Predictive Power of the ZEW Sentiment Indicator: Case of the German Automotive Industry

Lubor Homolka, Drahomíra Pavelková

Tomas Bata University in Zlín, Faculty of Management and Economics, Mostní 5139, 760 01 Zlín, Czech Republic

homolka@fame.utb.cz, pavelkova@fame.utb.cz

Abstract: This paper presents an analysis of German automotive industry and its connection to the market sentiment indicator ZEW. The analysis spans a period of the last decade and is divided into Pre-Crisis, Crisis and Post-Crisis periods. Research questions related to the predictive power of ZEW indicator on macro level indicators (composite DAX and GDP), sector indicator (technology-oriented companies TecDAX) and a selected automotive manufacturer (BMW) were answered. We found that ZEW index had foreseen the economic crisis starting in the March 2008 three months ahead of its start, but failed to see an upcoming economic recovery. We fit two models to estimate whether ZEW index can be used as a standalone forecasting instrument or whether inclusion of lagged values of other variables improves forecasting ability. We conclude that predictions from the ZEW-only models are worse in the test sample than those of the more complex model. We provide further evidence in form cross-correlations and causality analysis in the Granger sense. The study concludes with Impulse Response Function analysis. This analysis found that reaction of TecDAX on change of ZEW is strongest amongst studied variables.

Keywords: market sentiment; automotive industry; ZEW; DAX; VAR; Granger causality

1 Introduction

To efficiently utilise all available information is amongst the most challenging tasks in any analytical process. Those who are uncovering the fundamental value of a company must be concerned about the complete environment in which a company operates. Although it's clear that companies operate within a sector and a general macroeconomic framework, there is no clear link between macroeconomic variables' effects on a single company's performance. Many factors, mostly idiosyncratic, complicate the problem. How fast change in CPI translates into a change of interest rates of some financial instrument that company uses? Idiosyncratic effects represent a significant problem as they hide or completely neglect theoretical explanations and expectations. Creating a model based on a pool of similar companies, say, within an industry, to predict a single-company performance might work only in a limited way. This does not mean that single-company analysis is impossible. The analysis must be tailored to the specific case. There are approaches which mitigate this problem. For example, in the portfolio theory, idiosyncratic risks are removed through the process of diversification.

In this paper, an attempt of connecting a macro-level indicator to a single-company performance is made. Macroeconomic conditions are frequently described in terms of GDP growth, unemployment rate, inflation and other indicators. Peiró [1] highlights differences between macroeconomic research conducted in the US and in Europe. He concludes that changes in production and interest rates are the determinants of stock returns on the French, German and United Kingdom stock markets. These indicators explain about half of stock returns.

As a proxy the general economic condition can be also used as a latent indicator – market sentiment. Sentiment does not only reflect a current state of the economy but is also forward-looking.

For this paper, a macro-level indicator is considered an indicator of general economic conditions. We note that this is not exactly a fundamental analysis; as a proxy for company's performance are used stock price returns. In a real world, returns subject to many factors other than changes in fundamentals, such as speculative or inside trading. Moreover, given the rather low free float of European companies, the price of share might differ significantly in the case of full free float companies.

Finter et al. [2] divide sentiment measures to *implicit* and *explicit* sentiment proxies. *Implicit* proxies rely on market data. Baker and Wurgler [3] have developed composite indicator relying on six market variables: i) the total number of initial public offerings (IPO), ii) the average first-day returns of IPOs, iii) the dividend premium, iv) the close and fund discount, v) New York Stock Exchange turnover, and vi) the equity share in new issues. They adjusted this indicator by macroeconomic indicators to make the index comparable throughout the business cycle by regressing original sentiment values to the growth of industrial production, the real growth of goods and services growth. Shen at al. [4] expanded this model by adding total productivity, current inflation growth and market expectations, term and default premiums, aggregate market volatility and market excess returns and labour income growth. Lux [5] points to other less-frequent market indicators, such as the ratio of equity put to call trading volume or number of advancing to declining issues. A novel approach is provided in [6]. Sentiment indicator is decomposed to the rational (regression on economic indicators) and irrational parts. The irrational part contains residuals from the first regressions. Both parts have a predictive power on the near-term returns.

Survey-based indicators belong to the category of explicit proxie. Consumers and investors are being asked about their expectations related to either general economic conditions or stock index performance (market consensus released by rating agencies or other financial companies). ZEW Germany Expectation of Economic Growth Index (ZEW) belongs to the first category. ZEW index was chosen from available explicit proxies.

This paper aims to evaluate the forecasting power of ZEW. Last similar analysis was conducted by Spiwoks [7]. Our analysis is made on updated data sample ranging from 2005 to 2016. Time series are further divided according to the business cycle phases into three parts. We have designed two models which should provide a better estimation of the ZEW's predictive power. Predictive power is also assessed by Granger causality tests. First, is the simple model with only lagged explanatory variable of itself. The second model is the Vector Autoregressive model. We propose a dynamic prediction (rolling MSE) as the model's quality indicator.

2 Theoretical Framework

Companies operate in a larger framework described by macroeconomic or financial theory (e.g., Keynesian economics, Efficient market theory), represented by macroeconomic variables (e.g. inflation) and measured by macroeconomic indicators (e.g. Consumer Price Index). An information content of such indicators faces ongoing criticism for low validity and reliability. Researchers and practitioners are creating methodologies and tools to describe and summarise multidimensional nature of the reality [8]. From the practical side, some indicators, such as GDP growth, are usually released quarterly, while CPI is released on monthly basis.

