universität wien

MASTERARBEIT

Titel der Masterarbeit

„Measuring the effects of macroeconomic announcements on asset prices: An empirical analysis of their impact on European Stock Index Futures“

Verfasserin

Katerina Koule

angestrebter akademischer Grad

Master of Science (MSc)

Wien, 2015

Studienkennzahl lt. Studienblatt:	A 066 920
Studienrichtung lt. Studienblatt:	Master in Quantitative Economics, Management and Finance
Betreuer	Ao. Univ.-Prof. Dr. Jörg Borrmann

Declaration of Authorship

I, Katerina Koule, declare that this thesis with the title, „Measuring the effects of macroeconomic announcements on asset prices: An empirical analysis of their impact on European Stock Index Futures" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.

- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.

- Where I have consulted the published work of others, this is always clearly attributed.

- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

- I have acknowledged all main sources of help.

- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

Acknowledgments

A big thank you to my parents for the endless support and my brother for the valuable help with the data handling.

Last but not least, a big thank you to my professor Dr. Stephan Unger for the continuous guidance, mentoring and patience through all these months.

Abstract

This thesis investigates how new information from macroeconomic announcements affects different asset prices. In particular, it seeks to measure the surprise factor after macroeconomic releases and its relation to asset returns. After presenting the most significant models in the existing literature for measuring these effects, a new empirical model is developed for the European stock market, focusing on German, French and EU releases. To the best of the author's knowledge, this is the first attempt to analyse such effects in Europe. The model suggests that for some announcement types the effect is rather significant and it lasts for from thirty seconds to three minutes, depending on the type. This shows how quickly any new information is incorporated into the prices, which is mainly due to high-frequency trading. Moreover, it enables the measurement of the magnitude and the direction of the effect. Furthermore, it shows that these observed effects can be indeed attributed to the new information, irrespectively of the "regular" movements of the market. In addition, the results suggest that French announcements do not impact the market at all, in contrast to expectations. Lastly, the paper examines if these findings can be used to form a strategy for high-frequency trading and, at the end, it presents a sample strategy.

Zusammenfassung

Die folgende Arbeit untersucht die Auswirkungen neuer Informationen aus makroökonomischen Ankündigungen auf verschiedenen Assetpreise. Im Speziellen wird der sogenannte „surprise factor" nach makroökonomischen Veröffentlichungen sowie dessen Beziehung mit den Asset Return untersucht. Nach der Vorstellung der wichtigsten Modelle innerhalb der vorherrschenden Literatur, die für eine derartige Untersuchung geeignet sind, wird ein neues empirisches Model präsentiert. Dieses neue Model ist in der Lage, unter Berücksichtigung europäischer Ankündigungen, hier insbesondere aus Deutschland, Frankreich sowie der restlichen EU, die Auswirkungen auf makroökonomische Grundlagen zu messen. Nach bestem Wissen und Gewissen der Autorin ist dies der erste Versuch solche Auswirkungen im europäischen Raum zu messen. Das Model deutet auf signifikante Effekte hin, die zwischen 30 Sekunden und zwei Minuten, abhängig vom jeweiligen Typ, anhalten. Dies wiederum zeigt auf wie schnell neue Informationen in die Preisgestaltung eingearbeitet werden, was wiederum auf die Natur von „high-frequency trading" zurückzuführen ist. Zusätzlich erlaubt das Model die Messung der Größenordnung sowie die Richtung des Effekts. Weiters wird bewiesen, dass die beobachteten Effekte auf die neuen Informationen zurückführen sind. Ferner deuten die Ergebnisse darauf hin, dass Ankündigungen in Frankreich keine Auswirkungen auf den Markt haben. Schlussendlich untersucht die Arbeit ob die Ergebnisse für eine „high frequency trading" Strategie verwendet werden können und schlägt eine solche Strategie vor.

Contents

List of abbreviations

ARCH	Autoregressive Conditional Heteroscedasticity
CPI	Consumer Price Index
FESX	EUROSTOXX 50 ® Index Future
GARCH	Generalized Autoregressive Conditional Heteroscedasticity
MMS	Money Market Services
OLS	Ordinary Least Squares
PPI	Producer Price Index
RPI	Retail Price Index
SX5E	EUROSTOXX 50 ® Index
WLS	Weighed Least Squares

List of tables

List of figures

1. Introduction

Every year national statistic offices, research agencies other institutions put a lot of effort in to estimating and disclosing macroeconomic fundamentals as well as indices about the economic prosperity of each country. In addition, banks, funds and asset management companies invest both time and money to monitor and even predict their values before they are announced. This shows that announcements about macroeconomic fundamentals play an important role in economic decision making. But do these announcements affect prices of assets like equities, bonds and futures? And if they do so, what is their effect and what is its direction? These are questions that many researchers have tried to answer in the last few years.

Although in the early years there was little evidence of the relation between macroeconomic announcements and asset prices, recent and more advanced models, in combination with high frequency data, have shown that some announcements have a considerable but short-lived effect on certain asset prices. However, most of the existing literature focuses only on announcements of US macroeconomic fundamentals and their effect on either US or European markets. There is no research on how announcements of European macroeconomic indicators affect European or other markets. Lastly, to the author's best knowledge, there has been no attempt to investigate if this impact on the prices can be used for building a training strategy.

This paper tries to cover those aspects missing from current literature. It is the first one to take into account European macroeconomic announcements and it investigates their impact on the European stock market. In addition, it shows how this information can be applied to high-frequency trading and it proposes a sample trading strategy.

By using a dataset, consisting of tick-by-tick prices of a period of four years (2008-2012), it combines well-established models and suggests a methodology that can be easily applied to large datasets, without causing computational problems. This methodology does not only show which macroeconomic announcements are considered important from the market participants and hence affect the market, but also illustrates the magnitude, the direction and the duration of the effect. This last aspect indicates the speed of the price adjustment to the incoming information, i.e. the level of market liquidity, an important factor when it comes to deciding on a trading strategy.

The paper shows that there are certain macroeconomic announcements that affect the European stock market considerably, such as the ECB interest rate and German GDP, amongst others. In all cases, the direction of the effect is as pre-directed by the economic theory, while its duration varies between thirty seconds to three minutes. It is also proven that these effects are not caused by the "regular" movements of the prices, but are, indeed, results of the new information arriving in the market.

The rest of the paper is organised as follows: Section 2 gives an overview of the existing literature, presenting the most used methods for measuring announcement effects. Section 3 describes the methodology for the empirical analysis, while section 4 explains its results. Finally, section 5 proposes a high-frequency trading strategy, using the derived results.

2. Existing literature on measuring the effects of macroeconomic announcements

The existing research on measuring the effects of macroeconomic announcements on asset prices is long and dates back to the late eighties. Since then, several models have been developed, most of which have focused on the impact of US macroeconomic announcements on a wide range of assets and rates. This range includes interest rates, foreign exchange rates, stocks and/or stock indices worldwide, futures of all kinds etc. Little emphasis has been given on UK macroeconomic announcements and none on European. Apart from that, the literature has concentrated on answering which releases affect which prices and how long these effects last. The investigation has become easier and more detailed in the recent years, since high-frequency data for prices, bid-ask spreads and trading volumes are available. This is of high importance in an era of high-frequency and computerized trading, which leads the prices to the new equilibria within a few minutes of a shock.

The following review is structured chronologically and by the basic models that influenced similar works in different periods. It is worth mentioning that only models that are used for capturing the effects of scheduled macroeconomic announcements are presented, as well as those which aim to measure the duration of the effects. Moreover, special emphasis is given to the results that show the micro-dynamics of the futures markets, since the main instruments used in the empirical analysis in the next section are futures.

2.1 Early models: differentiation between the expected and the unexpected part of an announcement

Earlier works use mostly daily data for prices and they divide the announced values of the macroeconomic indicators in two parts: the expected part and the unexpected part. On the one hand, the expected part is defined by surveys showing the median forecasts of market participants, such as the one conducted by Money Market Services (MMS) International (Dwyer and Hafer (1989)). On the other hand, the unexpected part is defined as the difference between the actual reported value and the expected one. These two parts are used in a simple regression, in order to explain the changes in the prices. Two of the most characteristic works that use this methodology are Dwyer and Hafer (1989) and Cook and Korn (1991).

Dwyer and Hafer (1989) study the reaction of interest rates to the surprises from the releases of US money supply, producer price index (PPI), consumer price index (CPI), trade balance, unemployment rate and index of leading indicators. For this, they regress the change of the interest rate during a day, covering the announcement time, and the unexpected part of the announcement as follows:

$$\Delta R_t = a + \beta U_t + \varepsilon_t \tag{1}$$

where

ΔR_t is the interest rate change and

U_t is the unexpected part.

They use daily interest rates of three-months Treasury Bills and 30-years Treasury Bonds, as well as the forecasts from the MMS International survey. Their dataset consists of only those days when an announcement occurred. ΔR_t is measured as the difference between the closing prices from one day to the other for both, Treasury Bills and Treasury Bonds. U_t is defined as the percentage change between the actual value of the indicator and the forecasted one, over the actual. The regression then is estimated for each indicator separately, showing the effect of the announcement through the amplitude of β. Their results show that only the money supply affects both short-term and long-term interest rates.

Cook and Korn (1991) study the change in the interest rate as well, but only in relation to the US employment report. Driven by evidence that market agents use the report as an indicator for FED's upcoming targeted federal funds rate, which subsequently affects the Treasury bill rates, they examine the relation of the latter to the employment and unemployment rates. The regression that they use is as follows:

$$\Delta R_{n_t} = a + b_1 \Delta Expected\ Emp_t + b_2 \Delta Unexpected\ Emp_t + b_3 \Delta Expected\ UR_t + b_4 \Delta Unexpected\ UR_t + b_5 \Delta Rev_t + e_t \tag{2}$$

where

ΔR_{n_t} is the daily change of the n-month Treasury bill, covering the employment report release time,

Emp_t is the employment rate,

UR_t is the unemployment rate and

Rev_t is the revision of the previously reported value.

Their dataset consists of three-months, six-months and twelve-months Treasury bill yields, as well as employment and unemployment rate forecasts from MMS International for the period Feb. 1985–Apr. 1991. The results show that the expected part from the employment and unemployment rate does not influence any of the interest rates. On the other hand, there is strong evidence that the unexpected part, as well as the revision, do influence all interest rates. This indicates that any new information indeed affects the market.

2.2 Ederlington's and Lee's influence: identifying significant announcements using dummies

During the nineties researchers concentrated more on the market dynamics. The focus was not only on finding which announcements move the markets but also on estimating the market adjustment time and volatility persistence. This was possible by using intraday data and dividing the sample into announcement and non-announcement days. In addition, in most cases, futures were used as a close substitute for the underlying instrument in focus. The reason why futures are so popular in this kind of research is that they are heavily traded, their trading times are longer, covering a wider number of scheduled macroeconomic releases (compared to the underlying asset), and their prices are available tick-by-tick. In addition, transaction costs are lower in the future markets, making it easier to incorporate any new information (Andersen et al. 2007).

One of the earliest and most influential works of this kind is the one by Ederlington and Lee (1993). They study the effect of 19 scheduled US macroeconomic announcements (such as CPI, PPI, House Sales etc.) on the Treasury bond, Eurodollar and Deutsche Mark futures market and they generalize their results for the underlying instrument, i.e. interest rates and foreign exchange rates respectively.

They use tick-by-tick trading prices for the instruments mentioned above over a period of 775 trading days (07.11.1989–29.11.1991) and they divide the sample in two parts, announcement and non-announcement days. For their analysis they also calculate the log returns, defined as $R_{jt} = \ln\left(\frac{P_{jt}}{P_{(j-1)t}}\right)$, for every five-minute interval j over the trading day t. Since the announcements in the sample are released either on 8:30, 9:15, 10:00 or 14:00, only these intervals are taken into account[1]. Then they estimate the following regression, for every j-th 5-minutes interval of each trading day, over the whole sample:

[1] Considered intervals: 8:30-8:35, 9:15-9:20, 10:00-10:05 and 14:00-14:05.

$$\left|R_{jt} - \bar{R}_j\right| = a_{0j} + \sum_{k=1}^{K} a_{kj} D_{kt} + \varepsilon_{jt} \tag{3}$$

where

$\bar{R}_j$ is the mean return on the j-th five-minutes interval over the sample,

D_{kt} is a dummy variable for each announcement k, which takes the value 1 if the announcement was released on day t and 0 otherwise. If two announcements were made at the same day D takes the value 0.5 for both, hence giving the same weight to them.

k is the index of each of the announcements in the study (here K=19) and

a_{kj} is the variable to be estimated.

They conduct this regression separately for each of the three markets in question. It is clear that the size and the statistical significance of the coefficients show the impact of each announcement k on the respective market in each 5-minutes interval. However, the model cannot be applied if more announcements occur at the same time, or the same announcement has inconsistent release time over the years.

In contrast to Dwyer and Hafer (1989), Ederlington and Lee (1993) find that more announcements affect the interest rate market, not only the money supply (with significance level 0.005 or lower). This might be due to the use of intraday data. There is evidence that the new information is incorporated in the prices very quickly, right after the release. In most cases, for example, when an announcement is found to be significant in the 8:30-8:35 interval, it is no longer significant in the 9:15-9:20 interval. The same situation is observed in the foreign exchange market. Hence, the model of Dwyer and Hafer (1989) would have never captured this effect.

To estimate how long the new information needs to be integrated into the prices Ederlington and Lee (1993) calculate the standard deviation of the prices for each five-minute interval after the announcement, for both announcement and non-announcement days. They define an announcement day as any day in the sample, on which at least one announcement has occurred. Then they observe how long it takes for the volatility to return to its normal "non-announcement" levels. The volatility on the announcement days is significantly higher compared to the non-announcement days over the first five minutes and it declines thereafter. It returns to its normal levels within approximately twenty minutes. This is visible for all three markets in question.

Moreover, they regress the equation (3) for the first nine 5-minutes intervals after each major announcement. They find that the significance of the coefficients is gradually decreasing and, in most cases, disappears between the second and fifth interval, i.e. after ten to twenty-five minutes. Hence the speed of price adjustment depends on the kind of the announcement. It seems that market participants need more time to evaluate the new information coming from certain releases. The authors, for example, mention that after the release of the employment report the prices need forty to forty-five minutes to adjust because this report is a few pages long and contains much more information than others.

Lastly, they examine if any trading profits are possible, by testing market efficiency. Their main assumption is that the returns in each successive minute will be correlated until the equilibrium price is reached. Hence they calculate the covariance between the returns in each minute after the announcement and their successive ones. They find no significant correlation between the returns of the first minute after the announcement and the upcoming, at least for the Treasury bonds and the Deutsche Mark market. For the Eurodollar market, they find significant but negative correlation between the returns in the first five minutes. They explain this as a gradual correction of the overreaction of the market during the first minute. In general, they conclude that any possibility for profits through trading is possible only within the first minute after the release. However, they mention that experienced traders can assess if the market has over- or underreacted and make some profit even after the first minute, by making use of the persistence of the volatility.

Besides its restrictions, Ederlington and Lee's model (1993) was widely used in later research as a basis for identifying major announcements that influence different markets. Ap Gwilym et al. (1998), for example, use the exact same regression for identifying the effects of several UK macroeconomic releases on the UK stock and interest rate market and then use the most influential ones for deeper empirical analysis. Similarly to the original model, they use futures contracts as a proxy for the markets under consideration (FTSE 100 Index futures for the stock market and Short Sterling interest rate futures for the interest rate market). One of the major findings is that Retail Sales, PPI and Retail Price Index (RPI) are important for both markets. Interestingly, GDP, Balance of payments and Monetary Statistics seem to play no role in the price changes.

Subsequently, they examine volatility persistence in a similar way to Ederlington and Lee (1993), but with a slightly different setup. They first calculate the standard deviation of returns for each fifteen-second interval within a time-window of twelve minutes, which starts two minutes before the announcement and ends ten minutes after the release. This time-window is long enough to capture any effects prior to and after the announcement. Then they compare these values with the volatility of returns on the non-announcement days. They find that the volatility of the stock market is lower than normal up to fifteen seconds before the announcement, and then it begins to rise, reaching its peak almost one and a half minutes after the release. Afterwards, it starts falling but remains higher than the normal levels for the next ten minutes. In the interest rate market, the volatility is higher than normal even two minutes before the release and it peaks in less than a minute after the announcement. Note that only the announcements of the major indicators, as identified previously, are taken into account for the volatility analysis.

The authors suggest two factors for explaining this high volatility on announcement days. It can either be due to higher absolute returns or due to higher number of non-zero returns, as compared with non-announcement days. They conclude that both factors are responsible for the persistence in volatility. To prove that, they examine the mean absolute returns and the mean number of transactions for each fifteen-second interval within the announcement window and compare it with the respective means on non-announcement days. The results show that in the case of the stock futures market the mean absolute returns are lower than usual before the announcement, rise significantly within the first fifteen seconds after the announcement and remain at higher-than-usual levels for over five minutes. Looking at the mean transactions, the same picture can be seen, but the number of mean transactions remains significantly higher until the end of the announcement window. Similar results appear for the interest rate futures market, but the mean returns and number of transactions are higher even two minutes before the announcement. Since the number of transactions is related to the number of non-zero returns, it is evident that the higher volatility can be attributed not only to higher absolute returns but also to the larger number of non-zero returns.

Moreover, in order to identify the duration of the price adjustment to the new equilibrium, they compare the amount of continuations[2] per fifteen-second interval within the announcement window (calculated as a percentage over the total amount) with the "normal" one from the non-announcement days. They suggest that in case of a shock in the market, a temporary imbalance is caused and a number of successive continuations until the price reaches the equilibrium level can be observed. Hence, a higher than usual number of continuations implies this movement and the duration of these successive continuations shows the duration of the price adjustment. As a result, the stock futures market seems to adjust in about half a minute, while the interest rate futures market needs two and a half minutes to adjust. This is yet more evidence that scheduled macroeconomic news are incorporated into prices very quickly, even during the nineties, when the trading was still performed on the floor.

