KU LEUVEN

FACULTY OF ECONOMICS AND BUSINESS

Predicting Social Unrest and Financial Market Collapse

The Effect of Conflicts recorded by the GDELT Dataset on Indian BSE30

Michel Ballings

Master of Business Administration with main subject in strategic management

2018-2019

Content

Abstract

Firstly, in our study, we take a closer look at the effects of terrorist attacks on the European stock market. We test this by comparing the returns of the Eurostoxx50 with those of the MSCI with an event study. Our results show, however, that the differences in returns are insignificant; this partly can be explained by the nature of the index itself.

Secondly, this study investigates the effect of political conflicts on the variance of the returns in a country that enjoys a high level of violence. In our study, we choose the country India with its corresponding index BSE30. We examine the market returns using a count of conflicts which are recorded in the GDELT-dataset with GARCH-models as methodology. We show that taking into account the political conflicts as a variable; it can still improve the modeling quality of markets with high variances of violence.

Introduction

It's been in general interest to know how and in what extent a social and political conflict (such as a terrorist attack) affects financial markets. Especially after 9/11, researchers have largely been interested in how the markets respond. Among all the researches published to address a market reaction, the large majority of studies have failed to reject the null hypothesis that the markets do not respond significantly to variations in violence over time. However, these researches were mainly focused on western markets that are said to be efficient. Recent studies (Yonamine, 2013) failed to reject this null hypothesis for markets that are located in countries where large variations of violence are not uncommon (Israel for example).

In this research paper, we take a closer look at the effect of social and political conflicts in India. Even though the level of conflict and political violence has significantly fallen in recent years, India remains a high-risk country. India is rated 18 on places most risky to do business in according to the Conflict and Political Violence Index for 2014 compiled by London-based risk analytics firm Maplecroft (Dhoot, 2014). "India continues to enjoy a significant improvement in the level of conflict and political violence. A significant fall in the risk of terrorism has contributed to this […] in particular, by a significant fall in the frequency and lethality of Naxalite attacks following a surge of violence in 2009 and 2010," so the firm's principal political risk analyst Charlotte Ingham.

Thus, while the existing literature has provided consistent findings that financial markets in stable countries respond significantly to political violence, less is known about how financial markets in conflictual countries respond to violence. Therefore, we provide empirical tests to measure the extent to which variation in the level of violent attacks against India affects variance in returns of the BSE30.

Before testing the variances of the BSE30 in India, we'll first take a look onto the reaction on European stocks after a terrorist attack. Now we test whether our results are consistent with the existing literature and whether the European markets are, in fact, as integrated as we think they are. We now compare the returns of the Eurostoxx50 (the 50 biggest stocks traded on European stock exchanges) with the MSCI (which includes 1,649 stocks from companies throughout the world) while using an event study.

1 Literature Review and Hypotheses Formulation

At the beginning of this research, a deep analysis of the most recent studies (as for 2018) must be made. Since Google's dataset GDELT will be present through the whole research paper, the most logical decision will be to start the analysis with the capabilities of this dataset.

1.1 Global News Coverage

Kwak & An (2014) analyzed which disasters receive a great deal of attention from foreign news media. For the study, a large-scale news media coverage dataset, GDELT (Global Data on Events, Location, and Tone) was used. GDELT is a recently developed event dataset containing more than 200 million geolocated events with global coverage since 1979 (Leetaru & Schrodt, 2012). To select relevant variables, they used the Theory of Newsworthiness as a basis. Thereby Galtung & Ruge (1965) explained which factors influence the newsworthiness of an event. The suggested factors are frequency, intensity, unambiguity, meaningfulness, consonance, unexpectedness, continuity of an event, and some characteristics of an actor involved in the event.

The model of Kwak & An (2014) explains 25,4% of the variance of global news coverage, which is comparable with previous studies in the area (e.g. Wu, 2000). The study provides a basis for selecting relevant variables that predicts global news coverage. However, the number of news articles about disasters is readily influenced by US news media. In the model, they neutralized this possible bias by defining global attention to a disaster as the number of countries covering the disaster. Although, it does not say anything about the effect US news media has on getting global attention.

1.2 GDELT as Social Conflict Predictor

Yet, an important question remains unanswered, namely if GDELT is, indeed, capable of predicting any social conflict at all. An attempt to address the task of identifying, understanding and predicting when social unrest might occur have been made by Galla & Burke (2018). They found that news reflects society and can be used to detect building unrest. Besides that, their main contribution is to take into account a thoroughgoing picture of society to predict unrest level at the country level of location compared to previous work that looked at larger geographical regions. To test their hypothesis, they used GDELT which contains structured data that is mined from broadcast, print and web news sources in more than 100 languages since 1979.