Research directions of market sentiment can be categorised into two groups. The first concerns testing of theory. As Lux [5] states, under the Efficient Market Theory, publicly released sentiment measures should be reflected in the prices and should not, therefore, have any predictive power on prices. However, Lux also provides a literature review on "noise trader risks" which explains why some traders (non-fundamental traders) can make a profit in a long-term by following such indicators. The second group utilises sentiment as a predictor for other time-series or entity behaviour. This group relates to valuation and portfolio modelling. According to Baker and Wurgler [3, p. 1652] “market sentiment may cause systematic patterns of mispricing”.

Expected performance of any economy closely relates to the phase of the business cycle. Sentiment of market participants can provide another explanation why some companies and sectors exhibit different performance throughout the business cycle. Shen et al. [4] concluded that high-risk portfolios tend to earn more in the low sentiment periods than low-risk portfolios. Reverse works in the high sentiment periods. Yu and Yuan [9] analyse behaviour of sentiment traders and test the mean-variance relation in different periods. They conclude that in the low-sentiment period expected market returns positively relate to the conditional variance. This does not necessarily hold true in the high sentiment periods when the relation is weak. Stambaugh et al. [10] analyses market-wide sentiment and its influence on mispricing of financial assets.

AsFinter, Niessen-Ruenz&Ruenzi [2] point, market movers can be retail or institutional investors. The first is more common in the US, whereas the latter in the UK or Germany. Transferability of research results can be problematic as the investors' responses to shocks differ. Behaviour of smaller market participants was analysed Chackley et al. [11]. Main source of the data was content of micro-blogging sites. They analysed relations between sentiment and stock price, volatility and traded volume on a short trading frame. Extensive literature review on application of text mining tools in market prediction was conducted by Nassirtoussi et al. [12]. They provide a deep discussion about approaches to non-traditional sources of data, such as social networks to assess an instantaneous mood of the market.

3 Methodology of Research

German automotive industry for the empirical study was selected. Germany was selected for (i) its developed financial market, (ii) availability of sentiment indicators and (iii) a high importance of automotive industry, which is a central topic of the project this paper contributes to.

3.1 Data Selection

This study deals with macro, sector and individual company performance and its relation to the sentiment indicator ZEW. Data was acquired from the period of 2005 to 2016. This period is divided into three periods: Pre-Crisis, Crisis (March 2008 - June 2009) and Post-Crisis.

GDP and composite DAX index were selected to track the general economy. The technological sector index TecDAX was selected from available indexes tracked by Deutsche Börse. Its performance is based on the performance of the 30 largest companies from the technological industry listed on the Frankfurt Stock Exchange. On the company level, car manufacturer Bayerische Motoren Werke AG (BMW) was chosen. Selection of companies is explained in a dedicated paragraph.

Germany went through the whole economic cycle during analysed period. Grey areas in Figure 1 indicate short Crisis period. Left window of Figure 1 tracks development of BMW and Continental AG. Performance of both titles surpasses growth of the economy. The quarter growth of GDP [13], composite index DAX and TecDAXexhibit similar development. Time series were scaled to have index value 1 in 2005 as they differ in levels.

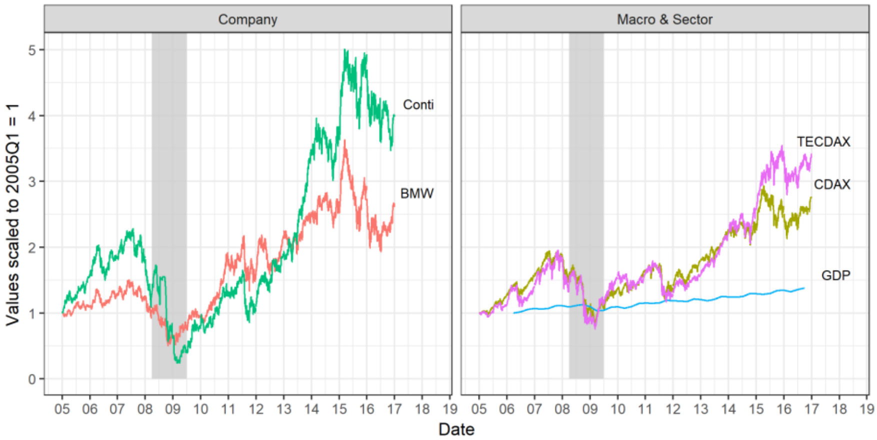

Figure 1

Development of performance of BMW, Continental AG, GDP, composite DAX and TecDAX.

Source: Own processing

In the initial research design, three car manufacturers were considered - BMW, Daimler (DAI) and Volkswagen (VOW). Looking at the development of the share prices we see a similar pattern in both BMW and Daimler. Three panels in Figure 2 reveal different paths of three major car producers throughout the analysed period. During the Crisis period Volkswagen (VOW) made a large acquisition and other changes in governance of the company which cannot be reflected in the general economic indicator immediately.

Volkswagen took a different path, especially during the crisis year. This was caused by extraordinary events, such as the acquisition by and of Porsche and following restructuring processes [14]. In the 2016 “Dieselgate” affair harmed the price, too. These company-related events would lead to unreliable outcomes as they cannot be depicted in the economic indicator designed to measure general condition. Therefore, we only analyse BMW in this paper. An inclusion of other sectors important for German’s economics, such as finance or energy, would require special treatment (such as seasonality adjustments), and is, therefore, omitted.