The Ederlington and Lee model (1993) is used even in more recent studies, although in most of them it is slightly altered and combined with other statistical models, such as GARCH. Kim, Mc Kenzie and Faff (2003), for example, use also dummy variables to

[2] Based on Gosnell (1995) as continuation is defined a price change that follows the same direction as the previous one.

identify the announcements, but do not use the absolute difference of returns from the mean return as the dependent variable. Instead, they use row daily returns. In order to count for "day-of-the-week effect", they use one more dummy for each day. In addition, they measure the volatility of the residuals by using a GARCH-like model. All in all, their model has the following form:

$$R_t = \mu_t + \sum_{i=MON}^{THU} \mu_i D_i + \sum_{j=1}^{6} \mu_j D_j + \varepsilon_t \tag{4}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} + \sum_{i=MON}^{THU} \mu_i D_i + \sum_{j=1}^{6} \mu_j D_j + \varepsilon_t \tag{5}$$

where

D_i is the dummy for each day of the week (taking the value 1 if it is MON-THU and 0 otherwise),

D_j is the dummy for each announcement being considered (taking the value 1 if one of the announcements j occurred on day t)

and the second equation accounts for the volatility of the residuals. This last part is an important improvement, as it shows which factors affect the volatility of the market.

In addition, in order to be able to measure the effects of expectations, they divide the announcements into positive and negative, and introduce a separate dummy for each category. As positive they classify the ones that the released value is "better" than the expected one. As expected the use the median value from the MMS International survey. The opposite holds for the negative announcements. In the end, their second model has the form:

$$R_t = \mu_t + \sum_{i=MON}^{THU} \mu_i D_i + \sum_{j=1}^{6} \mu_j D_j + \sum_{k=1}^{6} \mu_k D_k + \varepsilon_t \tag{6}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} + \sum_{i=MON}^{THU} \mu_i D_i + \sum_{j=1}^{6} \mu_j D_j + \sum_{k=1}^{6} \mu_k D_k + \varepsilon_t \tag{7}$$

where

D_j is the dummy for each of the six announcements if the news are positive (taking the value 1 if announcement j occurred and it is positive) and

D_k is the dummy for each of the six announcements if the news are negative (taking the value 1 if announcement k occurred and it is negative).

They regress both models for the JPY/USD and DEM/USD exchange rate market as well as US stock and bond market, over six US macroeconomic announcements. However, their results are rather mixed. They find only a few of the announcements to be significant while the largest part of the returns and volatility movement is affected by the days of the week and past volatility. This is probably due to the use of daily data. As it is evident from other research presented in this section, prices adjust within minutes to the new equilibrium after some a scheduled announcement. Hence, by using daily returns it is difficult to measure this effect.

Lastly, it is worth mentioning the model developed by Jones, Lin and Masih (2005), which enriched the Ederlington and Lee (1993) model by adding lagged returns and a GARCH model for the conditional volatility. In contrast to Kim et al. (2004), they use 5-minute returns. At first they identify the most important announcements using the original model. Then, for the major ones they estimate the following:

$$R_{tj} = a + \beta_1 R_{tj-1} + \beta_2 R_{tj-2} + \sum_{k=1}^{K} \gamma_{kj} D_{ktj} + \varepsilon_{tj} \quad (8)$$

$$h_{tj} = \delta + \lambda \varepsilon_{tj-1}^2 + \omega h_{t-1} + \sum_{k=1}^{K} \rho_k D_{ktj} \quad (9)$$

where

R_{tj} is the log return in the five-minute interval *j* on day *t*,

D_{ktj} is the dummy for announcement *k*, which takes the value 1 if the announcement occurs on day *t* in the five-minute interval *j*

and the second equation accounts for the volatility of residuals. In practice, this can be seen as the conditional volatility of the returns R_{tj}.

Jones et al. (2005) estimate this model for Short Sterling (future contract linked to UK interest rates), Ling Gilt (bond issued by the government of UK) and FTSE 100 index futures and a number of UK macroeconomic announcements. Their results show that there are several releases that affect the first two markets, such as national statistics, unemployment rate, UK monetary policy etc. They also find a negative effect from the lagged returns, which they ascribe to the bid-ask bounce. In addition, they find that the volatility is highly affected by the volatility of the previous term as well as the variance of the previous error term. On the other hand, they don’t find any significant results for the stock market, except the autocorrelation of residuals.

2.3 “Standardized” surprises and their use in more dynamic models

Although the model of Ederlington and Lee (1993) is good for identifying the most influential macroeconomic announcements, it does not show either the size or the direction of the effect. This problem was addressed by Balduzzi, Elton and Green (2001) with the introduction of the "standardized" surprises; a concept which is still used today in many models.

Balduzzi et al. (2001) study the effect of several announcements on US bonds of different maturity, by focusing on the surprise that the release causes to the investors. Since the announcements are always expected at a specific time, investors form some expectations about their values in advance. Hence, the reaction of the market is only due to the surprise element. This concept resembles the one of Dwyer and Hafer (1989) but Balduzzi et al. (2001) generalize it. Firstly they define the "standardized" surprises as:

$$S_i = \frac{E_i}{\sigma_i} = \frac{A_i - F_i}{\sigma_i} \tag{10}$$

where

A_i is the actual value of the indicator that is announced,

F_i is the forecasted value of the indicator as expected by the investors and

σ_i is the standard deviation of E_i

By dividing with the standard deviation of the "simple" surprise E_i, they made it possible to compare results from different macroeconomic indicators, which are usually measured in different units For instance, $S_i = 1$ means one standard deviation change in the surprise.

Secondly, for identifying the effect that each announcement has on the price, they use the following regression:

$$\frac{(P_{30it} - P_{-5it})}{P_{-5it}} = \beta_{0i} + \beta_{1i} S_{it} + \sum_{k=1}^{K} \beta_{k+1,i} S_{i_k t} + \varepsilon_{it} \tag{11}$$

where

P_{30it} is the price of the asset thirty minutes after the release *i* at time *t,*

P_{-5it} is the price five minutes before the release *i* at time *t,*

S_{it} is the surprise from the announcement *i* at time *t* and

$S_{i_k t}$ is the surprise from the announcement k that coincides with announcement *i* at least 10% of the time.

Estimating this model allows to identify the size and the direction of each announcement *i,* as it can be inferred by the size and the sign of β_{1i}. This is feasible even when other announcements occur at the same time. In addition, this model is more dynamic than the one of Ederlington and Lee (1993) and it can be used even if the release time is not constant over the sample. This is feasible because the standardized surprises can be calculated at any time and are not static like the dummies in the Ederlington and Lee (1993) model. Lastly, by calculating the absolute returns starting five minutes before the announcement, it is possible to capture any leakage of information prior to the release.

For conducting their model, Balduzzi et al. (2001) use the median of the MMS International survey as a proxy for investor's forecasts and 3-month Treasury bills, 2-year Treasury notes, 10-year Treasury notes and 30-years bond as assets classes. They estimate the regression for each asset type and announcement separately. The results show that bond markets are highly affected by surprises in macroeconomic indicators. Seventeen out of twenty-six announcements in question are found to be significant. In addition, for all significant announcements, the larger the maturity of the bond, the larger is size of the effect. This is consistent with the theory which suggest that the longer the maturity of a bond the more volatile the price.

Except for identifying the most important announcements and their effect, Balduzzi et al. (2001) model can be also used for estimating the time that the bond's market needs to respond to the surprise. For this reason, they calculate the absolute returns through different time frames, always keeping the end time constant at thirty minutes after the announcement but always modifying the starting time. Then they perform the same regression for every different return type and for those announcements that they find to be significant. The target is to find for which time frame the effect is not significant any longer. The results show that for most of the announcements the market seems to adjust within one minute and for only a few of them the effect lasts longer than three minutes. However, since they only performed this process on the 10-year notes, it cannot be assumed that every bond market adjusts so quickly. It is possible that lower-maturity bond markets adjust more slowly, since they are less volatile.

"Standardized" surprises of Balduzzi et al (2001) gained a lot of popularity in the upcoming years and they were used as a basic concept for many other models. In most recent works, a similar regression function is used, enriched with lagged returns as control variable. In addition, focus is given on modelling the volatility of the residuals, by estimating a secondary function which includes any factor other than the surprises that can explain the daily volatility of the prices. This procedure resembles the GARCH model. However, the secondary equation does not only contain lags of the residuals of the initial function (ARCH part) and/or lags of the dependent variable itself (GARCH part) but any other factor that can influence the daily volatility. Hence, the initial equation measures the conditional mean change of the price while the secondary the conditional volatility.

Andersen, Bollerslev, Diebold and Vega are the first who proposed such an approach. In Andersen, Bollerslev, Diebold and Vega (2003) they apply it for examining the effect of certain US and German announcements on a wide range of foreign exchange rates. Their initial model has the form:

$$R_t = \beta_0 + \sum_{i=1}^{I} \beta_i R_{t-i} + \sum_{k=1}^{K} \sum_{j=0}^{J} \beta_{kj} S_{k,t-j} + \varepsilon_t \tag{12}$$

where

R_t is the five-minute log return,

R_{t-i} is the five-minute log return, lagged by i,

$S_{k,t-j}$ is the standardised surprise (as defined by Balduzzi et al. 2001) from announcement *k* on time *t-j*. For $j = 0,\ S_{k,t}$ is the surprise on the actual time that it occurred. For $j \neq 0,\ S_{k,t}$ is the surprise *j* lags before.

i is the number of lags for the returns,

j is the number of lags for the surprises,

k is the number of announcements in the sample,

t the observation points.

In contrast to the Balduzzi et al. (2001) model, Andersen et al. (2003) use a time-series of five-minute log returns as dependent variable, not thirty-minute absolute returns calculated when the announcements occurred. High-frequency returns are important for a market which is proven to be very volatile and quickly adjustable to new information. In addition, they do not only take into consideration simultaneously announced news, but also the fact that the effect from the news might last longer than five minutes. For this reason, they include lagged surprises. Lastly, after performing Akaike's and Schwarz's information criteria, they apply five lags on the returns and two lags on the surprises (*I=5*, *J=2).*

Since their initial model is characterised by homoscedasticity, they estimate it by using a weighed least squares (WLS) method. Basic precondition for performing a WLS is the determination of the volatility of the residuals, so that it can be used for the weights. Hence, they firstly estimate the regression (1) by using a normal least square method (OLS) and then they use the estimated residuals for defining their five-minute volatility as follows:

$$|\hat{\varepsilon}_t| = c + \psi \frac{\widehat{\sigma}_{d(t)}}{\sqrt{288}} + \sum_{k=1}^{K} \sum_{j'=0}^{J'} \beta_{kj'} \left|S_{k,t-j'}\right| + (\sum_{q=1}^{Q} (\delta_q \cos\left(\frac{q2\pi t}{288}\right) + \varphi_q \sin\left(\frac{q2\pi t}{288}\right)) + \sum_{r=1}^{R} \sum_{j''=0}^{J''} \gamma_{rj''} D_{r,t-j''}) + u_t \tag{13}$$

where

$\hat{\varepsilon}_t$ is the estimated residual from equation (12) in the five-minute interval *t,*

$\hat{\sigma}_{d(t)}$ is the calculated daily volatility of the asset price for the day that includes the five-minute interval *t,*

$S_{k,t-j'}$ is the standardised surprise as defined above,

$D_{r,t-j''}$ is a dummy variable which takes the value 1 when one of the following happens during five-minute interval *t*:

- Japanese market opening time
- Japanese lunch time
- US late afternoon time during daylight saving period.

With this model, Andersen et al. (2003) suggest that the volatility of the disturbance factor consists of three parts:

- intraday volatility,
- calendar effect patterns and
- news effects.

For capturing the intraday volatility patterns they calculate daily volatility $\hat{\sigma}_{d(t)}$ by using a GARCH(1,1) model, as suggested by relevant literature (e.g. Bollerslev et al (1992), Diebold, Lopez (1995)). The term is divided by $\sqrt{288}$ to account for the five-minute volatility (each trading day in their sample consists of 288 five-minute intervals). In addition, they use a set of dummy variables to take into consideration all the above mentioned external factors that might affect the intraday volatility.

In order to allow for calendar effects they firstly use a Fourier term with trigonometric functions which captures any periodic patterns of each day (Fourier flexible form as defined by Gallant (1981)). The period used is 288 to capture the effects on the five-minute-interval level, while the parameter q is chosen to be 4, as determined by using Akaike and Schwarz information criteria.

Lastly, they consider news effects by including the standardized surprises and their lagged values. In contrast to equation (12), the lags in this equation are a result of experimentation.

Thus, they end up with $J'=12$ for the surprises, $J''=6$ for the Japanese opening dummy, $J''=0$ for the Japanese lunch time and $J''=60$ for the US late afternoon time.

Andersen et al. (2003) results will not be presented in detail; however their major findings are worth mentioning. Firstly, they observe that the majority of the announcements cause a rapid effect on the exchange rates. There is always a large jump on the returns during the first five minutes after the release, which is followed by smaller changes. Secondly, they find that scheduled announcements affect the prices more than the non-scheduled ones. For example, they notice that US macroeconomic releases, which are always announced on specific day and time, influence the DM/$ rate more than the German releases which, in their sample, are announced on certain days but on random times. This is plausible, since unforeseen announcements need some time to be processed by the market participants and be reflected in the prices. Hence, it might take more than 10 minutes, the timeframe after the announcement that this model uses. In contrast, the volatility of the exchange rates takes longer time to adjust. The model estimates that it takes up to one hour, as derived from the coefficients of the twelve lags of the surprises in the equation (13).

Although the above model approximates the volatility patterns of the exchange rates very well, it does not give a clear picture of the sole effect of each individual announcement. This is due to the fact that the biggest part of the change in returns is explained either by the lagged returns or the simultaneous announcements. Thus, Andersen et al. (2003) use in addition a simpler model, which focus on each announcement separately. More specifically, they estimate the following equation for each one of them, using only the observations at the actual time of the announcement:

$$R_t = \beta_k S_{kt} + \varepsilon_t \tag{14}$$

where

R_t is the five-minute return on the time of announcement t,

S_{kt} is the standardised surprise from announcement k on time t.

As a result, they find that more announcement surprises are associated with the exchange rate change and that the coefficient of determination R^2 is quite significant for the majority of them. This last observation shows that the standardised surprise alone is sufficient to explain the change in rates in such a short period of time after the release.

Later on, Andersen, Bollerslev, Diebold and Vega (2007) use a similar model to study the effects of US macroeconomic announcements on a wide range of markets, including foreign exchange rate markets, stock markets and bond markets. Similarly to older studies, they use futures as a close substitute of the underlying asset or rate. In addition, they do not only focus on US markets, but also on U.K., German and European markets. To the best of the author's knowledge, this is the first study that examines the European stock market. Hence they make use of EUROSTOXX 50 future contracts, the instrument used in the empirical analysis of this present paper. In total, they examine nine markets and twenty-five announcements. Lastly, they are interested in the markets reaction within the business cycle. Hence, they divide their sample into two parts: expansion years and recession years. Then they apply the model three times, firstly using the whole sample, and then using the expansion and the recession sample.

The model they conduct is slightly modified in comparison to that of Andersen et al. (2003), in order to account for interrelations between the markets. The main equation has the following form:

$$R_t^h = \beta_0^h + \sum_{h'}^{H} \sum_{i=1}^{I} \beta_{hi}^{h'} R_{t-i}^{h'} + \sum_{k=1}^{K} \sum_{j=0}^{J} \beta_{kj}^{h} S_{k,t-j} + \varepsilon_t^h \, , \quad t = 1, \ldots., T \qquad (15)$$

where

R_t^h is the five-minute log return of the asset *h* on time *t,*

$R_{t-i}^{h'}$ is the five-minute log return of any other asset *h´* on time *t,*

$S_{k,t-j}$ is the standardised surprise from announcement *k* on time *t-j,*

h is the indicator of the asset in question,

h´ is the indicator of all the other assets, with *H=9,*

k is the indicator of each announcement, with *K=25,*

i is the number of lags of returns,

j is the number of lags of surprises.

Using again Schwarz and Akaike criteria, they conclude to *I=2* and *J=3.* The regression is conducted for each of the nine asset classes. The sample contains five-minute log returns within a timeframe of one hour and forty minutes, starting ten minutes before each release. This window is long enough to capture the effects and the price-adjustment patterns of each market, as well as their interrelations. The size and the sign of $\beta_{hi}^{h'}$ show how a change in the price of asset h´ on each lag affects the price of the asset h.

Since the equation (15) is characterized by heteroscedasticity, Andersen et al. (2007) estimate it firstly by an ordinary least squares method (OLS). Then, by using the estimated residuals to define their volatility function, they conduct a WLS. The volatility function of the residuals consists of three parts, each one capturing a different effect, i.e. ARCH effects, intraday volatility effects and announcement effects. The formal representation of it is as follows:

$$|\hat{\varepsilon}_t^h| = \sum_{i=1}^{I'} \beta_{hi} |\hat{\varepsilon}_{t-i}^h| + \sum_{d=1}^{D} \gamma_d D_d + \sum_{k=1}^{K_d} \sum_{j'=0}^{J'} \gamma_{kj'}^h D_{k,t-j'} + u_t^h \qquad (16)$$

where

$\hat{\varepsilon}_t^h$ is the estimated residuals from equation (15) for the asset *h,*

$\hat{\varepsilon}_{t-i}^h$ is the estimated residuals, lagged by i ,

D_d is a dummy variable, which takes the value 1 for each five-minute interval within a day. Hence *D=38.*

$D_{k,t-j'}$ is a dummy variable, which takes the value 1 when announcement *k* occurs in the interval *t-j,*

i is the lag number of the residuals,

j is the lag number of the announcement-dummy.