The results of their study showed that Random forest is performing very well with good F1 scores at both State and County levels. This means that machine learning techniques are a good attempt to identify regions at state and county levels where social unrest might occur in the near future. They also indicate that GDELT essentially can be used as a tracker in studying events of interest; especially huge social unrest events. Even though the use of GDELT data is much promising, it can be used as a complimentary study to other recent studies which are based on an analysis of social networks.

1.3 Financial Markets and Terrorism

There has been done a lot of research that the effect of a political conflict (such as an act of terrorism or riots) statistically negatively influences the financial market (such as Rigobon & Sack, 2005; Schneider & Troeger, 2006 and Zussman, Zussman, & Nielsen, 2008). However, a terrorist attack doesn't only influence the financial markets as such, it does even more. Apergis & Apergis (2016) make an interesting attempt at whether international defense industries are significantly affected by such an attack. Here the empirical findings of Apergis & Apergis (2016) suggested that there was a positive significant reaction on the stock prices the next day, as well as on the days following the event.

Also after the bombing of the international airport in Brussels, Belgium, the question arose if airlines were also affected. An answer to this question was given by the research of Kolaric & Schiereck (2016) where they analyzed the reaction on airline stocks after the terrorist attacks (in Paris and Brussels). Here they concluded that the attacks had strong negative effects. However, the effects were statistically smaller following the Brussels strikes.

It's most likely that terrorist attacks negatively affect the returns of the country's financial markets. However, it would be interesting to see if the returns of the European market are also affected by a terrorist attack made in a country within Europe. An example could be whether the attacks on Nice in 2016 also affected the European market (e.g. Eurostoxx50). We can do this by comparing the CAR of the event period with a normal period, whereby the CAR stands for cumulative abnormal returns (as seen in Brown & Warner, 1985 and Yermack, 1997). Now we come to the first hypothesis, and we will test it for all attacks ISIS claimed responsibility for:

Hypothesis 1:

> The CAR of Eurostoxx50 is statistically not different from zero after a terrorist attack in an EU country.

1.4 Analyzing Variation in Index Returns by using GDELT

After ensuring that GDELT-dataset is a good predictor of social unrest and conflicts, the question then rises if there already has been published any studies analyzing variation in index returns by using GDELT.

In the previous researches, the authors only focus on financial markets where the level of violence is relatively low (e.g. Dow Jones Industry Index in the US). The research of Yonamine (2013) therefore makes an interesting attempt to find out the effect of violence on financial markets that are regularly affected by conflicts by using GDELT. In his study, he uses the TA100, which is the officially market index of the 100 biggest companies in Israel. Since Israel is often the victim of political violence, the TA100 is an ideal market to test the hypothesis on. Besides an analysis of the market, he chooses two particular (insurance) companies of the TA100 that might be affected by political violence.

He finds that, on average, variance in returns of TA100 is *not* significantly driven by levels of violence committed against Israel. Besides that, he strongly suggests that while the TA100 may not meaningfully respond to variation in attacks against Israel, specific companies do. In general, the research of Yonamine (2013) is a good counterpart to traditional literature. Even though his paper is a good attempt to explain the effect of political conflict on the TA100, there is some room for improvements. First, he uses the Dow Jones industry index as an indicator of the global economy as a control variable. A better option to measure the global economy might be the S&P500 index, where the biggest 500 companies of the world are listed. Second, he uses weekly returns of the TA100 instead of daily returns. This means that his research, indeed, doesn't find a statistically significant relationship of political conflict on TA100 on a *weekly* basis. However, it could be possible that there is a significant effect on a daily basis, but that the index recovers faster.

After analyzing the paper of Yonamine (2013), the author makes a good point that his findings may not be consistent with other countries where political violence is common but tends to be low-scale. One of these countries that fit this description can be India, where the index of the 30 biggest companies is called BSE30. Relating to this conclusion, we can formulate 2 other hypotheses:

Hypothesis 2:

> Variation in the level of political violence in India will have a statistically significant impact on the variance in returns of the BSE30 on a *weekly* basis.

Hypothesis 3:

> Variation in the level of political violence in India will have a statistically significant impact on the variance in returns of the BSE30 on a *daily* basis.

In another chapter of Yonamine's doctoral dissertation (2013), he tries to forecast future levels of violence by the GDELT dataset. An advantage of using Google's GDELT dataset is that it also provides geospatial information where a conflict happened. Before, this information was only available by human-coded data. To give an example, research has been done to predict conflicts events by Weidmann & Ward (2010) with ACLED's Bosnia dataset. The research was much promising, although the study outlined the slow, tedious nature of human-coded datasets. However, by using the GDELT dataset in the study of Yonamine (2013), the author provides a good basis for future research of predicting conflicts.