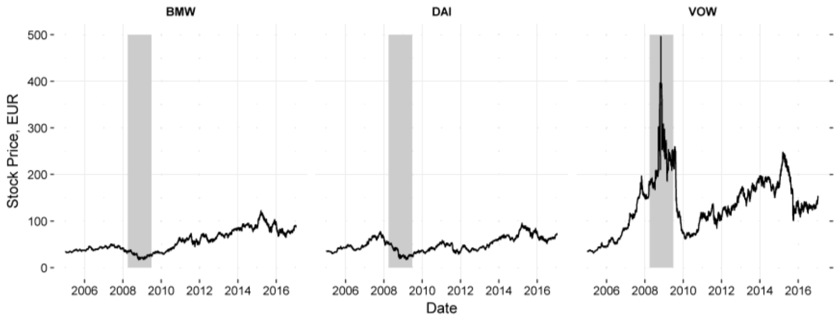

Figure 2

Development of stock prices of BMW, DAI and VOW during analysed period Source: Own processing

ZEW is a top-level indicator of expected Germany's economic performance. It is constructed as a composite indicator of views of 350 experts from various sectors. The experts are asked to express their opinions about general economic conditions during upcoming 6 months. Negative values indicate negative outlook. ZEW indicator is usually released in the middle of the month.

Huffner et al. [15] compared ZEW to other sentiment indicators, such as IfoExpectations. They concluded that for longer predictive horizons (3-12 month) ZEW-based models beat predictions of naive models and models based on the Ifo indicator. We consider the start of the foreseen three-months-period the first day of the month after the release.

We investigate whether sentiment index ZEW has had a good predictive power during the period 2006-2016. To assess the predictive ability, two research questions (RQ) were asked:

RQ 1: How did the ZEW Index predict BMW share price changes in the Pre- and Crisis periods?

RQ 2: How many false alarms did ZEW send to investors in the Post-Crisis period?

The first two research questions investigate whether ZEW index warned of Crisis before it had started and whether it predicted the end of the Crisis in time.

Next three research question focus on the Post-Crisis period and usefulness of ZEW as a standalone and complementary indicator in econometric models:

RQ 3: How good is a predictive ability of a model with only lagged ZEW values on DAX, TecDAX and BMW variables in the Post-Crisis period?

RQ 4: Does inclusion of lagged values of GDP, ZEW and modelled variable improve forecasting model (from RQ3) in the Post-Crisis period?

RQ 5: How does a change of ZEW value translate into the change of DAX, TecDAX and BMW values in the Post-Crisis period in the extended model from RQ4?

To answer the RQ1, visualisation of ZEW values and price returns computed using Equation 1 will be done. An indication of a good predictive power would be if combinations of ZEW values and returns with the same sign appear more often than combinations with mismatching signs.

Visual representation of time series may provide important insight although it does not provide any inferential test. It is, however, possible to assess predictive ability by looking at confusion matrix which summarises combinations of growth expectations and true values of growth. This approach was used by [16] in a 2x2 contingency table case (up-down) and further expanded in [17] to 4x4 case (Peak-Through-Upward-Downward tendency). Given the limited number of observations in the "Crisis period" Fisher Exact test was used to assess predictive power. Monte Carlo simulation with 2000 replications will be performed to estimate p-value. ZEW indicator will have a predictive ability if the positive association is found. RQ2 will be answered from the 2x2 confusion table described above.

Price analysis is based on logarithmic daily returns defined as:

$$\boldsymbol{R_t} = \ln(\boldsymbol{P_t}) - \ln(\boldsymbol{P_{t-1}}) \tag{1}$$

Difference between P_t and $P_{t-1}$is 1 trading day. Maximum difference between two trading days was 6 weekdays.

For the predictive model suggested in the RQ4 we use indicators that are released monthly. GDP, which is released quarterly, was smoothed by third polynomial spline to interpolate missing values. This imputation introduces autocorrelated values, which will be handled later in the Vector Autoregressive model.

Two models will be fitted and the predictive ability will be assessed by comparing mean squared error (MSE) on the test sample. Models will be trained on the period 7/2008 – 12/2015 + month ω. Prediction will be computed on the period 12/2015 + month ω to 12/2015 + month ω $+$ p. Average value of the MSE values on all predicted periods will decide the model's quality. We have 12 growth values to be predicted in 2016. If the best model has$\mathrm{p} = 3$ we will have first MSE for period 1/2016 – 3/2016. The second model will be trained on data from 7/2008 – 1/2016 and predictive period on which second MSE will be computed is 2/2016 – 4/2016. Best model will be found as an average of 10 MSE values.

First model (RQ 3) is a lagged model of form:

$$\boldsymbol{y_t} = \boldsymbol{\alpha_0} + \boldsymbol{\alpha_1 x_{t-1}} + \cdots + \boldsymbol{\alpha_p x_{t-p}} + \boldsymbol{\epsilon_t} \tag{2}$$

Where$\mathrm{y_t}$is a value of an endogenous variable. As stated in RQ3, those endogenous variables will be DAX, TecDAX and BMW. For each variable a separate model will be created.