After experimenting with different lagged values, the authors conclude that *I′=9* and *J′=14,* implying that the ARCH effect on volatility lasts up to forty minutes, while the announcement effects up to seventy minutes. They mention also that they conducted the WLS using different lag values as well as introducing standardized surprises to the volatility equation, but this form gave the best fit. This indicates that the sole occurrence of the announcement is enough to boost the volatility of the assets, irrespective of the presence or magnitude of the surprise.

Although Andersen et al. (2007) do not provide detailed results due to the complexity of the model, they summarize their main findings. To summarize the main points, they find that for all markets the effect of the announcements is larger over the first five minutes after the release, as it can be observed by the larger values of the respective coefficients in the first five-minute interval. This outcome is especially seen in the bond markets, where the effect is large and present irrespective of the phase of the business cycle. In the stock markets, the situation is a little different. When examining the results from the full sample, it seems that news about macroeconomic fundamentals do not affect the prices considerably. However, a closer look at the results from the expansion and contraction sample reveal that during the first phase negative surprises lead to positive reaction on prices, while during the second phase, negative surprises generate negative reaction. This explains the low effect when using the full sample. The effects from both phases cross out each other on average. As far as the exchange rate markets are concerned, the effect depends on the kind of the news. In general, the results here are not different from the ones under Andersen et al. (2003). Lastly, in the end of their paper, the authors investigate further the interdependencies between the different markets, but these results will not be presented in the present paper, since they are out of the scope of this research.

To finalize the present chapter, it is worth presenting the research from Rühl and Stein (2014). This is one of the few papers focusing solely on European stock market, although the macroeconomic announcements in question are American. In addition, it might be the only paper not using the MMS International Survey for estimating market's expectation, but the Financial Times Economic Calendar. As a proxy for the stock market, they consider the stocks that constitute EUROSTOXX 50 index and create a "self-made" index by using their mid-prices and weighting them based on their market capitalization. Lastly, they use a slightly different model than the one suggested by Andersen et al. (2003). Although the conditional mean equation is the same, they modify the conditional volatility function by adding a dummy variable for capturing the days-of-the week effects, as well as one for capturing the effects of the opening of US markets. In short, their model has the form:

$$R_t = \beta_0 + \sum_{i=1}^{I} \beta_i R_{t-i} + \sum_{k=1}^{K} \sum_{j=0}^{J} \beta_{kj} S_{k,t-j} + \varepsilon_t \tag{17}$$

$$|\hat{\varepsilon}_t| = c + \psi \frac{\hat{\sigma}_{d(t)}}{\sqrt{T}} + \sum_{k=1}^{K} \sum_{j=0}^{J} \beta_{kj}' \left|S_{k,t-j}\right| + \left(\sum_{q=1}^{Q} \left(\delta_q \cos\left(\frac{q2\pi t}{T}\right) + \varphi_q \sin\left(\frac{q2\pi t}{T}\right)\right) + \sum_{r=1}^{4} a_r D_r + a_5 NY\right) + u_t \tag{18}$$

where

D_r is a dummy variable which takes the value 1 for the days Monday to Thursday,

NY is a dummy which takes the value 1 when the US markets open

and the rest of the variables are specified as in Andersen et al. (2003).

Rühl and Stein (2014) estimate the above model by using WLS and they compare their results with the ones using Heteroscedasticity and Autocorrelation Consistent (HAC) coefficients method[3]. They report that quality-wise the results are the same, but HAC gives more significant coefficients. They proceed with WLS however, since it gives the possibility to analyse not only the conditional mean returns but also their volatility.

Concerning their results, they find that the index prices show strong autocorrelation patterns and that 12 out of 22 US announcements significantly affect the European market. Among them, there are indicators such as the Institute of Supply Management (ISM)

[3] The HAC method is used in the empirical analysis of this thesis, and it is described in detail in section 3.

Manufacturing Index, Retail Sales, and Home Sales. In general, positive surprises on indicators that show economic growth in the USA tend to have a positive effect on the European stock market. This implies that the robustness of the US economy is perceived as an indicator for the robustness of the global economy.

3. Empirical analysis

Continuing Rühl and Stein (2014) work, the third part of this paper studies how macroeconomic announcements affect the European stock market, with focus on releases of macroeconomic fundamentals in the Eurozone area. In particular, it examines: which indicators most affect prices, what is the magnitude and the direction of the effects and how quickly the prices incorporate the new information. The nature of the available data and the size of the dataset require a combination of well-established econometric methods, as well as newly introduced programming techniques.

The following subsections describe the data selection process and the methodology. The respective results are presented and discussed in section four.

3.1 Data selection

Before addressing the model itself, an overview of the dataset is provided. In order to study the movements of European stock market, I selected the EUROSTOXX 50 ® Index Future (FESX), which underlines the EUROSTOXX 50 ® Index (SX5E). This index consists of stocks of the 50 largest Eurozone companies in terms of market capitalization and is updated every year. Table 1 shows its composition as of September 2014.

Based on the literature (e.g. Ederlington and Lee (1993), Andersen et al. (2003, 2007)), futures can be used as a proxy for the underlying asset and their results can be generalized for the asset itself. In this case, they are preferred to the index itself, since they are heavily traded, thus incorporating any new information easily. In addition their trading hours are longer, hence covering the announcement time of US indicators. FESX are traded on EUREX Exchange between 8:00 and 22:00 CET, while most of the US scheduled macroeconomic announcements occur at 16:00 CET. The sample for the empirical analysis consists of intraday data from January 2008 to December 2012, including the trading price per second for each trading day. This leads to an initial sample of 64.108.800 observations (1.272 trading days × 50.400 seconds.).

Concerning the macroeconomic announcements, several online databases exist, which give information about scheduled releases, such as the Financial Times and Bloomberg. These websites inform investors about upcoming macroeconomic announcements of every country, indicating the date and time of the release, the previous value of the respective macroeconomic indicator, as well as a prediction about its upcoming value. This feature is known as economic calendar. Since these websites are widely used by traders, their forecasts can be presumed as the investors' expected value of the indicators. In addition, most of these websites contain information about past announcements, showing the released values at that time.

For the following analysis I selected the economic calendar of eltee.de. Eltee is a German website with real-time Futures trading information and its calendar is the largest calendar available online, containing information from February 2008. The Financial Times' calendar for example only contains information from 2012. For consistency reasons, I compared Eltee's forecasts with those of Financial Times and Bloomberg, but I did not find any significant deviations. In most of the cases the predictions match, leading to the assumption that such sites use more or less the same methodology.

Having in mind the composition of the EUROSTOXX Index, I used only those announcements which concern the macroeconomic fundamentals of Eurozone, Germany and France. I included German and French announcements for two reasons: The first one is straightforward, since the Index mostly consists of French and German stocks. The second one is mainly intuitive. During the years of 2008–2012 these two countries proved to be the largest and most stable economies in the Eurozone and undertook the task of guaranteeing the stability in the whole area. In case of any event questioning the Euro's strength, Germany and France would be the first to step in and bear the risk. This led to a popular opinion that they (especially Germany) were the ones who dominated Eurozone policy. Hence, it is reasonable to assume that investors pay attention to announcements of German and French indicators to inform themselves about Eurozone prosperity.

In addition, I included some US Macro-announcements, since there are studies which have proven that European and UK markets are indeed affected by them. Rühl and Stein (2014), for example, examined which US announcements affect the European stock markets. Among others, they used the EUROSTOXX 50 Index for their analysis. Since this is the underlying asset of the future in scope, some of these announcements might also affect the future price. For simplicity reasons, I selected only those which Rühl and Stein found to be statistically significant at 1% confidence level. However, data were available only for ISM Manufacturing, Retail Sales, Industrial Production and Home Sales.

Table 2 describes the final selected announcements, grouped in two categories: the ones that have a direct link to the prices (e.g. CPI, interest rates) and the ones that indicate the status of the overall economy (e.g. GDP, PMI etc.). Most of them are released monthly, while six of them-Eurozone CPI, Eurozone manufacturing Purchasing Managers' Index (PMI), Eurozone services PMI, German CPI, German manufacturing PMI, German services PMI-are announced every two weeks. This class of announcements is a special case. About two weeks prior to the official monthly release, a preliminary value of the indicator is disclosed. However, in the model they are both treated as actual announcements. The intuition is that these preliminary values are also taken into consideration by the investors, since they are marked as important in every trading website. Lastly, there are two indicators–the French manufacturing PMI and the French services PMI-which were announced monthly until May 2010 and April 2010, respectively, and biweekly afterwards.

What should also be noticed is the time of the announcement. Most of them have a fixed release time. Nevertheless, for some of them (see table 2) the release time varies, although it is always publicly known at least one week ahead. Lastly, in some cases the announcement time for German CPI is tentative and not known in advance. These observations are excluded from the sample. As it will be explained in the next section, without knowing the exact time of the announcement, it is not possible to observe its effect on the future prices.

Last but not least, some of the announcements are released at the same day and time. Table 3 shows how many times each announcement has overlapped with each of the rest. The diagonal shows the number of times the announcement itself has occurred; hence this is the total number of observations for this announcement.

All the above specifications led to a total announcements' sample of 1996 observations from 2008–2012.

Table 1: Composition of EUROSTOXX 50 Index

No.	Name	Sector	Country
1	AIR LIQUIDE	Chemicals	FR
2	AIRBUS GROUP NV	Industrial Goods & Services	FR
3	ALLIANZ	Insurance	DE
4	ANHEUSER-BUSCH INBEV	Food & Beverages	BE
5	ASML HLDG	Technology	NL
6	ASSICURAZIONI GENERALI	Insurance	IT
7	AXA	Insurance	FR
8	BASF	Chemicals	DE
9	BAYER	Chemicals	DE
10	BCO BILBAO VIZCAYA ARGENTARIA	Banks	ES
11	BCO SANTANDER	Banks	ES
12	BMW	Automobiles & Parts	DE
13	BNP PARIBAS	Banks	FR
14	CARREFOUR	Retail	FR
15	DAIMLER	Automobiles & Parts	DE
16	DANONE	Food & Beverages	FR
17	DEUTSCHE BANK	Banks	DE
18	DEUTSCHE POST	Industrial Goods & Services	DE
19	DEUTSCHE TELEKOM	Telecommunications	DE
20	E.ON	Utilities	DE
21	ENEL	Utilities	IT
22	ENI	Oil & Gas	IT
23	ESSILOR INTERNATIONAL	Healthcare	FR
24	GDF SUEZ	Utilities	FR
25	GRP SOCIETE GENERALE	Banks	FR
26	IBERDROLA	Utilities	ES
27	Industria de Diseno Textil SA	Retail	ES
28	ING GRP	Banks	NL
29	INTESA SANPAOLO	Banks	IT
30	L'OREAL	Personal & Household Goods	FR
31	LVMH MOET HENNESSY	Personal & Household Goods	FR
32	MUENCHENER RUECK	Insurance	DE
33	NOKIA	Technology	FI
34	ORANGE	Telecommunications	FR
35	PHILIPS	Industrial Goods & Services	NL
36	REPSOL	Oil & Gas	ES
37	RWE	Utilities	DE
38	SAINT GOBAIN	Construction & Materials	FR
39	SANOFI	Healthcare	FR
40	SAP	Technology	DE
41	SCHNEIDER ELECTRIC	Industrial Goods & Services	FR
42	SIEMENS	Industrial Goods & Services	DE
43	TELEFONICA	Telecommunications	ES
44	TOTAL	Oil & Gas	FR
45	UNIBAIL-RODAMCO	Real Estate	FR
46	UNICREDIT	Banks	IT
47	UNILEVER NV	Personal & Household Goods	NL
48	VINCI	Construction & Materials	FR
49	VIVENDI	Media	FR
50	VOLKSWAGEN PREF	Automobiles & Parts	DE

Table 2: Description of selected macroeconomic announcements (observations 2008-2012)

No.	Macroeconomic Indicator	Frequency	Regular time	Observations
	Price related			
1	ECB Interest rate	Monthly	13:45	59
2	Eurozone Consumer Price Index (CPI)	Every 2 weeks*	11:00	112
3	French Consumer Price Index(CPI)	Monthly	08:45**	60
4	German Consumer Price Index (CPI)	Every 2 weeks*	Usually 08:00 or 14:00	104
5	German Producer Price Index (PPI)	Monthly	08:00	59
6	USA Consumer Price Index (CPI)	Monthly	14:30	57
	Economic - state related			
7	Eurozone GDP	Monthly	11:00	58
8	Eurozone Indicator of Economic Sentiment ZEW	Monthly	11:00	59
9	Eurozone Industrial production	Monthly	11:00	58
10	Eurozone Purchasing Managers Index (PMI)	Every 2 weeks*	10:00	117
11	Eurozone Retail sales	Monthly	11:00	58
12	Eurozone Service Purchasing Managers Index (PMI)	Every 2 weeks*	10:00	115
13	Eurozone Unemployment rate	Monthly	11:00	57
14	French Consumer Spending	Monthly	08:45	55
15	French GDP	Quarterly	Either 07:30 or 08:45	25
16	French manufacturing Purchasing Managers Index (PMI)	Monthly until 05/2010 - Every 2 weeks* afterwards	Either 09:00 or 09:50	78
17	French service Purchasing Managers Index (PMI)	Monthly until 04/2010 - Every 2 weeks* afterwards	Either 09:00 or 09:50	80
18	German GDP	Quarterly	08:00	38
19	German Gfk consumer climate index	Monthly	08:00	58
20	German Ifo business climate index	Monthly	10:00	59
21	German Industrial Production	Monthly	12:00	59
22	German Manufacturing Purchasing Managers Index (PMI)	Every 2 weeks*	Usually 09:30 or 09:55	90
23	German Retail Sales	Monthly	08:00	59
24	German Services Purchasing Managers' Index (PMI)	Every 2 weeks*	Usually 09:30 or 09:55	87
25	German unemployment rate	Monthly	09:55	51
26	USA Home Sales	Monthly	16:00	59
27	USA Industrial production	Monthly	15:15	58
28	USA ISM Manufacturing Purchasing Managers' Index (PMI)	Monthly	16:00	58
29	USA Retail Sales	Monthly	14:30	58

Table 3: **Number of times that the announcements coincide (observations 2008-2012)**

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1	59				4												9								1				
2		58					7																						
3			38										4			1		3											
4				25				5	1																				
5	4				58												4					6			5		2		
6						59							1																
7		7					57																						
8				5				59	1																				
9				1				1	55																				
10										115	84	9																	
11										84	114	9																	
12										9	9	59																	
13			4			1							58																
14														59															
15															51				1		2								
16			1													58													
17	9				4												112					4			6		34		
18			3															104											
19															1				90		61								
20																				58									
21															2				61		87								
22					6												4					58					4		
23																							58						
24																								58					
25	1				5												6								58				
26																										58			
27					2												34					4					57		
28																												78	55
29																												55	80

1 **EU* IndicatorEconSentiment ZEW**
2 **US RetailSales**
3 **DE GDP**
4 **FR GDP**
5 **EU* GDP**
6 **DE PPI**
7 **US CPI**
8 **FR CPI**
9 **FR ConsumerSpending**
10 **EU* ManufacturingPMI**
11 **EU* ServicesPMI**
12 **DE Ifo BusinessClimateIndex**
13 **DE Gfk ConsumerClimateIndex**
14 **US HomeSales**
15 **DE UnemploymentRate**
16 **DE RetailSales**
17 **EU* CPI**
18 **DE CPI**
19 **DE ManufacturingPMI**
20 **US ManufacturingPMI**
21 **DE ServicesPMI**
22 **EU* RetailSales**
23 **ECB InterestRate**
24 **DE IndustrialProduction**
25 **EU* IndustrialProduction**
26 **US IndustrialProduction**
27 **EU* UnemploymentRate**
28 **FR ManufacturingPMI**
29 **FR ServicePMI**

*EU here stands for Eurozone

3.2 Methodology

3.2.1 Identifying the effects of announcements and the speed of price adjustment

The variation of release time for some indicators makes it impossible to apply a model such as the one proposed by Ederlington and Lee (1993). As described in section 1, Ederlington and Lee's (1993) model requires fixed announcement times through the whole sample. It seems that an approach similar to the one used by Balduzzi, Elton, and Green (2001) is more appropriate for this sample, i.e. regressing price changes on the surprise created by the announcement.

As mentioned in section 2, Balduzzi et al. (2001) define the "standardized" surprise measure as $S_i = \frac{A_i - F_i}{\sigma_i}$, where A_i the value of announcement i, F_i the median of the MMS forecast survey and σ_i the standard deviation of $A_i - F_i$ across all observations. This paper uses a similar but simpler approach for measuring the surprise of each announcement.

The surprise measure is defined as

$$S_i = \frac{A_i - F_i}{|F_i|} \tag{19}$$

where

A_i is the actual value of the announcement i and

F_i the forecasted value of the announcement i, as taken from Eltee.

This expression can be seen as a percentage deviation from the expected value of an announcement. In this way it is possible to compare announcement surprises of indicators that are measured in different units. In addition, by using the absolute value of F_i in the denominator, the sign of surprise S is preserved and, as it will be explained later, the direction of the effect can be estimated. A negative surprise means that the released value is lower than expected, while a positive that it is higher than expected. However, a positive surprise does not necessarily mean that the actual value of the macroeconomic indicator is positive. For example, consider the case when the expected value for the upcoming GDP release is, say, -1.5% (compared to previous quarter) and the actual GDP is announced to be -1.0%. The surprise in this case will be $S_i = \frac{-0.01+0.015}{0.015} \cong 0.33 \text{ } or \text{ } 33\%$. Intuitively this means that the economy is not as bad as expected and this is a positive surprise for the investors.

