td>
<td>5</td>
<td>0.079</td>
</tr>
<tr>
<td>Bond10us</td>
<td>r</td>
<td>.15189</td>
<td>1</td>
<td>0.697</td>
</tr>
<tr>
<td>Bond10us</td>
<td>CPIus</td>
<td>.22346</td>
<td>1</td>
<td>0.636</td>
</tr>
<tr>
<td>Bond10us</td>
<td>ConConfUs</td>
<td>.70525</td>
<td>1</td>
<td>0.401</td>
</tr>
<tr>
<td>Bond10us</td>
<td>IndConfEZ</td>
<td>.57091</td>
<td>1</td>
<td>0.450</td>
</tr>
<tr>
<td>Bond10us</td>
<td>Bond10ez</td>
<td>.33121</td>
<td>1</td>
<td>0.565</td>
</tr>
<tr>
<td>Bond10us</td>
<td>ALL</td>
<td>11.813</td>
<td>5</td>
<td>0.037</td>
</tr>
<tr>
<td>Bond10ez</td>
<td>r</td>
<td>1.0857</td>
<td>1</td>
<td>0.297</td>
</tr>
<tr>
<td>Bond10ez</td>
<td>CPIus</td>
<td>.00057</td>
<td>1</td>
<td>0.981</td>
</tr>
<tr>
<td>Bond10ez</td>
<td>ConConfUs</td>
<td>3.757</td>
<td>1</td>
<td>0.053</td>
</tr>
<tr>
<td>Bond10ez</td>
<td>IndConfEZ</td>
<td>1.6661</td>
<td>1</td>
<td>0.197</td>
</tr>
<tr>
<td>Bond10ez</td>
<td>Bond10us</td>
<td>.71619</td>
<td>1</td>
<td>0.397</td>
</tr>
<tr>
<td>Bond10ez</td>
<td>ALL</td>
<td>8.8423</td>
<td>5</td>
<td>0.116</td>
</tr>
</tbody>
</table>