Second model (RQ 4) will take a form of Vector Autoregressive Model (VAR). Matrix representation of the model:

$$\mathbf{y_t} = \mathbf{v} + \mathbf{A_1}\mathbf{y_{t-1}} + \mathbf{A_p}\mathbf{y_{t-p}} + \mathbf{D} + \boldsymbol{\epsilon}_\mathbf{t} \tag{3}$$

where $\mathrm{y_t}$ is a $\mathrm{m} \times 1$ dimensional random vector of m variables. $\mathrm{A_k}$ are coefficient matrices for $\mathrm{k} = 1, \dots, \mathrm{p}$ of size $\mathrm{m} \times \mathrm{m}$. Matrix D contains seasonal dummy variables multiplied by corresponding seasonal coefficients. To test RQ4 we select ZEW, GDP and one of DAX, TecDAX or BMW variables. Therefore, in our model $\mathrm{m} = 3$. [18]

A value of maximum lag$p$ will be estimated automatically by optimisation rule based on AIC as described in [19]. Theory-based selection might not lead to the best predictive model. This does not represent a problem as RQ4 asks about prediction and is not concerned about real-world meaning of the parameters.

Correlations between variables will be analysed by means of cross-correlation up to lag ± 12 ($\pm$one-year period). Statistical testing of correlations is highly influenced by the underlying nature of the time series. The analysis will be done on detrended values. Time series can be considered as independent if the cross-correlations are centred around lag 0. [20]

Predictive power will be used to assess causality of time series on another in the Granger's sense (whether an addition of predictor's lagged values in addition to lagged values of original series improves predictions). As some of the original time series are non-stationary, an adjusted version of the Granger test based on Wald's test developed by Toda and Yamamoto [21] will be computed.

Effect of change of endogenous variable (RQ 5) in the model from RQ 4 will be estimated by Impulse Response Function (IRF). IRF measures reactions of modelled variables on the change of the selected variable over the $t + j$ predictive horizon. IRF is used to estimate response reaction only. It shall not be interpreted as a cause-effect analysis as a change of ZEW does not affect change of BMW prices or GDP growth.

To ensure that VAR model is estimated correctly and the IRF is meaningful, time-series will be transformed to be covariance-stationary. Augmented Dickey-Fuller test will be used to test for the presence of stochastic trend (unit root). Tests will be conducted for in three settings, i) unit root tests only, ii) unit root and bias (drift) test and iii) as a combination of unit root, bias and trend.

4 Results

To answer the RQ1 Figure 3 was constructed. In this figure the relation between values of ZEW index and mean values of returns is depicted. Mean value is computed on the three months period starting after the ZEW value is released. To emphasize varying uncertainty of the daily returns, point size shows standard deviation of returns in the corresponding period. Larger size means higher uncertainty as the prices were more volatile.

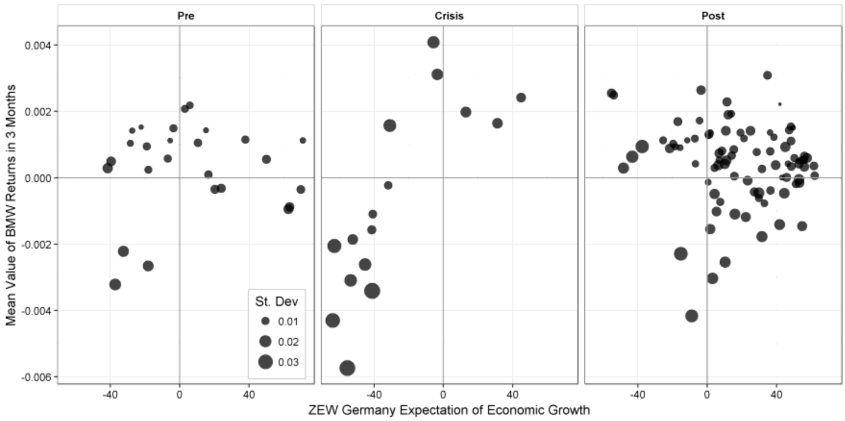

Figure 3

Three windows contain scatter plots of ZEW-Returns variables. Source: Own processing

Visual inspection reveals that lower ZEW values are associated witha higher variance of returns. Table 1 summarises expected direction of BMW's returns in the period of positive/negative ZEW value. There were 13 months with positive ZEW values in the Pre-Crisis period. Out of those releases, 8 following three months periods resulted in positive growth. Conversely, there were 13 months with negative ZEW figure but in 10 months turned to have positive returns. This error represents an upside risk. It expresses higher pessimism of the general market conditions than the actual performance of the stock price.

This erroneous evaluation during the Pre-Crisis period can be attributed to the foreseen crisis. As can be seen from the Figure 4, an expectation of BMW price drop was wrongly timed. The first point in Figure 4 labelled as 2006 corresponds to the late December value of 2005 and is considered as a starting value for the first year. The first point after 2006 summarises values achieved in January 2006 or in the period starting from January (bottom panel). Inspection of the upper panel of monthly returns does not provide much evidence for price-drop in the Crisis period. Returns oscillated around value 0. Reading the second panel, ZEW started losing from February 2006 on the value 69.8. Such a high value was not reached since then.ZEW rose for the first time after ten months in December 2005, when it was already in negative values. Despite the sharp rise in 6 consecutive months in the Crisis period, only in three months positive returns were achieved (top panel). Yet, average returns on the three months predictive window were all positive (bottom panel). When the ZEW reached local maximum in May 2007, future returns were below zero. Since this time to one month before Crisis period, ZEW was steadily decreasing and returns we rather negative across the period. Lowest values of returns were not predicted at all as they occurred in the same month ZEW was released (see the lowest values of the first and second panel).