In contrast to Balduzzi et al (2001), where 35-minutes absolute returns are used, this model uses logarithmic returns of different time frames. Since the sample consist of a continuous variable (tick prices), logarithmic returns are more suitable. In addition, the 35-minutes time frame is too large for such a liquid assets as futures. The existing literature shows that future prices are adjusted within the first three minutes after the announcement - see for example ap Gwilym, Buckle, Clare and Thomas (1998). Based on that, I use six different time intervals for the calculation of the returns, namely: 30 sec., 1 min., 1,5 min., 2 min., 2,5 min., 3 min. However, if any announcement seems to have a longer effect on the returns, the analysis is continued up to nine minutes after the announcement time.

Following Andersen et al (2007) methodology for identifying the sole effects of each release, the following equation is estimated, for each interval t and each announcement *k*:

$$R_{tki} = a_0 + a_k * S_{ki} + \varepsilon_i \quad (20)$$

where

$R_{tki} = \log\left(\frac{P_{t_1}}{P_{t_0}}\right)$ is the return on the inteval *t* after the announcement,

t_0 is the time of the announcement *k,*

$t_1 \epsilon \{30, 60, 90, 120, 150, 180\}$ – seconds after the announcment *k,*

S_{ki} is the surprise of announcement *k,*

i the number of observations.

In addition, for each regression, I use only those returns which were realized at the time of announcement *k*. For example, when running the regression for the ECB interest rate, only the returns realized on the announcement day at time 13:45 are used. This gives 58 observations for this regression, and so on. Moreover, I exclude those announcements that occurred during a bank holiday, since there was no trading in place and the impact of the surprise on the prices cannot be measured. One could argue that the impact could be measured on the opening price of the day after, but this would lead to biased estimates, firstly because it will insert in the model the effects of the market opening after a holiday and, secondly, because it will make the sample inconsistent. All the other observations measure the effect on the returns immediately after the announcement. Adding one or more observations which measure the effect in a different time frame will make the model unreliable.

By estimating the above model, it is possible to discover which macroeconomic announcements mostly affect prices, the magnitude and direction of the effects, as well as the speed that the prices adjust to the new equilibrium. In particular, the coefficient a_k shows both the importance and the effect of announcement i. If a_i is statistically significant (as indicated by a t-test with $a_k = 0$ as null hypothesis) then there is a relation between the change in prices and the surprise from the announcement. This intuitively means that the investors consider this announcement as important and react by buying or selling. On the other hand, if no significance is found, then the price change is not related to the surprise, meaning that this announcement is not considered as important.

Furthermore, the size and the sign of the coefficient a_k show the magnitude and the direction of the effect. Consider for example the case where the sign of a_k is negative. If the surprise is positive (i.e. the macroeconomic indicator happens to be larger than expected), then the returns will be negative, indicating a price reduction right after the announcement–negative effect. A positive surprise results in an exactly opposite effect. The following table summarises all of the possible effects of surprises.

Table 4: Different effects of the surprises, based on their coefficient

$\alpha_i<0$

Surprise			Returns			Price	
Larger than expected	$S_i>0$	$\Longrightarrow$	Negative	$R<0$	$\Longrightarrow$	Reduced	$P_{t1}<P_{t0}$
Lower than expected	$S_i<0$	$\Longrightarrow$	Positive	$R>0$	$\Longrightarrow$	Increased	$P_{t1}>P_{t0}$

$\alpha_i>0$

Surprise			Returns			Price	
Larger than expected	$S_i>0$	$\Longrightarrow$	Positive	$R>0$	$\Longrightarrow$	Reduced	$P_{t1}>P_{t0}$
Lower than expected	$S_i<0$	$\Longrightarrow$	Negative	$R<0$	$\Longrightarrow$	Increased	$P_{t1}<P_{t0}$

Lastly, by using different time frames for calculating the returns, it is possible to identify the time when the prices will have adjusted to the new equilibrium level. The assumption is that, as time passes by, the effect of the surprise will fade and at some point there will be no relation between the returns and the surprise. Hence, after some time the coefficient a_k will not be significant any more.

3.2.2 Case of simultaneous announcements

As mentioned in section 3.1, some announcements coincide in time. Consider for example Eurozone Manufacturing PMI. It coincides 89 times with Eurozone Services PMI, i.e. 77% (= 89/115) of the time, and 9 times with German Ifo Business Climate Index, i.e. 7,8% of the time (Table 3). One can therefore argue that the price change that occurs during the announcement of Eurozone Manufacturing PMI is also due to the co-release of the other two announcements. In this case, the true coefficient a_k might be different than that resulting from the simple regression described above. For this reason, these cases are treated differently.

In order to reduce complexity, for those announcements that coincide at least 30% of the time, I apply a multivariate linear regression. Following the logic of Balduzzi et al (2001), this regression has the form:

$$R_{tki} = a_0 + a_k S_{ki} + \sum_{k'=1}^{K'} a_{k'} * S_{k'i} + \varepsilon_i \tag{21}$$

where R_{tki}, S_{ki} are defined as in section 3.2.a.

Moreover, k' is the index of the announcement that coincides at least 30% of the time with announcement k, while K' is the total number of those announcements. If in a certain time interval the announcement k' did not coincide with k then $S_{k'i}$ takes the value 0. As an example, consider once more the case with Eurozone Manufacturing PMI. German Ifo Business Climate Index is not included in the multilinear regression, since it coincides only 7,8% of the time. This regression has 115 observations.

In total, four such multilinear regressions are conducted for the following announcement pairs:

- Eurozone Manufacturing PMI & Eurozone Manufacturing Services PMI
- Eurozone Unemployment Rate & Eurozone CPI
- German Manufacturing PMI & German Services PMI
- French Manufacturing PMI & French Services PMI

Similarly to the procedure of individual regressions, I calculated the returns on different time frames, in order to measure the duration of the effect. The analysis starts from thirty seconds up to three minutes after the announcement. In case the effect is still preserved in the third minute, the analysis continues up to nine minutes.

Detailed results of this process are described in section 4.

3.2.3 Considering autocorrelation of returns

It is known from the literature that individual stock returns as well as stock indices returns show not only daily but also intraday autocorrelation patterns for reasons such us irrationality of some investors, information asymmetries, transaction costs etc. For the same reasons, autocorrelation can appear in future prices. In addition, future prices are strongly correlated to the underlying stock prices, and hence follow the same patterns. Ahn et al. (1999), for example, have shown that, for 15 different indices worldwide, autocorrelation of returns and the respective futures returns exist.

In order to ensure that the change in prices after each significant announcement occurs only due to the new information and not because of "natural" patterns that the prices follow, I also take into consideration autocorrelation of returns. In the existing literature, it is usual to model 5-minute asset returns as depended on their lagged values and on the announcement surprises, such as the model introduced by Andersen et al. (2003). Since the present study concentrates on high-frequency analysis, I used 1-minute returns for studying the autocorrelation patterns of the market.

At first I created a time series consisting of 1-minute returns for the period covering the first and the last announcement in the sample, namely from 12.02.2008 11:00:00 to 28.12.2012 08:45:00. This results in a time series of 1.038.506 points. Secondly, I estimated autocorrelation models AR(p) of the order 1 up to 6 and compared them using the Akaike's information criterion (AIC) (Akaike, 1974). The AR(p) models have the form:

$$\boldsymbol{AR(p)}:\ R_t = \beta_0 + \sum_{i=1}^{p} \beta_i * R_{t-i} + \varepsilon_t \tag{22}$$

where

R_t the 1-minute return at time t and

p the number of lags.

The AIC is used extensively in the literature to assess the quality of a model. Formally AIC is defined as:

$$AIC = 2k - 2\ln(L) \tag{23}$$

where

k is the number of the estimated parameters and

L the maximum value of the likelihood function in the model.

The lower is the AIC the better the model.

Table 5 shows the results of each AR model: the estimators for the intercept β_0 and each coefficient β_i, their standard errors, their t- and p-values of the Student t-test for the null-hypothesis $\beta_i = 0$ against the alternative $\beta_i \neq 0$ and their significance codes, as described in the end of the table. For each model, the respective ACI value is shown as well. As it can be seen, the first 5 lags are statistically significant, meaning that the 1-minute price change at each time is correlated to the price change up to 5 minutes earlier. In addition, AR(4) and AR(5) are the models with the lowest AIC. However, since the fifth lag is less significant than the previous four and so as to keep complexity low, four lags are sufficient for the following analysis.

Table 5: Autoregressive analysis of returns

AR(1)	Estimate	Std. Error	t value	Pr(>|t|)		AIC
lag1	-63.86	0.97	-66.00	0.00	***	-14,316,432
intercept	0.00	0.00	-0.64	0.53		

AR(2)	Estimate	Std. Error	t value	Pr(>|t|)		AIC
lag1	-0.06	0.00	-66.59	0.00	***	-14,316,557
lag2	-0.01	0.00	-11.32	0.00	***	
intercept	0.00	0.00	-0.64	0.52		

AR(3)	Estimate	Std. Error	t value	Pr(>|t|)		AIC
lag1	-0.06	0.00	-66.64	0.00	***	14,316,571
lag2	-0.01	0.00	-11.56	0.00	***	
lag3	0.00	0.00	-4.18	0.00	***	
intercept	0.00	0.00	-0.65	0.52		

AR(4)	Estimate	Std. Error	t value	Pr(>|t|)		AIC
lag1	-0.06	0.00	-66.67	0.00	***	-14,316,639
lag2	-0.01	0.00	-11.66	0.00	***	
lag3	0.00	0.00	-4.71	0.00	***	
lag4	-0.01	0.00	-8.42	0.00	***	
intercept	0.00	0.00	-0.65	0.52		

AR(5)	Estimate	Std. Error	t value	Pr(>|t|)		AIC
lag1	-0.06	0.00	-66.69	0.00	***	-14,316,644
lag2	-0.01	0.00	-11.67	0.00	***	
lag3	0.00	0.00	-4.74	0.00	***	
lag4	-0.01	0.00	-8.57	0.00	***	
lag5	0.00	0.00	-2.63	0.01	**	
intercept	0.00	0.00	-0.65	0.51		

AR(6)	Estimate	Std. Error	t value	Pr(>|t|)		AIC
lag1	-0.06	0.00	-66.69	< 2e-16	***	-14,316,641
lag2	-0.01	0.00	-11.67	< 2e-16	***	
lag3	0.00	0.00	-4.74	0.00	***	
lag4	-0.01	0.00	-8.58	< 2e-16	***	
lag5	0.00	0.00	-2.65	0.01	**	
lag6	0.00	0.00	-0.49	0.62		
intercept	0.00	0.00	-0.65	0.51		

Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**

Subsequently, I combined the autocorrelation patterns of the returns with the surprises from the macroeconomic announcements under the single model:

$$R_t = \beta_0 + \sum_{n=1}^{4} \beta_n * R_{t-n} + \sum_{k=1}^{29} \alpha_k * S_{kt} + \varepsilon_t \quad (24)$$

where

R_t the 1-minute log returns at timestamp *t,*

S_{kt} the surprise from the announcement *k* in the five-minute interval *t* (as defined in the sections above) and

n the number of lags.

For the estimation of this model, I used the same data frame as for the creation of the timeseries, which results in 1.038.506 observations. In addition, if the announcement k occurs on timestamp *t,* the variable S_{kt} takes the value of the surprise measure; otherwise it takes the value zero. In contrast to Andersen et al. (2003, 2007), I did not include lagged surprises in the model. Due to the size of the dataset, any additional control variable created computational issues. Moreover, the inclusion of the lagged surprises would not give any additional information in this stage of the research. The focus here is not to model the path that returns follow (as in Andersen et al (2003, 2007)), but to examine if the surprise that occurs at the very first minute after the announcement is responsible for the observed price-change or not.

After the first estimation, the model was tested for autocorrelation and heteroscedasticity of the residuals, using the Breusch-Godfrey (1978) and the Breusch-Pagan (1979) tests, respectively.

One of the assumptions of the classical OLS method is that the covariance between the error terms is constant. If, for instance, the equation

$$y_n = \beta_0 + \sum_{k=1}^{K} \beta_k * x_{nk} + \varepsilon_n, \ \forall n \in [1, N] \quad (25)$$

is to be estimated by an OLS, then the assumption implies that for its residuals the following will hold:

$$Var(\varepsilon_k, \varepsilon_m) = 0 \Leftrightarrow E(\varepsilon_k \varepsilon_m) = 0, \ \forall \ k \neq m \ . \quad (26)$$

If this assumption does not hold, then autocorrelation exists between the residuals, which causes problems with the estimators. While they are still linear and unbiased, their variances are biased (Johnston–DiNardo (1997), Greene (1997)). Hence, any statistical testing of the model, which includes variances of estimators, will be unreliable. In this case, since the variances are needed for testing the significance of the coefficients (t-test), I need to test for autocorrelation and correct it, if it is present.

As the sample used for the model (24) includes a time series of 1-minute returns and four lags, I expected an autocorrelation of a high order. The most appropriate test, in this case, is the one developed by Breusch (1978) and Godfrey (1978), because it allows testing for autocorrelation of any order and can be used even if lags of the dependent variable are included in the model. The testing procedure works as follows:

- Firstly, the model (24) is estimated using the OSL method and the residuals are supposed to have autocorrelation of order p, i.e.:

$$\varepsilon_t = \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \cdots + \rho_p \varepsilon_{t-p} + u_t \tag{27}$$

- Subsequently, a secondary regression is estimated, using the standard errors from the first estimation:

$$\hat{\varepsilon}_t = \gamma_0 + \sum_{n=1}^{4} \gamma_n * R_{t-n} + \sum_{i=1}^{29} \delta_k * S_{kt} + \rho_1 \hat{\varepsilon}_{t-1} + \rho_2 \hat{\varepsilon}_{t-2} + \cdots + \rho_p \hat{\varepsilon}_{t-p} + u_t \tag{28}$$

- Lastly, the hypothesis $\boldsymbol{H_0}: \rho_1 = \rho_2 = \cdots = \rho_p = 0$ (i.e. no autocorrelation) is tested using the statistic $(T-p)R^2$, where R^2 is the coefficient of determination of the secondary regression. This statistic under H_0 follows asymptotically the X^2 distribution with p degrees of freedom. The null hypothesis is rejected if $(T-p)R^2 > X_{\alpha}^2$, where α *is* the selected significance level.

Table 6: Results from Breusch-Godfrey test for autocorrelation

Order of autocorrelation	B-G statistic	p-value
p=1	2,17	0,141
p=2	5,25	0,073
p=3	42,19	0,000 ***
p=4	72,15	0,000 ***

I tested the model (24) for autocorrelation of order up to four, using the **bgtest()** method in the statistical package “lmtest” of R-Studio (Zeileis and Hothorn (2002)) and the results can be seen in Table 6. The null hypothesis was rejected for p=3 and p=4, showing autocorrelation of an order higher than three.

Similarly, the model was tested for heteroscedasticity. An OLS-estimated model is said to have heteroscedasticity, when the variance of the residuals, in contrast to the classical assumptions, is not constant, but varies based on the value of the explanatory variables. In other words, this means that in a regression such as (24) the following will hold:

$$Var(\varepsilon_t) = E(\varepsilon_t^2) \neq \sigma^2 \tag{29}$$

Once again, the resulting estimators are linear and unbiased, but their variances not, making them inappropriate for use in statistical testing.

For determining if the model (24) exhibits heteroscedasticity, a Breusch-Pagan test is performed, making use of the following procedure:

- The model is firstly estimated using the OLS method, while the variance of the residuals is supposed to have the form

$$\sigma_t^2 = \gamma_0 + \gamma_1 Z_{t1} + \cdots + \gamma_m Z_{tm} + u_t \tag{30}$$

- Next, a secondary regression is estimated, using the residuals from the first regression:

$$\frac{\hat{\varepsilon}_t}{\sigma_t^2} = \gamma_0 + \gamma_1 Z_{t1} + \cdots + \gamma_m Z_{tm} + u_t \tag{31}$$

  Since the Z's are unknown, the explanatory variables of the model itself were used, as this is a common practice.

- Lastly the statistic $\frac{SSR}{2}$ is used to test the hypothesis $\boldsymbol{H_0}: \gamma_1 = \gamma_2 = \cdots = \gamma_p = 0$ (i.e. the variance is constant). This statistic follows the X^2 distribution with *m* degrees of freedom. H_0 is rejected if $\frac{SSR}{2} > X_\alpha^2$, for a selected significant level α.

The test is performed using the **bptest()** in the statistical package "lmtest" of R-Studio (Zeileis and Hothorn (2002)). In this case, the residuals of the model were found to be heteroscedastic, since the null hypothesis was rejected for significance level α=0,01 (since *B-P=2539,04, p-value=0,00*).

Since heteroscedasticity and autocorrelation is detected in model (24), it requires the estimation of heteroscedasticity and autocorrelation consistent (HAC) coefficients. This has been an issue that has generated a lot of research, dating back to the 1980's (White (1980), White and Domowitz (1984), Newey and West (1987), Andrews (1991) amongst others). The focus of all related literature is to find an efficient way to estimate the variance-covariance matrix of the estimators, when the form of heteroscedasticity and autocorrelation is unknown. Given a regression in the form of (24), the covariance matrix can be expressed as:

$$Var(\widehat{\boldsymbol{\beta}}) = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} \tag{32}$$

$$= (\frac{1}{N}\boldsymbol{X'X})^{-1}\frac{1}{N}\boldsymbol{\Phi}(\frac{1}{N}\boldsymbol{X'X})^{-1} \qquad (33)$$

where

$\widehat{\boldsymbol{\beta}}$ is the estimated vector of coefficients,

$\boldsymbol{X}$ **is** the matrix of the observations of the independent variables and

$\boldsymbol{\Phi} = \boldsymbol{n^{-1}X'\Omega X}$.