Table 1
Prediction of the returns' trend in relation to absolute value of ZEW in the Pre-Crisis, Crisis and Post-Crisis periods

	ZEW					
	Pre-Crisis		Crisis		Post-Crisis	
Returns	Negative	Positive	Negative	Positive	Negative	Positive
Negative	3	5	10	0	2	22
Positive	10	8	3	3	16	47

Source: Own processing

Although ZEW started growing since November 2008, the first value indicating positive market sentiment has appeared in April 2009. By this time, monthly returns were positive, so the future returns. If the analyst would not be interested only in the absolute value, but also about the rate of change of ZEW, future positive returns in the upcoming recovery would be discovered sooner. From the economic perspective, positive change in negative ZEW values is a sign of improvement but should be interpreted as the negative consequences will bear lower costs than in the previous period. However, it still talks about expected costs rather than profit. BMW recovered faster than ZEW expected, as the last six points in the bottom window of Crisis period confirm (from those only three months reached positive ZEW values). These results show that ZEW index warned before declination of prices, but (perhaps) too early and did not recognise the price recovery. In the After-Crisis period, ZEW valueswere mostly positive, except of two periods in the end of 2012 and second quarter of 2013.

Formally, three Fisher tests were performed. Inconclusive results were reached in the Pre-Crisis period (sample odds-ratio of $t = 0.493$ on the sample size 26 values is not strong evidence enough to reject null of true odds $\theta = 1$ which indicates no association (p-value is 0.678). In Post-Crisis period $t = 0.27$, p-value=0.1363). Since there is a cell in the confusion matrix in the Crisis period which contains 0 combinations, odds ratio of the sample is not computed. We can, however, provide a p-value, which is derived from marginal sums. This p-value=0.0357 leads to rejection of the null hypothesis on standard $\alpha = 0.05$ level.

Qualitative analysis in Table 1 answers the RQ2. Contingency table suggests that 3 months ahead predictive power was reasonably strong in the Post-Crisis period. From the 63 periods with positive returns, 47 (75%) reference ZEW values were positive. Remaining 25% represent False alarm rate. Economic conditions were expected to be negative while the returns in the forecasted period were positive.

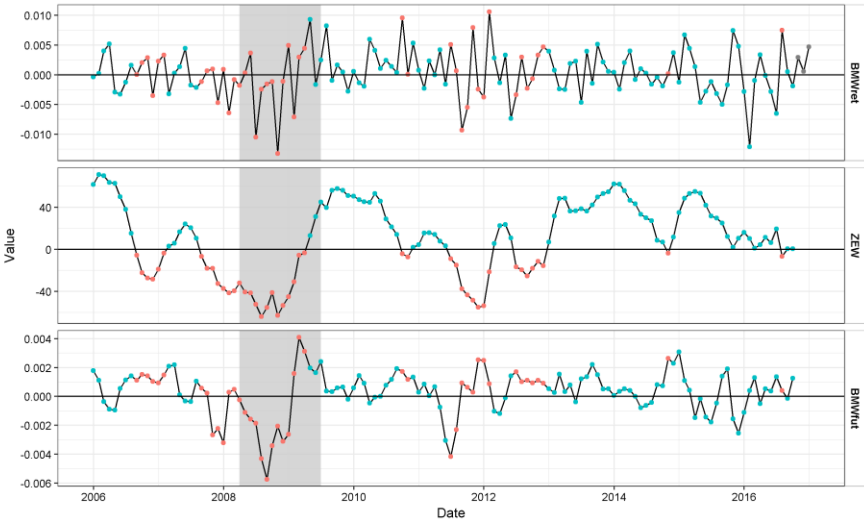

Figure 4

Three windows provide comparison of monthly returns (upper), ZEW value and mean value of future returns (bottom). Colour of the point indicates whether ZEW value was positive. Source: Own processing

Another perspective on usefulness ZEW indicator provides single-variable model defined in Equation 2. Model was estimated on integrated data to allow comparison of VAR models' MSE. These models need to be estimated on stationary times series as the IRF was computed in the last step. To achieve the stationarity, first differences were computed, see Figure 5. Visual inspection suggests that TecDAX is heteroscedastic. This might negatively affect predictive ability of the VAR model.

Stationarity was checked by ADF test with lag up to 3 months. Results ofτ, the test statistics of unit-root stationarity, are provided in Table 2. Tau values indicate whether integrated time series do not contain a unit root. Stochastic trend column contains results of ADF test which was based on lag one value. The last column identifies whether time series is stationary after drift and trend are removed. In all differenced time series drift and trend were missing. The last row presents critical values for test on the given the time series length.

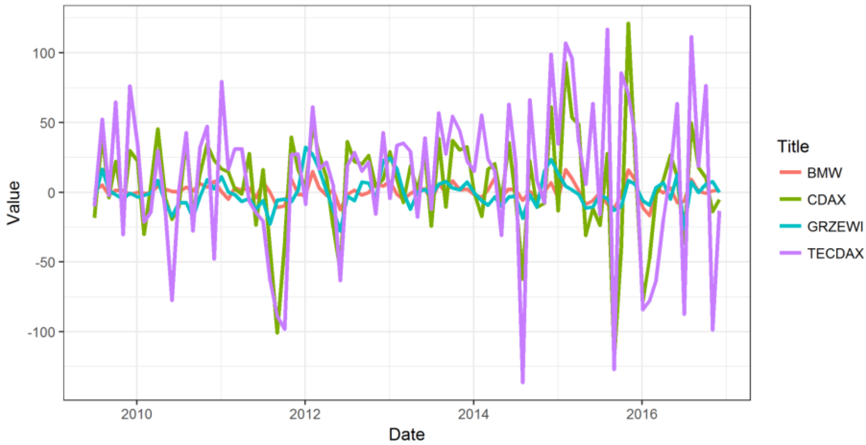

Figure 5
Values of differenced time series. Values of TecDAX are more volatile in the end of the analysed period. This might negatively affect predictive ability of the variable. Source: Own processing