When the residuals are homoscedastic and independent then $\boldsymbol{\Omega} = \sigma^2\boldsymbol{I}_n$, leading to the classical OLS $Var(\widehat{\boldsymbol{\beta}}) = \sigma^2(\boldsymbol{X'X})^{-1}$. When not, then an estimator of $\boldsymbol{\Phi}$ needs to be found, which usually has the form:

$$\widehat{\boldsymbol{\Phi}} = \frac{1}{N}\sum_{n,j=1}^{N}\boldsymbol{w}_{|n-j|}\,\widehat{\boldsymbol{V}}_n\widehat{\boldsymbol{V}}_n{}' \qquad (34)$$

where

$\boldsymbol{w}$ is a vector of weights and

$\widehat{\boldsymbol{V}}_n$ is the estimator of the function $V_n(\boldsymbol{\beta}) = x_n(y_n - x'_n\beta)$.

Hence, the estimation of the covariance matrix requires the weights to be specified. Newey and West (1987, 1994), Andrews (1991) and Lumley and Heagerty (1999) have all suggested different approaches to indicate the weights. This paper applies the approach suggested by Andrews, as it is one of the most widely used. It can be also easily employed, since it is a predefined filed of the method **vcovHAC()** in the statistical package "Sandwich" in the R-Studio, developed by Zeileis A. (2004).

In short, Andrews' approach suggests using weights of the form $w_l = K(\frac{l}{B})$, where K() is a kernel function and B determines a bandwidth (smoothing) parameter. While he studies different kernel functions, he concludes that the most efficient one is the quadratic spectral, which gives weights of the form

$$w_l = \frac{3}{z^2}\,(\frac{\sin(z)}{z} - \cos(z)) \qquad (35)$$

where $z = \frac{6\pi}{5}\frac{l}{B}$.

As for the selection of the bandwidth parameter, he suggests the use of AR(1) or ARMA(1,1) process (Andrew 1991, Zeileis 2004). The quadratic spectral function and the bandwidth selection using AR(1) are both used by this paper, as they are defaults in the **vcovHAC()** method.

Having found the HAC covariance matrix, one can use it for testing the significance of the re-estimated coefficients. This is feasible using the method **coeftest()** of the package "lmtest" of R-studio (Zeileis, Hothorn 2002). This method allows the user to choose their own covariance matrix, with which the t-test for estimators' significance is performed. Hence, in this case I used HAC matrix as calculated by **vcovHAC()** method.

The results of this approach and a detailed analysis can be found in the next section. This procedure ensures unbiased estimators and variances, leading to reliable results when it comes to statistical testing. In addition, it is quite simple and it does not require large computational power or time, which is usually the problem when such large samples are involved. Lastly it is worth mentioning that one could approximate the volatility of the residuals by a combination of ARCH or GARCH model with a number of control variables and then perform a WLS method to correct for heteroscedasticity (as in Andersen (2003) for example). I did not choose this approach for two reasons: Firstly, the size of the dataset did not allow for the computation of such a complex volatility function. Secondly, as tested by Rühl and Stein, the two methods give results of the same quality, but HAC focuses more on the coefficients than on the variance of the residuals. Andersen et al. (2003) also comment that HAC is a good alternative if a more detailed analysis of the volatility is not required. Since the goal of this paper is to show that the coefficients remain significant even after the introduction of lagged returns as control variables, WLS is not needed.

4. Results

4.1 Significant announcements – duration of the effect

As explained in section 3, I performed the regression (19) for each announcement that does not coincide with any other, and for each different timeframe up to three minutes after the announcement time. Table 7 shows for each regression the coefficient a_i with its significance level code, the estimated probability of rejecting the null hypothesis $a_i = 0$ of the Student t-Test ($P<|t|$) and the R^2. Table 8 shows the results of the analysis up to nine minutes after the announcement time, only for those indicators whose significance was preserved at the third minute.

As it can be seen not all announcements have an effect on the returns During the first 30 seconds after the announcement, eleven out of twenty-one have a significant coefficient, namely:

- German GDP

- German Ifo Business Climate Index
- German Industrial Production
- German Retail Sales
- German Unemployment Rate
- ECB Interest Rate
- Eurozone Industrial Production
- Eurozone Retail Sales
- Eurozone Unemployment Rate
- US Home Sales
- US ISM Manufacturing

In addition, all coefficients have the expected sign, showing the direction of the effect. For example, a 1% surprise on the ECB interest rate (interest rate higher than expected) leads to a decrease of 1,3% (since $a_i = -0{,}013$) in the future prices. This is consistent with the economic theory, because an increase in interest rates makes borrowing more expensive and lending more profitable. Hence, investors will start to sell. In general, positive surprises on indicators that show prosperity of the overall economy have a positive impact on prices, while positive surprises on indicators that show stagnation have a negative impact on prices.

From an economic perspective, the indicators that affect the prices the most (looking only at the first 30 seconds after the announcement) are the ECB interest rate, German unemployment rate, Eurozone unemployment rate and the US ISM manufacturing index (-1,3%, -1,2%, -1,03% and 1,08% change in the price respectively per 1% surprise). It is also interesting that none of the French macroeconomic indicators are considered important by investors, although France is one of the largest economies in EU and French stocks constitute a large part of the EUROSTOXX Index. This clearly indicates that Germany is considered the leading economy of EU and that investors pay attention to its macroeconomic fundamentals to evaluate European economy.

Moreover, for most of the regressions where significant coefficients were found, R^2 is over 10% and in some cases exceeds 20%. This means that 10–20% of the price change is a result of the surprise in the announcement. Unfortunately, comparisons with existing literature are not possible, since in similar studies the values of R^2 are rarely disclosed. A good point of reference could be the work of Andersen et al (2003), where they use a similar approach for foreign exchange rates. Their R^2 values are on the same levels and they claim that they are high enough, since macroeconomic announcements have a "short-lived impact on exchange rates". Therefore, it is rational to expect similar R^2 values for the futures, because their markets share some similarities with foreign exchange markets. They are both very liquid and both are used for speculation or liquidity purposes. Lastly, in such volatile markets, surprises from macroeconomic announcements cannot be the only factors that lead to price change. Hence, these R^2 values are mostly as expected.

Furthermore, it is interesting to examine how the effect of the surprises and the R^2 values develop in time. Figure 1 shows these developments graphically for all significant indicators. The dotted lines show the development of the p-values over the first three minutes (nine minutes for those where the effect persists in the third minute) using the left axis, while the grey lines show the development of R^2 (right axis).

In general, the effect of the surprise lasts until the moment when the p-value is larger than the significance level 0,01. Until that moment the change in the price is related significantly to the surprise from the announcement. Hence, one can expect low p-values right after the announcements, which increase as time passes. Indeed, for half of the selected indicators (six out of twelve), the p-values increase right after the first 30 seconds and at some point exceed the 0,01 level (see for example the graph of German unemployment rate and ECB interest rate). For some others (four out of twelve) this value decreases during the first 1-1,5 minutes and then rises. This means that the effect initially grows during the first 1,5 minutes and then starts to decline (see for example the graph for German GDP). On the contrary, there are two indicators, namely the German Ifo business climate index and the US ISM Manufacturing index, for which the p-value remains steadily low or fluctuates, but never surpassing the 0,01 significance level. One can assume that the effects of these announcements will last longer than nine minutes. All in all, for the remaining ten indicators the announcement effect lasts between thirty seconds and 2,5 minutes, with the only exception being German GDP, which lasts for five minutes.

4.2 Simultaneous announcements – duration of the effect

As described in section 3.2.2, simultaneous announcements are treated differently, using the multivariate linear regression (20). Tables 9 and 10 show the coefficient a_k and $a_{k'}$ of the simultaneous announcements k and k' with their significance codes, the p-values of the Student t-test for the null-hypothesis $a_k = 0$ and $a_{k'} = 0$ against $a_k \neq 0$ and $a_{k'} \neq 0$ respectively, as well as the R^2 for each multilinear regression. As in tables 7 and 8, these values are shown for different returns, in order to examine the duration of the effect. More precisely, the effect of the announcement stops when the relevant coefficient is not significant anymore. Table 9 shows the duration of the effect for the first three minutes after the announcement. If an announcement is still significant on the third minute, the same procedure continues up to nine minutes, as shown in Table 10.

Out of the eight indicators being examined, three of them seem to be significant, namely:

- EU Manufacturing PMI
- EU Unemployment rate
- German Services PMI

However, their effects show different patterns. The effect of the EU Unemployment rate fades after 1,5 minutes, the one of EU Manufacturing PMI is preserved until the sixth minute, while the effect of German Services PMI becomes significant after the second minute and fades out on the fourth minute.

In addition, for all three of them, the direction of the effect is as expected, as seen by the sign of the coefficient. EU unemployment rate for instance affect the future price negatively, since it shows a downturn in the overall EU economy, which makes investors more reluctant to invest in the underlying European stocks.

Last but not least, the two announcements which concern the French economy do not affect the prices at all. This is yet more evidence that the situation of the French economy does not influence investors' decisions of investing in European Futures.

<u>Table 7:</u> Results from individual regressions – first three minutes after the announcement

	30sec.-Returns			1min.-Returns			1.5min.-Returns			2min.-Returns			2.5min.-Returns			3min.-Returns		
	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2
DE CPI	0.00000	0.995	0.00	0.0000	0.851	0.00	0.000	0.704	0.00	0.0000	0.605	0.00	0.0000	0.984	0.00	0.0000	0.680	0.00
DE GDP	0.00018 *	0.036	0.12	0.00029 **	0.006	0.19	0.00039 **	0.001	0.25	0.00038 **	0.007	0.19	0.00052 ***	0.001	0.28	0.00051 **	0.001	0.25
DE Gfk ConsumerClimateIndex	-0.00004	0.940	0.00	0.00026	0.614	0.00	0.00003	0.961	0.00	-0.00009	0.890	0.00	-0.00048	0.534	0.01	-0.00059	0.443	0.01
DE Ifo BusinessClimateIndex	0.00012 ***	0.000	0.20	0.00011 *	0.013	0.10	0.00013 **	0.001	0.16	0.00010 *	0.022	0.09	0.00012 *	0.018	0.09	0.00010 *	0.039	0.07
DE IndustrialProduction	0.00001 *	0.042	0.07	0.00002 *	0.042	0.07	0.00003 **	0.004	0.14	0.00003 **	0.007	0.12	0.00003 **	0.005	0.13	0.00002 .	0.071	0.06
DE PPI	0.00001	0.582	0.01	-0.00003	0.294	0.02	-0.00002	0.622	0.00	0.00000	0.917	0.00	-0.00002	0.582	0.01	-0.00003	0.491	0.01
DE RetailSales	0.00002 *	0.049	0.07	0.00003 .	0.067	0.06	0.00003	0.132	0.04	0.00002	0.192	0.03	0.00003 .	0.078	0.06	0.00003	0.147	0.04
DE UnemploymentRate	-0.01202 ***	0.000	0.23	-0.01441 **	0.001	0.19	-0.01389 **	0.006	0.15	-0.01236 *	0.039	0.09	-0.01040 .	0.076	0.06	-0.00849	0.122	0.05
ECB InterestRate	-0.01305 ***	0.000	0.20	-0.01790 ***	0.000	0.26	-0.01210 **	0.004	0.14	-0.01007 *	0.020	0.09	-0.00502	0.269	0.02	-0.00605	0.165	0.03
EU GDP	0.00005	0.539	0.01	0.00019	0.107	0.05	0.00023	0.110	0.05	0.00024	0.119	0.04	0.00024 .	0.065	0.06	0.00027 .	0.076	0.06
EU IndicatorEconSentiment ZEW	0.00001	0.326	0.02	0.00001	0.458	0.01	0.00001	0.512	0.01	0.00001	0.538	0.01	0.00000	0.637	0.00	0.00000	0.610	0.00
EU IndustrialProduction	0.00002 .	0.091	0.05	0.00003	0.127	0.04	0.00003	0.189	0.03	0.00001	0.549	0.01	0.00001	0.758	0.00	0.00001	0.730	0.00
EU RetailSales	0.00002 *	0.017	0.10	0.00002	0.192	0.03	0.00001	0.683	0.00	0.00000	0.761	0.00	0.00002	0.302	0.02	0.00003	0.159	0.04
FR ConsumerSpending	0.00000	0.486	0.01	0.00001	0.327	0.02	0.00001	0.367	0.02	0.00001	0.362	0.02	0.00001	0.503	0.01	0.00001	0.299	0.02
FR CPI	0.00000	0.978	0.00	0.00000	0.933	0.00	-0.00004	0.506	0.01	-0.00004	0.541	0.01	-0.00002	0.733	0.00	-0.00004	0.550	0.01
FR GDP	-0.00001	0.915	0.00	0.00003	0.594	0.01	0.00005	0.569	0.01	0.00008	0.378	0.03	0.00010	0.355	0.04	0.00009	0.419	0.03
US CPI	0.00010	0.303	0.02	0.00009	0.385	0.01	0.00025 .	0.064	0.06	0.00013	0.304	0.02	0.00018	0.218	0.03	0.00012	0.444	0.01
US HomeSales	0.00219 *	0.015	0.10	0.00263 **	0.005	0.13	0.00271 *	0.021	0.09	0.00235 .	0.062	0.06	0.00224 .	0.088	0.05	0.00213	0.151	0.04
US IndustrialProduction	-0.00001	0.640	0.00	0.00002	0.315	0.02	0.00001	0.516	0.01	0.00002	0.337	0.02	0.00003	0.261	0.02	0.00004	0.135	0.04
US ISM ManufacturingPMI	0.01083 ***	0.000	0.22	0.01060 ***	0.001	0.21	0.01314 ***	0.000	0.29	0.01515 ***	0.000	0.31	0.01358 ***	0.000	0.27	0.01415 ***	0.000	0.29
US RetailSales	0.00010	0.106	0.05	0.00014 .	0.053	0.07	0.00016 .	0.051	0.07	0.00014	0.110	0.04	0.00014	0.146	0.04	0.000161	0.111	0.04

Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**

<u>Table 8</u>: Results from individual regressions - after the third minute

	4min.-Returns			5min.-Returns			6min.-Returns			7min.-Returns			8min.-Returns			9min.-Returns		
	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2
DE GDP	0.00057 **	0.01	0.19	0.00061 **	0.005	0.20	0.00063	0.115	0.08	0.00057	0.174	0.06	0.00060	0.155	0.07	0.00066	0.138	0.07
DE Ifo BusinessClimateIndex	0.00009 .	0.08	0.05	0.00011 *	0.042	0.07	0.00013 *	0.023	0.09	0.00012 *	0.043	0.07	0.00018 **	0.005	0.13	0.00018 **	0.008	0.12
DE IndustrialProduction	0.00002	0.12	0.04	0.00001	0.341	0.02	0.00002	0.173	0.03	0.00001	0.484	0.01	0.00000	0.771	0.00	0.00001	0.746	0.00
US ISM ManufacturingPMI	0.01703 ***	0.00	0.30	0.01550 ***	0.000	0.23	0.01626 ***	0.000	0.23	0.01660 ***	0.000	0.22	0.01496 **	0.002	0.18	0.01568 **	0.002	0.17

Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**

Tables 7 & 8 show the results of the individual regressions of the form $\boldsymbol{R_{tki}} = \boldsymbol{a_0} + \boldsymbol{a_k} * \boldsymbol{S_{ki}} + \boldsymbol{\varepsilon_i}$, *where S the announcement surprise defined as* $\boldsymbol{S_i} = \frac{A_i - F_i}{|F_i|}$ *for each announcement k and time frame t. Whenever the surprise is found to be significant for more than three minutes, the analysis continues until nine minutes after the announcement, and it is shown in table*

<u>Table 9:</u> Results from multivariate regressions (announcements that coincide) – first three minutes. after announcement

	30sec.-Returns			1min.-Returns			1.5min.-Returns			2min.-Returns			2.5min.-Returns			3min.-Returns		
	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2
EU CPI	-0.00004	0.153	0.08	-0.00005	0.168	0.06	-0.00005	0.257	0.03	-0.00004	0.369	0.02	-0.00004	0.365	0.03	-0.00005	0.308	0.02
EU UnemploymentRate	-0.01049 **	0.003	0.08	-0.01137 *	0.013	0.06	-0.00932 .	0.080	0.03	-0.00598	0.278	0.02	-0.01045 .	0.070	0.03	-0.00643	0.287	0.02
EU ManufacturingPMI	0.00232 *	0.036	0.03	0.00430 **	0.003	0.07	0.00565 **	0.004	0.05	0.00405 .	0.071	0.02	0.00523 *	0.025	0.04	0.00616 *	0.015	0.04
EU ServicesPMI	0.00000	0.685	0.03	0.00005	0.111	0.07	0.00002	0.675	0.05	0.00002	0.610	0.02	0.00005	0.302	0.04	0.00005	0.291	0.04
DE ManufacturingPMI	0.00043	0.694	0.02	0.00000	0.390	0.01	0.00000	0.849	0.01	0.00000	0.520	0.04	0.00000	0.690	0.05	0.00000	0.524	0.08
DE ServicesPMI	0.00129	0.241	0.02	0.00086	0.581	0.01	0.00150	0.333	0.01	0.00409 *	0.041	0.04	0.00461 *	0.021	0.05	0.00610 **	0.003	0.08
FR ManufacturingPMI	0.00000	0.430	0.01	0.00000	0.455	0.01	0.00000	0.719	0.01	0.00197	0.258	0.01	0.00122	0.554	0.02	0.00185	0.343	0.02
FR ServicePMI	0.00110	0.291	0.01	0.00000	0.790	0.01	0.00000	0.463	0.01	0.00000	0.565	0.01	0.00000	0.219	0.02	0.00000	0.305	0.02

Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**

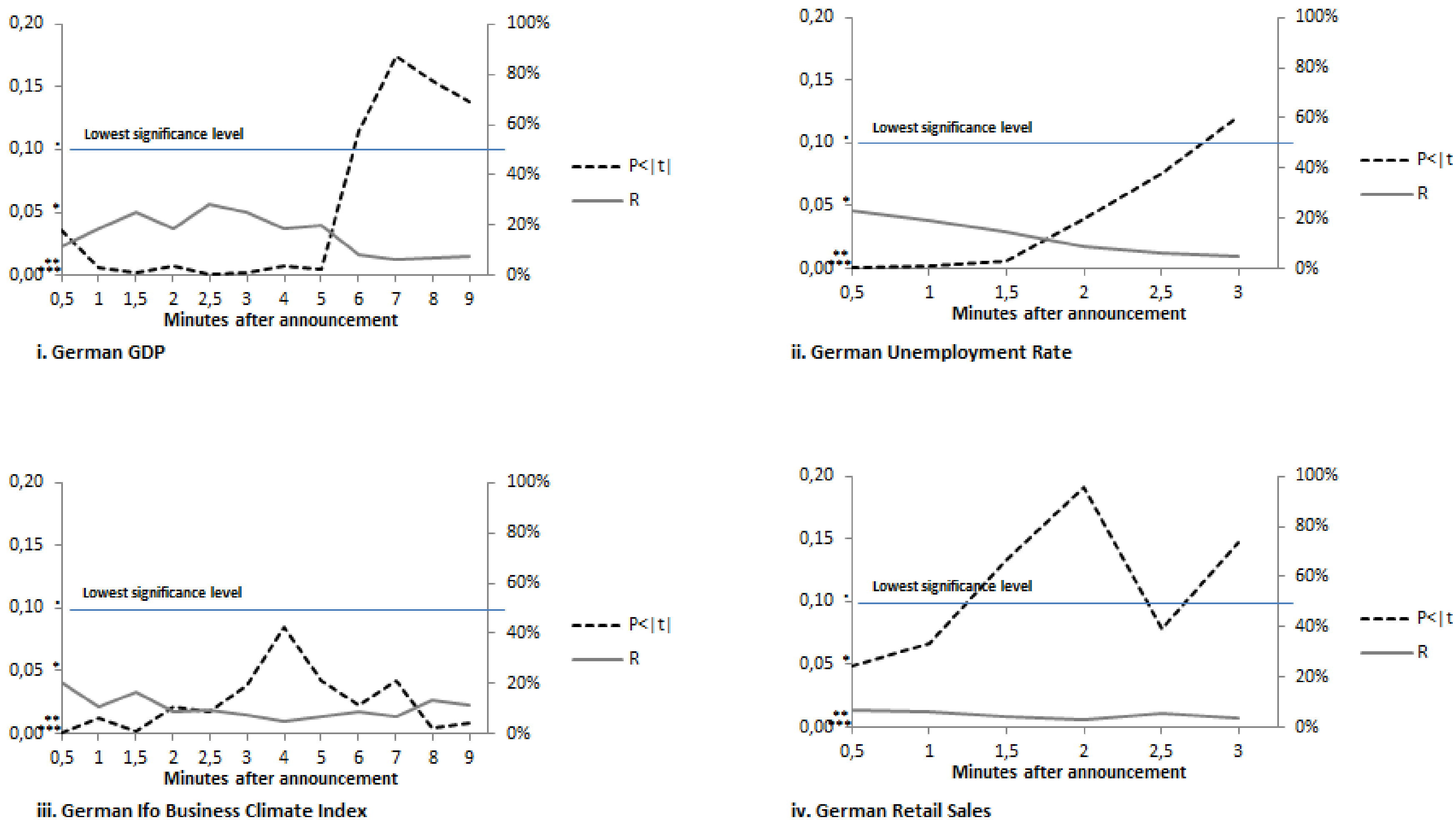
<u>Table 10</u>: Results from multivariate regressions (announcements that coincide) – after the third minute.

	4min.-Returns			5min.-Returns			6min.-Returns			7min.-Returns			8min.-Returns			9min.-Returns		
	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2	Coef.	P>|t|	R^2
EU ManufacturingPMI	0.00568 *	0.03	0.03	0.00754 **	0.008	0.04	0.00953 **	0.002	0.05	0.00358	0.248	0.01	0.00403	0.211	0.01	0.00694 *	0.038	0.03
EU ServicesPMI	0.00001	0.83	0.03	-0.00004	0.535	0.04	0.00001	0.933	0.05	-0.00004	0.586	0.01	-0.00002	0.815	0.01	-0.00003	0.676	0.03
DE ManufacturingPMI	0.00071	0.77	0.06	0.00000	0.673	0.05	0.00000	0.892	0.01	0.00000	0.817	0.01	0.00000	0.971	0.01	0.00197	0.645	0.01
DE ServicesPMI	0.00571 *	0.02	0.06	0.00682	0.022	0.05	0.00025	0.209	0.01	0.00023	0.261	0.01	0.00023	0.282	0.01	0.00019	0.420	0.01

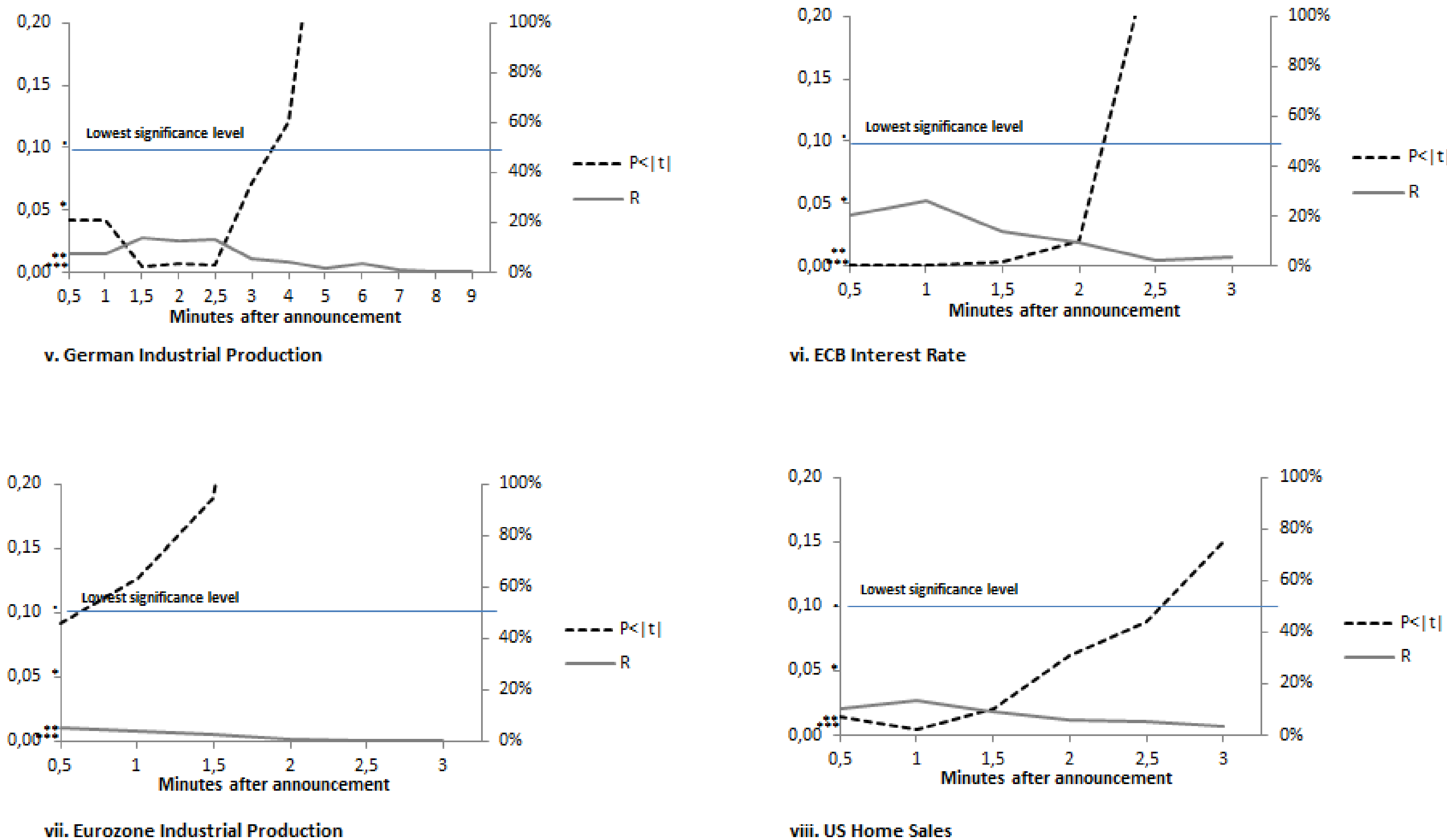
Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**

Tables 9 & 10 show the results of the multivariate regressions of the form $\boldsymbol{R_{tki} = a_0 + a_k S_{ki} + \sum_{k'=1}^{K'} a_{k'} * S_{k'i} + \varepsilon_i}$ *, where S the announcement surprise defined as* $\boldsymbol{S_i = \frac{A_i - F_i}{|F_i|}}$ *for each announcement k, its simultaneous announcements k' and time frame t. Whenever the surprise is found to be significant for more than three minutes, the analysis continues until nine minutes after the announcement, and it is shown in table 10.*

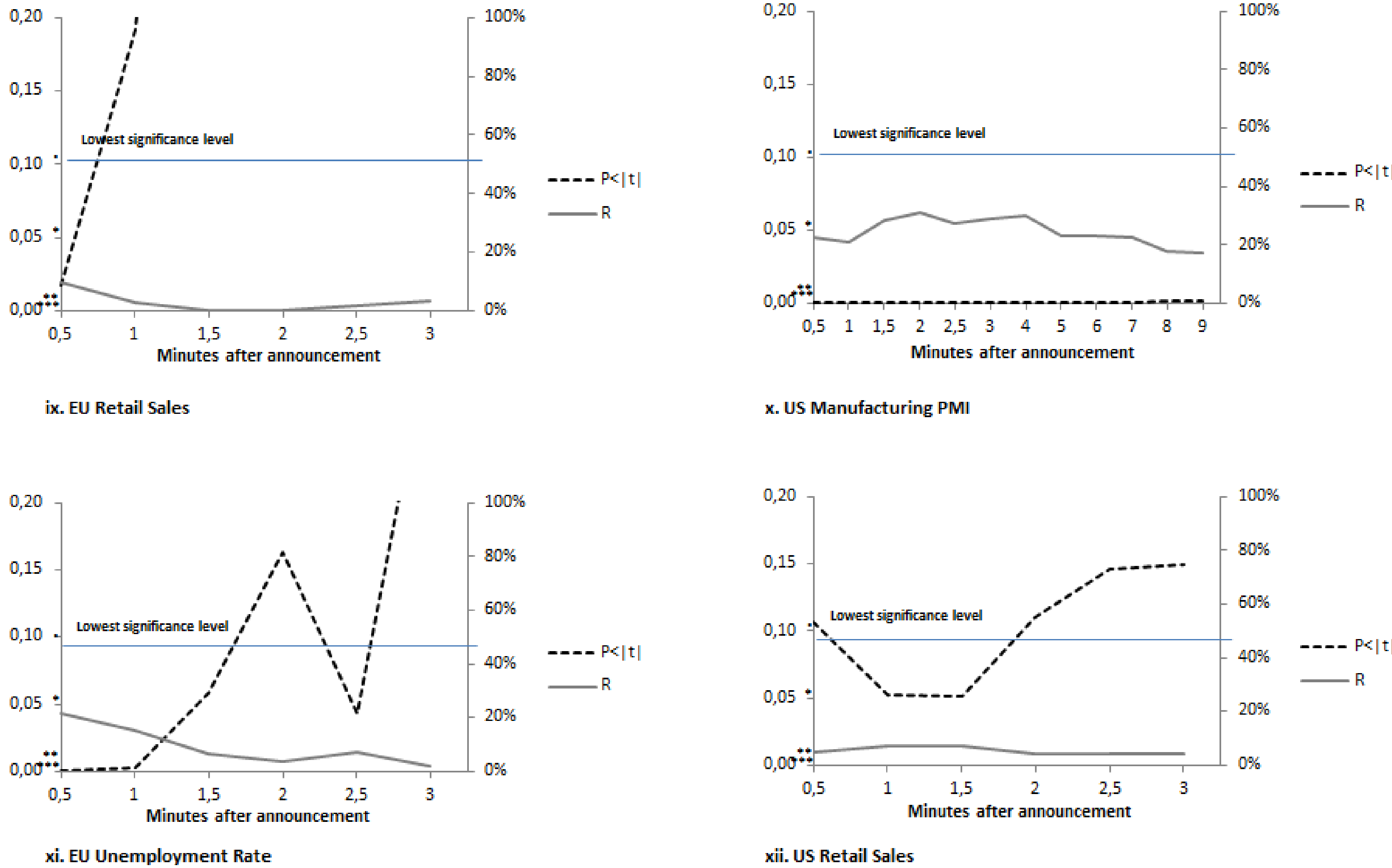
Figure 1: **Duration of the announcement effect**



<u>Figure 1:</u> Duration of the announcement effect (continues)



<u>Figure 1:</u> Duration of the announcement effect (continues)



Development of P-values (dotted line – left axis) and R^2 *(grey line – right axis) over time for each indicator. Asterisks on the left axis show significance on 0,001, 0.01, 0,05, 0,1 levels as used in Table 6. The announcement effect stops when the P-value exceeds the significance level 0,1. For those indicators for which the announcement effect is still preserved in the third minute, the analysis continues until the ninth minute.*

4.3 Effects after taking into account autocorrelation of returns

As described earlier, one of the goals of this paper is to prove that the observed effects of the various macroeconomic announcements on futures prices are not a result of the "normal path" of the prices themselves, but rather a result of the new information in market. For this reason, I developed the model (23), which includes four lags of the one-minute returns. The results after correcting for autocorrelation and heteroscedasticity of the residuals are shown in in Table 11.

The table shows the HAC coefficients for each announcement / lag of returns, their t- and p-values from testing the null hypothesis $\alpha_i = 0$ or $\beta_n = 0$ respectively (Student t-test for significance of the coefficients) and their significance codes, as described at the bottom. Similarly to the previous analysis, the lower the p-value, the more important the effect of the announcement.

As expected, not only the lags of the returns were found to be significant, but also several of the surprises. Fifteen out of twenty-nine announcements have a significant effect on prices, while twelve of them have already been identified as important from the models (19) and (20), in which the "natural" movements of the prices are not taken into consideration. This is the first demonstrates that any price changes during the first one minute after the announcement can indeed be accounted to the release of new information. These twelve indicators have been marked in light grey in the table.

In addition, three indicators, namely EU CPI, EU Indicator of Economic Sentiment (ZEW) and EU Services PMI, appear now to have a significant effect, while they previously did not. One reason could be that when they are analysed separately through individual regression, their effect on the price is overwhelmed by the autocorrelation of returns. Since the latter is not included in the original model, this could explain the initial low significance of these three announcements.

Moreover, there is no announcement that was found previously significant but not in the latest model. This shows once more that whenever a macroeconomic announcement appears to have an effect on the price, this is solely due to the new information that it brings. If this was not the case, then it would have been possible to identify some indicators as important using model (19) but as not important using model (20).

By emphasising the type of macroeconomic indicator, some remarkable conclusions can be drawn. Firstly, once more it is evident that the announcements related to the French-economy fundamentals do not affect the European stock-index future prices, even though the majority of the stocks comprising this index are French. This means that they are not essential to investors when it comes to assessing the state of the European economy. Secondly, as mentioned in section 3.1, there are two categories of macroeconomic indicator in the sample, the ones that are directly related to the prices and the ones that show the overall situation in the economy. From the first category, it seems that the investors pay attention only to the Eurozone CPI and the interest rate of the ECB. These two are essential for asset pricing, since the first one indicates the inflation level and hence the real prices while the latter is used for discounting the future values. Hence, their coefficients were expected to be statistically significant, since the underlying index comprises of stocks from different EU countries. On the contrary, German CPI and PPI do not play any role in formulating the prices of this stock-index future. The direction of the effect is also as expected. A positive surprise on CPI, for example, shows that inflation is larger than estimated, meaning that real prices are lower. This leads to selling of stocks and futures, causing further price reductions.

Concerning the second category, investors consider only some German and almost all Eurozone macroeconomic indicators important for assessing the general state of the EU economy. All of them appear to have the expected effect on prices (expected direction). Finally, three out of five US macroeconomic announcements are found to be significant. As pointed out by Rühl and Stein (2014) US indicators are important to investors even when they trade in European markets, because they demonstrate the state of the overall global economy.

<u>Table 11</u>: Results from unique regression, accounting for autocorrelation of returns (heteroscedasticity and autocorrelation corrected coefficients)

	Coef.		t value	P> | t
Intersept	0,00000		-0,42	0,67
Returns t-1	-0,06705	***	-21,89	0,00
Returns t-2	-0,01133	***	-4,33	0,00
Returns t-3	-0,00445		-1,17	0,24
Returns t-4	-0,00766	**	-3,00	0,00
DE CPI	-0,00001		-0,24	0,81
DE GDP	0,00030	***	6,48	0,00
DE Gfk ConsumerClimateIndex	0,00047		1,20	0,23
DE Ifo BusinessClimateIndex	0,00013	***	11,93	0,00
DE IndustrialProduction	0,00002	*	2,41	0,02
DE ManufacturingPMI	-0,00123		-0,44	0,66
DE PPI	-0,00003		-1,03	0,30
DE RetailSales	0,00002		1,43	0,15
DE ServicesPMI	0,00100		0,78	0,43
DE UnemploymentRate	-0,01495	**	-3,24	0,00
ECB InterestRate	-0,01797	**	-2,92	0,00
EU CPI	-0,00005	***	-10,40	0,00
EU GDP	0,00018	*	2,11	0,04
EU IndicatorEconSentiment ZEW	0,00001	*	2,19	0,03
EU IndustrialProduction	0,00002		1,57	0,12
EU ManufactoringPMI	0,00473	*	2,05	0,04
EU RetailSales	0,00002	.	1,77	0,08
EU ServicesPMI	0,00005	.	1,70	0,09
EU UnemploymentRate	-0,00899	*	-2,31	0,02
FR ConsumerSpending	0,00001		1,24	0,21
FR CPI	-0,00001		-0,44	0,66
FR GDP	0,00000		-0,08	0,93
FR ManufacturingPMI	-0,00105		-0,54	0,59
FR Service PMI	-0,00029		-0,22	0,82
US CPI	0,00008		0,73	0,47
US HomeSales	0,00310	**	3,17	0,00
US IndustrialProduction	0,00003		1,20	0,23
US ISM ManufacturingPMI	0,01170	***	4,11	0,00
US RetailSales	0,00015	***	3,46	0,00

Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**

Table 11 shows the results of the individual regressions of the form $\boldsymbol{R_t} = \boldsymbol{\beta_0} + \sum_{n=1}^{4} \boldsymbol{\beta_n} * \boldsymbol{R_{t-n}} + \sum_{k=1}^{29} \boldsymbol{\alpha_k} * \boldsymbol{S_{kt}} + \boldsymbol{\varepsilon_t}$ *where S the announcement surprise defined as* $\boldsymbol{S_i} = \frac{\boldsymbol{A_i - F_i}}{|\boldsymbol{F_i}|}$.

5. Application to high-frequency trading

Having conducted an in-depth analysis of the impact of macroeconomic announcements on stock future prices, I examine whether this information can be applied to intra-day trading. Since futures are a very liquid asset, any opportunity for speculation disappears within minutes or even seconds. In an era with such advanced information technology, it is rather simple to program an investment strategy that will open and close a position within a few seconds. Ironically, Ederlington and Lee (1993) wrote in the end of their analysis "In order to profitably trade on the market's initial reaction to the information release, one must be in a position to receive the information and place and execute an order within one minute. We doubt that anyone off the floor can do this effectively". Fortunately, after 22 years this is possible.

This last section shows how the method proposed in sub-section 3.2.1 can be used for predicting possible returns after each announcement. Unfortunately, due to the lack of data on future prices, the prediction procedure could not be fully tested using data from 2013-2014. However, a few cases are presented, which compare the predicted values to the realized ones, using a short dataset from January 2013–mid May 2013. Lastly, a trading strategy is proposed, along with some applied examples. Basic assumptions are that trading costs are negligible, the market is very liquid and any order can be executed immediately.

Having estimated the OLS coefficients in linear regression (19), the calculation of the predicted (or fitted) values is straight forward. Knowing the measured surprise, one can simply calculate the expected price returns. However, coefficient values are always related to the sample used and are the best possible estimate for the relation between the surprise and the return. In order to make more secure predictions, confidence intervals and/or prediction intervals are needed.

Provided a simple linear regression

$$\boldsymbol{y} = a_0 + a_1\boldsymbol{x} + \boldsymbol{\varepsilon} \tag{36}$$

where y, x and ε are vectors, the fitted value for y for a given x^* is given by:

$$\hat{y} = a_0 + a_1 x^* \; . \tag{37}$$

Then the confidence interval for $\hat{y}$ is defined as:

$$c = \hat{y} \pm t_{(n-2,\ 0{,}05)} * s\sqrt{\frac{1}{n} + \frac{(x^*-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}} \tag{38}$$

where

$t_{(n-2,\ 0.05)}$ is the value of t-distribution for n-2 degrees of freedom and 0,005 significance level,

s is the standard deviation of the residuals and

n is the number of observations in the sample.

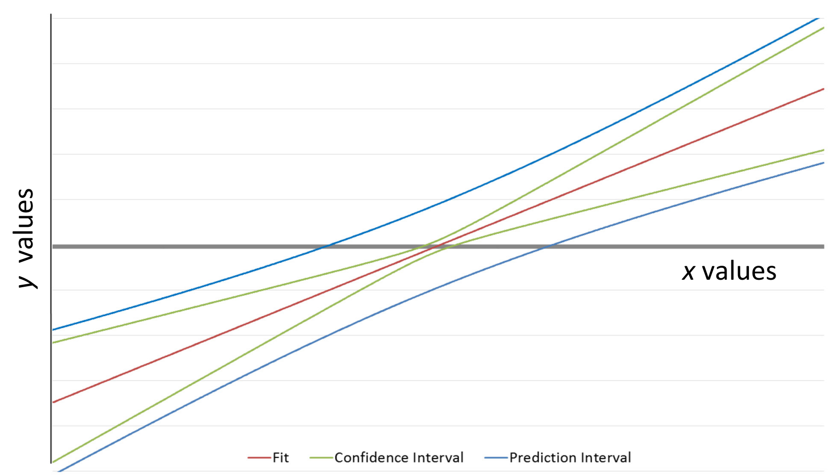
Similarly, the prediction interval for $\hat{y}$ is given by:

$$p = \hat{y} \pm t_{(n-2,\ 0.05)} * s\sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum(x_i - \bar{x})^2}} \qquad (39)$$

On the one hand, the confidence interval covers the average value of y, given a certain x, i.e. gives an upper and lower value for the estimate of $E(y|x)$. On the other hand, the prediction interval covers the value of the estimated $\hat{y}$ given certain x. In other words, the confidence interval covers the fluctuations of the estimated mean value of the dependent variable, while the prediction interval covers the fluctuations of the fitted value itself. Hence, the prediction interval is always larger, since it needs to cover a much more volatile variable (see Draper and Smith (1981) or Wooldridge (2009)).

Figure 2 represents the prediction and the confidence interval for a random simple regression such as (35). The red line represents all fitted values of y for a given x, while the green and the blue represent the confidence and prediction interval respectively. As it can be seen, the confidence interval is smaller around the mean value of x and it becomes wider as the new $x*$ value moves away from it. By definition, the larger the deviation of $(x^* - \bar{x})^2$, the larger the interval.

Figure 2: Graphical representation of Confidence and Prediction Interval



By using the values mentioned above, one can estimate an upper and a lower boundary for the realised returns at the time when an announcement occurs. However, this depends on the chosen time interval, since each announcement affects the prices differently at different times after the release, as shown in section 4. One therefore needs to choose the time frame within which the price change is mostly explained by the surprise, i.e. where the R^2 is larger. A R^2=0,20 or larger is typically considered sufficient for making predictions. Hence, from all announcements that are identified as important using the model (19), only those regressions that give a R^2 equal or larger than 0,20 are selected. This also means that for each announcement the time frame is used in which the effect is larger. As proven in section 4, the value of R^2 is positively related to the significance of the coefficient. However, note that for German Ifo Bus. Climate Index and German Unemployment rate, the largest R^2 occurs when regressing with thirty-second returns. Since no trading can be done in such a short time frame, the regression with the next lower-level R^2 is selected. As a result, the following announcement and time frames are used for testing the predictive power of the model and developing a trading strategy (Table 12):

Table 12: Regressions used for building the proposed strategy

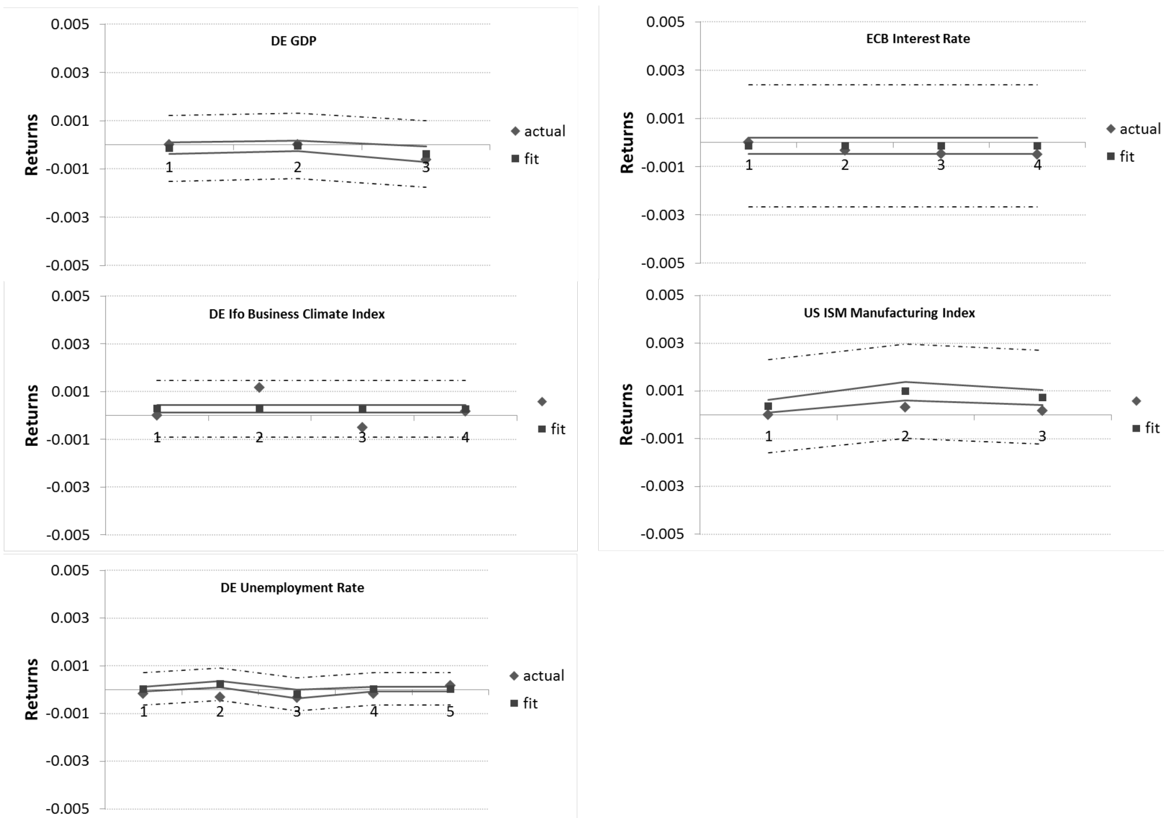
Announcement	Minutes after release	Regression used
German GDP	2,5	$\hat{R} = -0{,}00004 + 0{,}00052 * S'$
German Ifo Bus. Climate Index	1,5	$\hat{R} = 0{,}00028 + 0{,}00013 * S'$
German unemployment rate	1	$\hat{R} = 0{,}000022 - 0{,}01441 * S'$
ECB Interest rate	1	$\hat{R} = -0{,}00014 - 0{,}0179 * S'$
US ISM Manufacturing	2	$\hat{R} = 0{,}00024 + 0{,}01515 * S'$

A good way to test the predictive power of the model is by forward testing. By using the realised surprises (S') in the years after 2012, it is possible to compare the realized returns with the fitted ones ($\widehat{R}$). In Figure 3 these comparisons are shown together with the confidence and prediction intervals. Due to a lack of data for the index prices for the years 2013-2014, only the results from the first five months of 2013 are shown.

The x-axis depicts the observations since January 2013, while the y-axis shows the level of returns. The rhombus represents the actual realised returns for each observation and the square the fitted value. Lastly, the continuous lines illustrate the confidence interval and the dotted one the prediction interval for each observation. For German GDP there are three observations: preliminary end-of-year, revision of end-of-year and first-quarter value. For the ECB Interest Rate and German Unemployment rate, data for May was not made available, while for the US ISM Manufacturing index the observations for April and May were excluded, since the release was made on bank holidays in Europe.

As it can be seen, the predicted values do not deviate much from the fitted ones, with the largest deviation being 0,09% (second observation point of DE Ifo Business Climate Index). Additionally, although the confidence interval does not always manage to cover every realized value, the prediction interval does. Generally, the model must not be used for making exact predictions of future values, but for giving a margin of the possible higher and lower future returns in a specific time interval after the announcement.

<u>Figure 3:</u> Comparisons of fitted with realised returns and their confidence and prediction intervals



What is therefore the new information from this model that one can use in one's trading strategy? Firstly, it determines which macroeconomic announcements are significant and hence require attention. As seen in the previous section, there are 11 such macroeconomic indicators. This indicates that applying a trading strategy which takes advantage of the high volatility right after the release might prove profitable. Secondly, for any surprise level an upper and lower boundary for the expected returns can be estimated by using confidence and prediction intervals. This means that at the exact time of the announcement one knows if they should expect positive or negative returns, as well as their maximum and minimum values. One can therefore decide if they will enter the market or not, if they will take a long or short position, and even when is the right time to leave the market. Surely any actions must be taken within one or two minutes, but this is possible by applying computerized trading. One needs only to define and program a strategy.

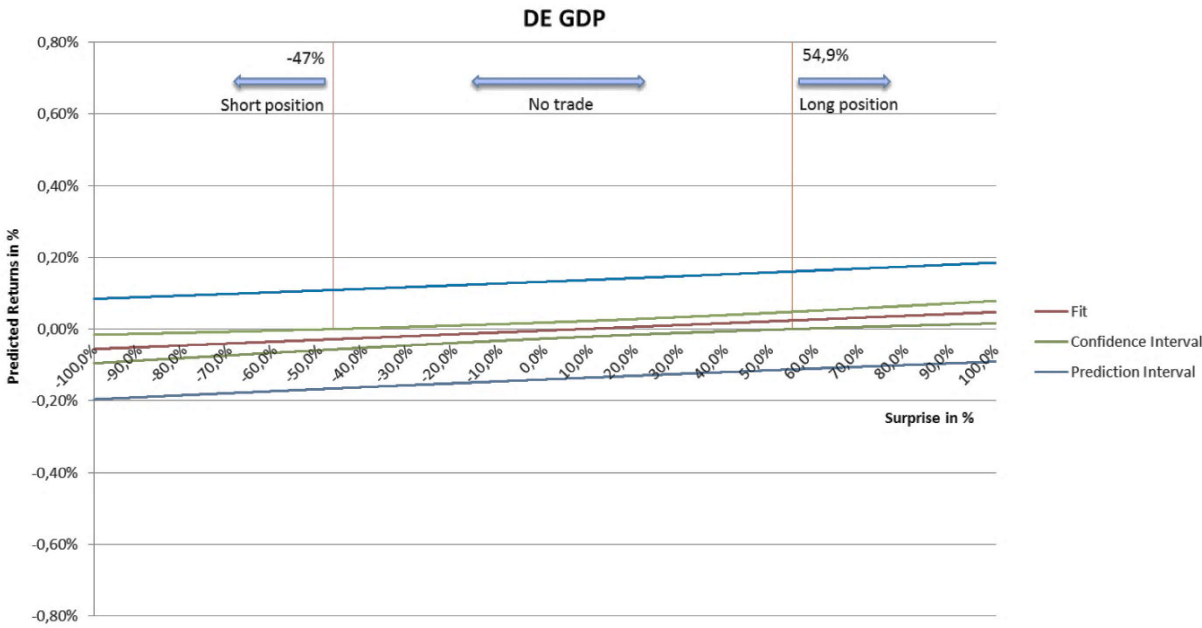
In order to structure the strategy, entry and exit conditions are defined, as well as the position to be taken, by making use of the confidence and prediction intervals. Firstly, for each of the five final selected announcements and for each level of surprise, in a range between -100% and 100%, the fitted values of the returns and these intervals are calculated. Since showing all these calculated values within a table would be rather counterproductive, they are presented graphically in Figure 4.

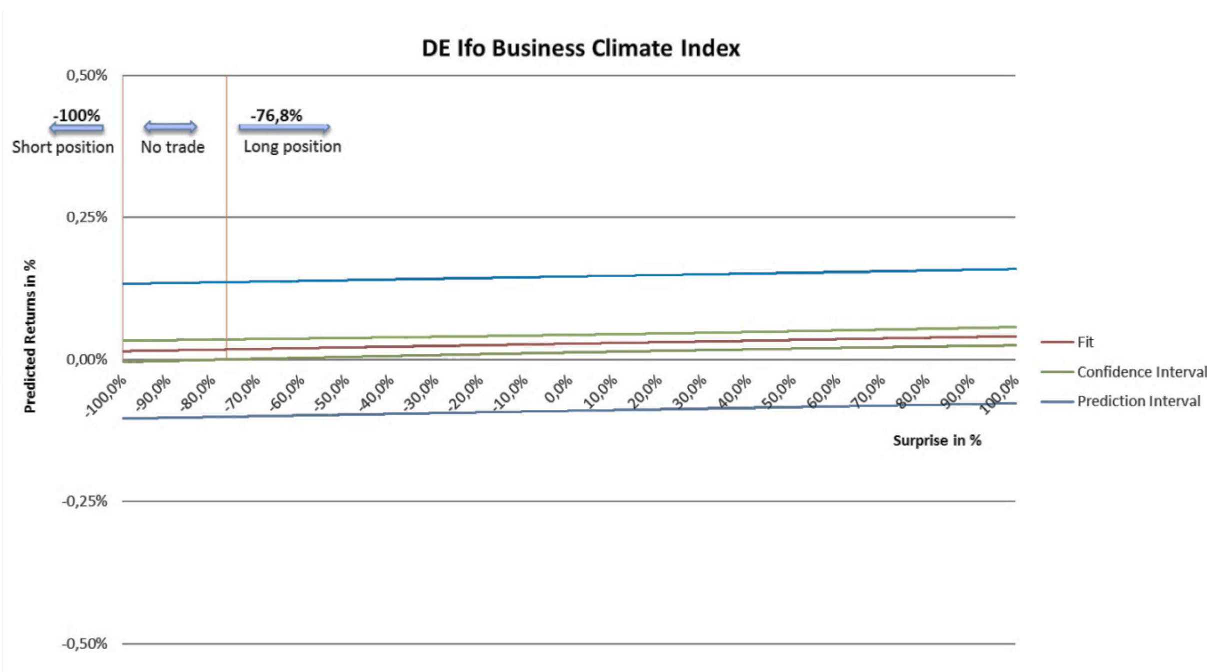
Subsequently, the announcements are divided into two groups: those for which the surprises have a positive relation to the returns (DE GDP, DE Ifo Index and US ISM Index) and those where the relation is negative (ECB interest rate and DE unemployment rate). For each group a different entry condition is defined. For the first group, the "entry level" surprise is defined as the surprise level at which the lower bound of the confidence interval is equal/larger than zero or the upper bound of it is equal/lower than zero. In the first case, the position should be long in order to make use of the expected positive returns. In other words, the investor should go long only if the average predicted return is non-negative. In the second case, the position should be short in order to deal with the expected negative returns. Similarly the intuition is that the investor should invest short when the average predicted return is negative. For any case in between, the confidence interval leans between negative and positive values, making it impossible to decide for a long or short position. Hence, if the realised surprise leans between the two entry level surprises, then no trading should be done.