Table 2
Results of ADF test. Presence of stochastic trend (unit root test), drift and deterministic trend was tested

	Lag	Unit root (UR)	UR+Drift	UR+Drift+Trend
TecDAX	1	-5.13	-5.53	-5.57
	2	-3.85	-4.22	-4.26
	3	-3.56	-3.92	-4.00
CDAX	1	-7.43	-7.68	-7.63
	2	-5.49	-5.79	-5.76
	3	-4.47	-4.75	-4.72
BMW	1	-7.85	-7.89	-7.97
	2	-6.11	-6.18	-6.33
	3	-5.27	-5.36	-5.54
GRZEWI	1	-5.07	-5.06	-5.03
	2	-5.28	-5.26	-5.24
	3	-5.14	-5.13	-5.11
Crit. Val 5 pct		-1.95	-2.89	-3.45

Source: Own processing

Figure 6 shows predicted values of the single model with p = 3. As expected, the worst performance was achieved on TecDAX index which has the highest variability in the training sample. Data points are scattered manually on to demonstrate differences on forecasted models in both plots. These models were identified on different sample sizes, depending on the valueω.

All predicted time series were mean-reverting. Model which utilised only one lagged value as a predictor performed the best on the DAX series. This can be seen in Figure 6 where the predicted values are copying the latest development of the time series, change of DAX. In case of TecDAX lagged values turned to have only a limited predictive power and the forecasting trajectory fluctuates around long-term mean value of change.

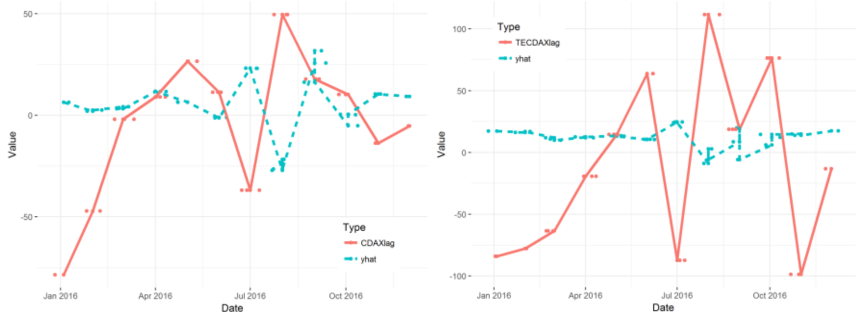

Figure 6

Predicted values from simple lagged model. Dashed lines are predictions based on the past data of lag 3. Solid line shows real value of time series. Composite DAX and Technological DAX are plotted. BMW is not plotted as shows similar pattern as TecDAX. Source: Own processing

Relation of ZEW to other time series is described by cross-correlation plots (with 5% significance bounds). This analysis was conducted on the Post-Crisis period due to the data availability.

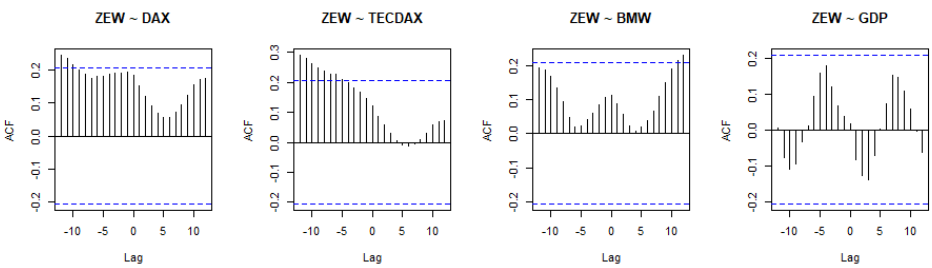

Figure 7

Cross correlations spanning periods of 24 months. Although DAX and TECDAX exhibit similar growth patterns throughout the whole Post-crisis period, shape of correlogram differs. Source: Own processing

Symmetry (based on visual inspection) in the windows of Figure 7 is reached only in case of BMW. If the ZEW takes above average values, BMW tends to have above average returns 12 months later. For both DAX and TECDAX ZEW is a leading indicator. Statistically significant lags start near the lag 6. This corresponds with the aim of the ZEW indicator to capture future trends occurring half a year after issuing. In the last window, ZEW and GDP exhibit a strong seasonal pattern. This confirms our decision to include seasonal dummy variables in VAR models.

Granger test based on the Wald test with an adjusted number of parameters [21] were conducted on models with two variables. Tests were conducted for all pairs of variables with the lag length of 3 and 6 months and with the option of monthly seasonality. ZEW does not Granger-cause (and is not Granger-caused) by any variable except GDP. The strongest evidence is for GDP granger causes ZEW on the 3 months lag model ($\chi^2 = 11.85$, p-value=0.01) and ZEW is Granger causing GDP on the 6 lags ($\chi^2 = 15.16$, p-value=0.02).

Whether the inclusion of lagged values of GDP and a dependent variable itself improves predictive power was investigated by using VAR model, specified in Equation 3. For all VAR models lag $\mathrm{p} = 1$ minimised information criterions. Such parsimonious models were more successful than long-lagged models. The quality of forecasts is in Table 3. VAR models were more accurate than simple univariate models. Trailing MSE of a simple model was about 14% higher in the DAX and 17% in the TecDAX case. Interestingly, simple model outperformed VAR in the BMW series. Values in the table refer to an average value of test-sample MSE on updated models. Value of BMW is substantially lower due to different scale [EUR], whereas DAX and TecDAX are measured in points.

Table2

Comparison of prediction quality on the test sample

	Simple lag model (RQ:2)	Vector Autoregressive Model (RQ:3)
DAX	1486	1304
TecDAX	5755	4913
BMW	55	58

Source: Own processing

Table presented above does not convey the whole message. It only evaluates distance from point estimate to the actual values. Following figure contains a probabilistic assessment of a predictive uncertainty in fan charts [22]. The darkest line is a point estimate similar to those that were computed above. The difference is that point estimates in Figure 8 are computed on the whole training period and on the predictive horizon of 12 months. In the detailed analysis, VAR model was re-estimated when new information about remaining two indicators appeared. Moreover, predictive horizon was set to 3 months. All figures revert to the mean value of the integrated time series. As expected, time series with higher volatility has wider 90% intervals.

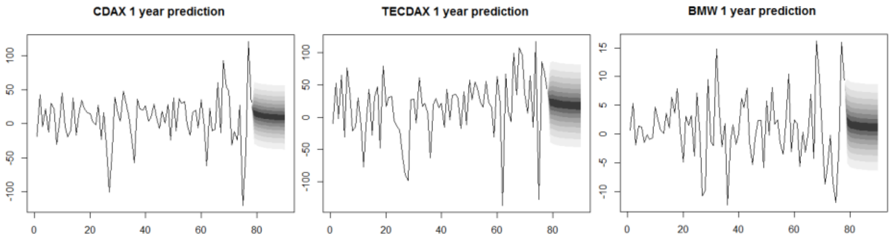

Figure 8

Estimated values of change in the last 12 months. This forecast covers the whole test period. Range covers interval with probability 0.1 to 0.9 (lightest range). Source: Own processing

Outcomes of Impulse Response functions are shown in Figure 9. All variables follow a similar pattern. There is a large change in the first month after the ZEW index is increased. The largest shock is caused to technological companies. Usually, when ZEW increases, TecDAX is increasing by 21 points in the first month. This ZEW's change is also associated with changes in following months, but with the lower magnitude.

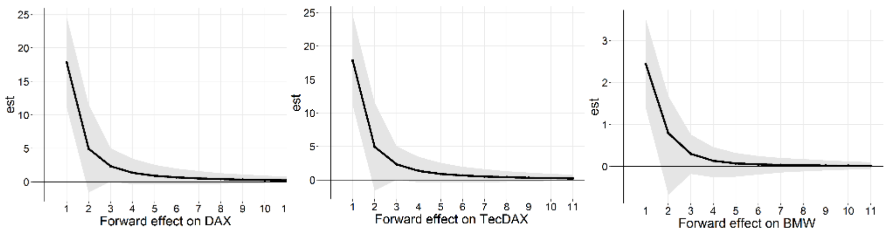

Figure 9

Plots of IRF functions for DAX, TecDAX and BMW (grey area indicates 95% confidence interval of the forward effect). Source: Own processing

Discussion and Conclusions

Five research questions were asked to analyse predictive power of selected sentiment. Answer to *RQ1* revealed that ZEW index warned before declination of prices, but (perhaps) too early. On the other hand, ZEW failed to recognise price recovery in time. Both outcomes are based on absolute value of ZEW. If analysis of changes would be considered, forecasters would discover changing momentum earlier for future returns declinations and recoveries.

RQ2 aimed at predictive power in the Post-Crisis period. Although returns of BMW were positive in 73% of all forecasting periods, relying on ZEW would lead to losses caused by both expecting better market conditions (in 31% of all cases when ZEW presented positive outlook). Performance of ZEW indicator as a standalone predictor has improved in the Post-Crisis period compared to previous periods. False alarm rate reached 25%. One third of three-month forecasted periods were expected to be negative, but BMW price returns were positive.

RQ3 concerned simple univariate models with lagged values. As the best performing models were identified models with lag 1. The importance of lagged variable differed. Its importance in composite DAX time series was higher than in remaining two.

Sentiment indicator ZEW can be considered as a good predictive indicator when accompanied by other indicators, such as lagged values of GDP. Although ZEW indicator is designed as 6 months-ahead indicator, it showed good classification accuracy of the trend direction in the Crisis period on the3-months prediction window. As assigned in *RQ4*, we have quantified the performance difference. Our setup on the rolling MSE on the test sample of the year 2016 showed that in market indexes VAR model outperformed univariate lagged model on the test sample. This does not apply to BMW returns where training MSE was almost identical. Although VAR model is more accurate, provided predictions are quite uncertain. This uncertainty was visualised in fan charts. The price volatility of BMW resulted in the widest range of predictive intervals relative to the scale.

An answer to the last research question *RQ5* provided a sensitivity analysis of shocks in time series. We see diminishing effect of shocks over time in all series. The largest shock is caused to technological companies. Shock in ZEW is associated with changes in other variables for upcoming 3-4 months. This is consistent with design of ZEW indicator which foresees economic sentiment in a short term.