On the other hand, when the relation between the surprise and the return is negative, exactly the opposite strategy should be followed. A long position should be taken if the realised surprise is smaller than the surprise where the upper bound of the confidence interval is equal or larger than zero, and short position should be held if the realised surprise is smaller than the surprise where the lower bound of the confidence interval is equal or smaller than zero. In between, no trading should take place.

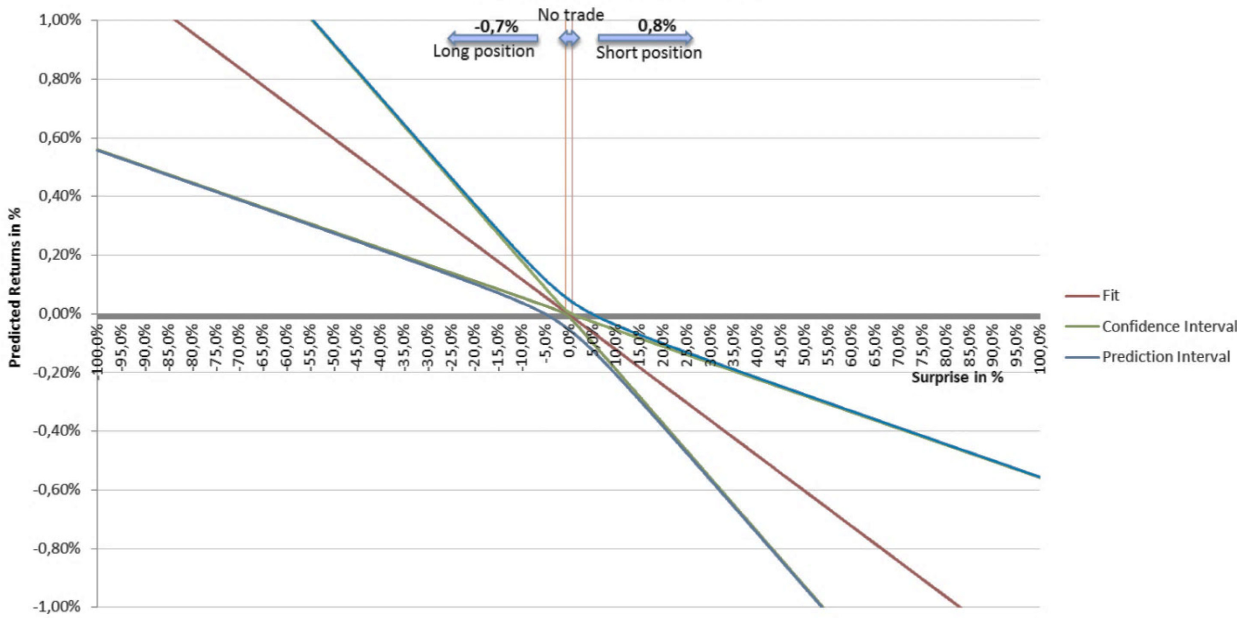
In order to avoid confusion, the proposed strategy is summarised in Table 13 and the trading criteria for each announcement type are shown in the respective graph of Figure 4.

Figure 4: **Trading criteria based on the announcement type**

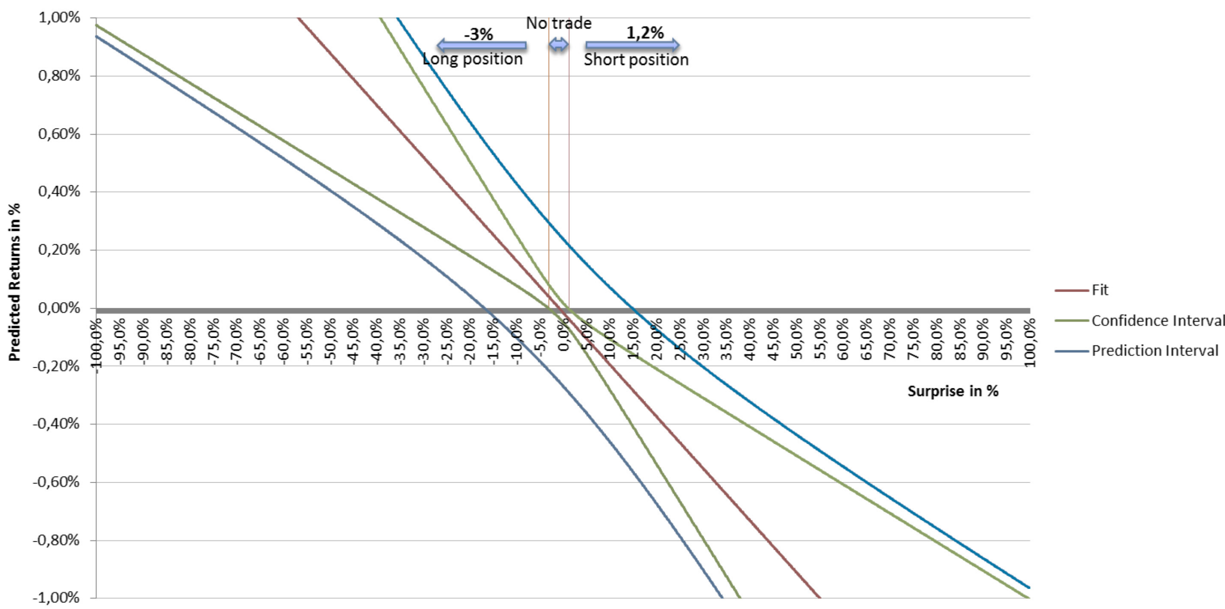


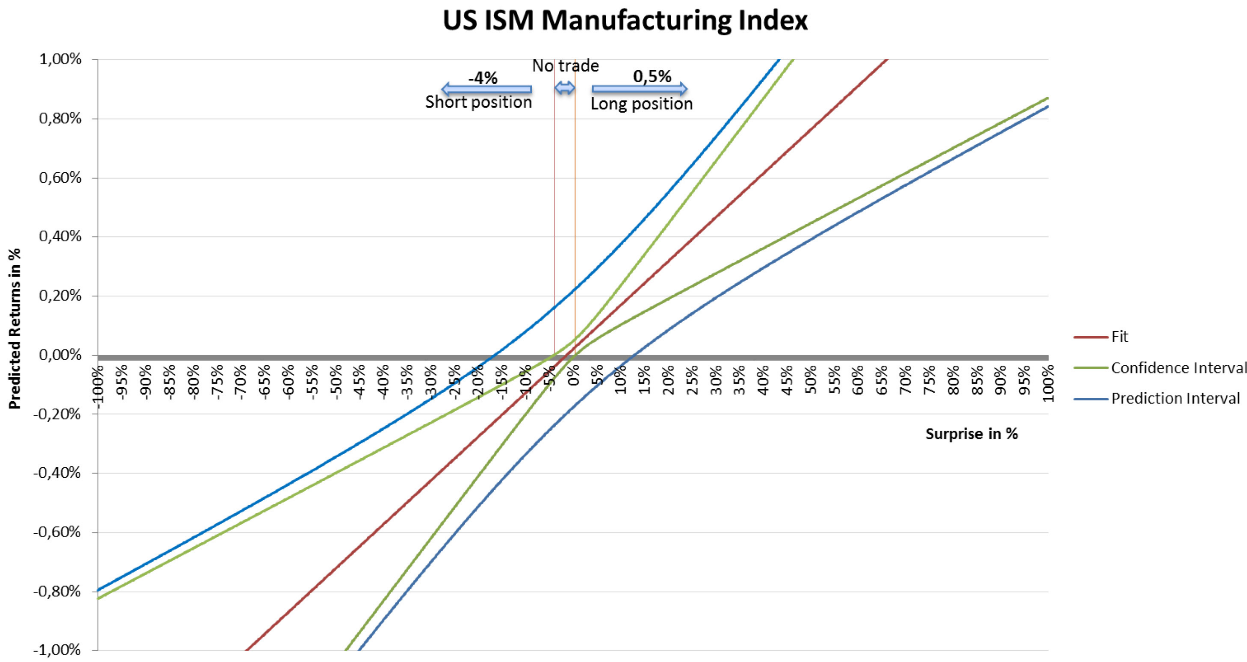


DE Unempoyment Rate



ECB Interest Rate





Subsequently, the exit condition should be defined, i.e. a time or return level at which the investor closes the position. An exit time is easier to determine. One should not hold the position longer than the time frame used for estimating the initial regression (see Table 12) and hence the confidence interval. For example, when a surprise lower than -3% comes from the announcement of the ECB Interest rate, the investor will go long for one minute, since for this level and time frame the model predicts on average positive returns. After this time the position should be closed, since the model cannot guarantee any subsequent development.

On the other hand, the bounds of the prediction interval can also serve as an exit condition (exit-level return). This means that if the position is long (short) and the realized returns reach the upper (lower) bound of the prediction interval the investor should close the position. Since this is the maximum return that the model predicts for the specific time frame, keeping the position longer cannot guarantee larger returns. Finally, these two conditions can be combined as follows: the investor should close the position when the respective time frame after the announcement is reached or when the exit-level return is realised, whichever comes first. The exit conditions are summarized in Table 13 as well.

<u>Table 13</u>: Summary of the proposed investment strategy

Relation between surprise and returns	Critical surprise level (for entering)	Position	Exit condition
Positive	1st: Lower bound of confidence interval equals zero	Long if realized surprise higher than 1st critical	Hold position until the respective time frame used for regression or until the realized return equals the <u>upper bound of prediction interval</u> (maximum predicted returns), whichever comes first.
	2nd: Upper bound of confidence interval equals zero	Short if realized surprise lower than 2st critical	Hold position until the respective time frame used for regression or until the realized return equals the <u>lower bound of prediction interval</u> (maximum predicted returns), whichever comes first.
	Within confidence interval	No trade	-
Negative	1st: Lower bound of confidence interval equals zero	Long if realized surprise lower than 1st critical	Hold position until the respective time frame used for regression or until the realized return equals <u>upper bound of prediction interval</u> (maximum predicted returns), whichever comes first.
	2nd: Upper bound of confidence interval equals zero	Short if realized surprise higher than 2st critical	Hold position until the respective time frame used for regression or until the realized return equals the <u>lower bound of prediction interval</u> (maximum predicted returns), whichever comes first.
	Within confidence interval	No trade	-

6. Conclusion

This paper shows that announcements regarding European macroeconomic fundamentals do affect the European future stock market and their effect is rather quick. Two different approaches are used based on whether the announcements are simultaneously released or not. In both cases the effect is a decreasing function of time and in most cases it lasts for 1,5–2,5 minutes. The list of the most important announcements consists of announcements about Eurozone, German and a few US macroeconomic fundamentals. Surprisingly, not one of the French announcements is found to affect the prices.

Moreover, it shows that even when the autocorrelation of returns is taken into consideration, these announcements are still significant. This means that the price change at the time when the announcement occurs is a result of the new information reaching the market and not only of the natural patterns the returns show.

Lastly, it shows that by estimating the direction and the magnitude of the effect for each important announcement, it is possible to make predictions about the price change after each future release. This information, in combination with the estimation of confidence and prediction intervals, can be used for setting up a high-frequency trading strategy, mostly used for speculation. What is left for future research is the testing the proposed strategy (using forward testing) in order to assess its profitability. This testing might be combined with actual bid-ask prices, trading volumes and trading cost, in order to evaluate if it is profitable in every-day trading. Lastly, it can be extended to other markets, such as the bond market, which are less liquid and the effect might last longer, so that there is enough time for opening and closing a position.

7. References

Ahn, D. H., Boudoukh, J., Richardson, M., Whitelaw, R. F., 1999 July. Behavioralize this! International evidence on autocorrelation patterns of stock index and features returns. National Bureau of Economic Research, Working paper 7214.

Ap Gwilym, O., Buckle, M., Clare, A., & Thomas, S., 1998 Winter. The transaction-by-transaction adjustment of interest rate and equity index futures markets to macroeconomic announcements. Journal of Derivatives 6, 7–17.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. 19, 716-723.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Vega, C., 2003. Micro effects of macro announcements: Real price discovery in foreign exchange. American Economic Review 93 (1), 38-62.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Vega, C., 2007. Real-time price discovery in global stock, bond and foreign exchange markets. Journal of International Economics 73, 251–277.

Andrews D. W. K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817–858.

Balduzzi, P., Elton, E. J., Green, T. C., 2001. Economic news and bond prices: Evidence from the U.S. treasury market. Journal of Financial and Quantitative Analysis 36, 523-543.

Bollerslev, T., Chou, R. Y., Kroner, K. F. 1992. ARCH modeling in finance: A selective review of the theory and empirical evidence. Journal of Econometrics 52, 5-59.

Breusch, T. S., 1978. Testing for autocorrelation in dynamic linear models. Australian Economic Papers 17, 334-355.

Breusch, T. S., Pagan, A. R., 1979. A simple test for heteroscedasticity and random coefficient variation. Econometrica 47, 1287-1294.

Cook, T., Korn, S., 1991. The reaction of interest rates to the employment report: The role of policy anticipations. Economic Review, Federal Reserve Bank of Richmond 77, 3-12.

Diebold, F. X., Lopez, J., 1995. Modeling Volatility Dynamics. Macroeconometrics: developments, tensions and prospects, Boston: Kluwer Academic Press, 427-472.

Draper, N. R., Smith, H., 1966. Applied regression analysis. Wiley series in probability and mathematical statistics, John Wiley & Sons Inc., Chapter 1.

Durbin, J., Watson, G. S., 1971. Testing for serial correlation in least squares regression III. Biometrika 58, 1-19.

Dwyer, G., Hafer, R. W., 1989. Interest rates and economic announcements. Review, Federal Reserve Bank of St. Louis 71, 34-46.

Ederlington, L. H., Lee, J. H., 1993. How markets process information: News releases and volatility. Journal of Finance 48, 1161–1191.

Gallant, A. R., 1981. On the bias in flexible functional forms and an essentially unbiased from: The Fourier flexible form. Journal of Econometrics 15, 211-245.

Godfrey, L. G., 1978. Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables, Econometrica 46, 1293-1302.

Gonell, T. F., 1995. The distribution of reversals and continuations and tests for intraday market efficiency. Journal of Business Finance and Accounting 22, 255-243.

Greene, W. H., 1997. Econometric Analysis, Prentice-Hall.

Jones, B., Lin, C. T., Masih A. M., 2005. Macroeconomic announcements, volatility, and interrelationships: An examination of the UK interest rate and equity markets. International Review of Financial Analysis 14, 356-375.

Johnston, J., DiNardo, J., 1997. Econometric Methods, McGraw Hill International.

Kim, S. J., McKenzie, M. D., Faff, R. W., 2004. Macroeconomic news announcements and the role of expectations: evidence for US bond, stock and foreign exchange markets. Journal of Multinational Financial Management 14, 217-232.

Newey W. K., West K. D., 1987. A simple, positive - definite, heteroscedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Rühl, T. R., Stein, M., 2014. Discovering and disentangling effects of US macro-announcements in European stock markets, Ruhr Economic Papers 500, available at SSRN: http://ssrn.com/abstract=2462515.

Wooldridge, J. W., 2009. Introductory econometrics, South-Western Cengage Learning, 4th Edition, Chapters 2, 3, 6, 8.

White H., 1980. A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity. Econometrica 48, 817–838.

White H., Domowitz I., 1984. Nonlinear regression with dependent observations. Econometrica 52, 143–161.

Zeileis, A., Hothorn T., 2002. Diagnostic checking in regression relationships. R News 2,7-10.

Zeileis, A., 2004. Econometric computing with HC and HAC covariance matrix estimators. Journal of Statistical Software11, 1-17.

Curriculum Vitae

Name: Katerina Koule

E-mail: katerina.koule@rbinternational.com , catherine.koule@gmail.com

EDUCATION

09/2012-09/2015	**University of Vienna** MSc in Quantitative Economics, Management and Finance
09/2007-11/2011	**National & Kapodistrian University of Athens** Undergraduate degree in Economic Science (Ptychion), G.A.S.: "Very Good" 7,08/10 *Major*: Business Economics and Finance *Minor*: Quantitative Methods
02/2011-06/2011	**Karl Franzes University of Graz** Erasmus Exchange Program
09/2004-06/2007	**28th General Senior High School of Athens** Certificate of Graduation, G.A.S.: "Excellent" 18/20

WORKING EXPERIENCE

02/2015-today	**Raiffeisen Bank International–Global Exposure & Limit Management** Junior Risk Data & Reporting Manager
09/2013-01/2015	**Raiffeisen Bank International–Credit Risk Control** Freelancer in Credit Portfolio Analysis team
07/2010-11/2010	**Invest in Greece Agency–Ministry of Finance** Intern-Assistant to the President and the CEO.
06/2008-12/2008	**National Bank of Greece** Bank Clerk in Mortgages - Basic Banking Products

PROJECTS

02/2015-07/2015	**Raiffeisen Bank International–Global Exposure & Limit Management** GEM Project - Implementation of IBM Algo Credit Manager software
01/2014-04/2014	**Raiffeisen Bank International–Credit Risk Control** Asset Quality Review - Assessment of bank's assets quality by ECB