This study has several limitations. Proposed models necessarily suffer from an omitted variable problem as they were intended to demonstrate the usefulness of ZEW indicator. More sophisticated models can be constructed by including more time series and implicit sentiment indicators, interactions or better error-variance handling. For the future research, other sentiment indexes, sectors and companies might be considered.

Acknowledgement

The authors are thankful to the Czech Science Foundation (GA CR), project No. 16-25536S: “Methodology of Developing a Predictive Model of Sector and Company Performance in the Macroeconomic Context”, for financial support of this research.

References

[1] Peiró, A. (2016) Stock prices and macroeconomic factors: Some European evidence. *International Review of Economics & Finance*, 41, 287-294. doi:10.1016/j.iref.2015.08.004

[2] Finter, P., Niessen-Ruenzi, A. &Ruenzi, S. (2012) The Impact of Investor Sentiment on the German Stock Market. *Zeitschriftfür Betriebswirtschaft,82*(0), 133-163, doi:10.2139/ssrn.1650164

[3] Baker, M. & Wurgler, J. (2006) Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of Finance, 61*(4), 1645-1680. doi:10.1111/j.1540-6261.2006.00885.x

[4] Shen, J., Yu, J. & Zhao, S. (2017) Investor sentiment and economic forces. *Journal of Monetary Economics, 86*, 1-21. doi:10.1016/j.jmoneco.2017.01.001

[5] Lux, T. (2010). Sentiment dynamics and stock returns: the case of the German stock market. *Empirical Economics, 41*(3), 663-679, doi:10.1007/s00181-010-0397-0

[6] Verma, R. & Soydemir, G. (2006) The Impact of U.S. Individual and Institutional Investor Sentiment on Foreign Stock Markets. *Journal of Behavioral Finance*, *7*(3), 128-144. doi:10.1207/s15427579jpfm0703_2

[7] Spiwoks, M. (2004) Die Verwendbarkeit der ZEW-Aktienindex-Prognosenfür active Portfoliomanagement-Strategien / The Usefulness of ZEW Stock Market Forecasts for Active Portfolio Management Strategies. *Jahrbücher fü rNationalökonomie und Statistik, 224*(5) doi:10.1515/jbnst-2004-0503

[8] Zhong, X., &Enke, D. (2017) Forecasting daily stock market return using dimensionality reduction. Expert Systems with Applications, 67, 126-139, doi:10.1016/j.eswa.2016.09.027

[9] Yu, J. & Yuan, Y. (2011) Investor sentiment and the mean–variance relation. *Journal of Financial Economics, 100*(2), 367-381, doi:10.1016/j.jfineco.2010.10.011

[10] Stambaugh, R., Yu, J. & Yuan, Y. (2012) The Short of It: Investor Sentiment and Anomalies. *Journal of Financial Economics* (104) 288-302, doi:10.1016/j.jfineco.2011.12.001

[11] Checkley, M., Higón, D. A., &Alles, H. (2017) The hasty wisdom of the mob: How market sentiment predicts stock market behavior. *Expert Systems with Applications*, 77, 256-263, doi:10.1016/j.eswa.2017.01.029

[12] Nassirtoussi, A. K., Aghabozorgi, S., Wah, T. Y., & Ngo, D. C. (2014) Text mining for market prediction: A systematic review. *Expert Systems with Applications*, 41(16), 7653-7670, doi:10.1016/j.eswa.2014.06.009

[13] Germany GDP, Current Prices: Eurostat, Gross Domestic Product for Germany© [CPMNACNSAB1GQDE] retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPMNACNSAB1GQDE, Febr. 17, 2017

[14] Akpinar, M. & Vincze, Z. (2016) The dynamics of coopetition: A stakeholder view of the German automotive industry. *Industrial Marketing Management, 57*, pp. 53-63, doi:10.1016/j.indmarman.2016.05.006

[15] Hüfner, F. P., &Schröder, M. (2002) Prognosegehalt von ifo-Geschäftserwartungen und ZEW-Konjunkturerwartungen: EinökonometrischerVergleich / Forecasting German industrial Production: An Econometric Comparison of ifo- and ZEW-Business Expectations. *Jahrbücher für Nationalökonomie und Statistik,222*(3) doi:10.1515/jbnst-2002-0303

[16] Schnader, M. H. &Stekler, H. O. (1990) Evaluating Predictions of Change. *The Journal of Business, 63*(1), 99-107, doi:10.1086/296486

[17] Tsuchiya, Y. (2016) Do production managers predict turning points? A directional analysis. *Economic Modelling*, *58*, 1-8, doi:10.1016/j.econmod.2016.05.019

[18] Pfaff, B. (2008a) *Analysis of integrated and cointegrated time series with R*. New York: Springer, doi:10.1007/978-0-387-75967-8

[19] Pfaff, B. (2008b) VAR, SVAR and SVEC Models: Implementation Within R Package vars. *Journal of Statistical Software, 27*(4) doi:10.18637/jss.v027.i04

[20] Dean, R. T., & Dunsmuir, W. T. (2015) Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models. *Behavior Research Methods*, 48(2), 783-802, doi:10.3758/s13428-015-0611-2

[21] Toda, H. Y., & Yamamoto, T. (1995) Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, 66(1), 225-250, doi:10.1016/0304-4076(94)01616-8

[22] Britton, E., Fisher, P., & Whitley, J. (1998) The Inflation Report projections: understanding the fan chart. *Bank of England Quarterly Bulletin*, 38:30-37