In other words, we try to highline the advantage of GDELT's future existence and uniformity. An example is the study of Zammit-Mangion, Dewar, Kadirkamanathan, & Sanguinetti (2012) whereby the researchers use Wikileak's dataset of predicting future levels of violence in Afghanistan. They find that 62.5% of actual levels of violence fall within 95% confidence intervals of predicted levels. Although the study applies innovative methods, it faces the problem that Wikileaks isn't an ideal database since it doesn't provide real-time updates. Also, the fact that a lot of entries are not legal also raises questions of sustainability. While the accuracy is likely to be greater for the WikiLeaks dataset (since it is based on firsthand accounts), the ongoing debate still exists regarding the accuracy of human-coded and machine-coded datasets. This suggests that neither human-coded nor machine-coded datasets have a clear advantage (see Chojnacki, Ickler, Spies, & Wiesel, 2012; Eck, 2012; King & Lowe, 2003; O'Brien, 2010; O'Loughlin, Witmer, Linke, & Thorwardson, 2010; Schrodt, 2012 for these discussions). In the study of Yonamine (2013), the author tries to predict future levels of violence in Afghanistan, albeit it is possible by using GDELT to choose other countries. He trains 3 different ARIFMA-models (on district-, province- and countrywide level). Whereas the results are promising on a district level, here, the model outperforms his naïve model in 47 out of the 48 out-of- sample months. On the other levels, the model is still an improvement; however, it's less effective than on the district level.

1.5 Other Independent Variables

After finding enough evidence that (political) conflict (provided by the GDELT-dataset) is a good indicator for predicting social- and financial unrest, we will expand our research by looking for other possible independent variables.

Finding independent variables to predict the stock markets are difficult. However, in a study by Zhang, Fuehres, & Gloor (2011), the researchers try to find a correlation of the usage of certain words that explains mood in Twitter-messages. According to a previous study by Gilbert & Karahalios (2010), the emotional state can influence the decisions (which also includes stock market investment decisions). While keeping this study in mind, Zhang et al. (2011) find a (negative) correlation between the words 'fear', 'hope' and 'worry' in tweets and the returns of the Dow Jones stock market on the next day. Thus, the study highlights that it is possible to predict the returns of the stock market by checking emotional outbursts on twitter. However, the research doesn't give a solution to the closing of the financial markets on the weekends. Twitter doesn't close on the weekends, but it isn't realistic to take the tweets that have been tweeted on Sunday as an indicator on the returns of the stock markets on Monday.

In this context, a recent study of Elshendy, Fronzetti Colladon, Battistoni, & Gloor (2018) is an interesting example of an attempt to developing a forecast model of crude oil prices. What made this study extraordinary was the fact that they trained an ARIMAX-model by using not only the information from one platform but a combination of different platforms. In contrast, a lot of previous studies only train their model by using one information platform. An example of such a study is the one of Zhang et al. (2011), where the auteurs collected 6 months of Twitter feeds and finding a significant correlation between the crowd's emotion and all the dependent variables (which were gold price, crude oil price, currency exchange rates and stock prices). In the study of M. Elshendy et al. (2018), however, the authors gave evidence to the advantages of integrating information from different platforms to realize their predictive models. In other words, training a predictive model by just relying on a single information platform proved to be less accurate.

According to their combined model, they also proved that a combination of Twitter, Wikipedia and GDELT can lead to forecast crude oil prices with a reasonably high level of accuracy. Whereas a higher number of articles on GDELT, as well as more messages related to the oil price on Twitter and a higher number of page views on the Wikipedia pages, showed a negative relation with the oil price. They also show that Google Trends emerges as the best time series predictor for the WTI crude oil price, followed by Twitter. Even though the authors of the article make an interesting point on the combination of multiple information platforms, there are some limitations. A limitation on the Twitter-data is that the authors didn't include the weight of the popularity of the users. This is, however, a logical conclusion since a popular Twitter-user has a larger influence than a passive user. In relation to the GDELT-data, an improvement could be including the number of mentions, which counts the number of information sources containing mentions of the event.

We could, therefore, include multiple datasets (such as Google Trends and Twitter) in our model to test our hypothesis. However, we'll start our research by mainly using the GDELT-dataset which was much promising in studies such as Yonamine (2013) and Elshendy et al. (2018). Collecting more data such as a detailed sentimental analysis of Twitter wouldn't be possible in the timeframe given for our research. We, therefore, continue with building our model based on the GDELT-dataset.

1.6 Forecasting Methods

Another question that arises during the paper is if it is actually possible to forecast the returns of a financial market such as the Indian BSE 30 market index. As discussed in supra, Galla & Burke (2018) proved that machine learning is a good method to predict social unrest. This leaves us the question of what the best option is for our model. Dunis, Laws, & Karathanasopoulos (2012) investigated the use of alternative novel neural network architectures while forecasting the ASE 20 Greek Index. Their conclusion collaborates with those of Dunis, Laws, & Sermpinis (2010) and Lindemann, Dunis, & Lisboa (2005) where HONN (Higher Order Neural Network) demonstrates superior forecasting in comparison with Multilayer Perceptron (MLP) and Recurrent Neural Network (RNN).

Also using hybrid models or combining several models is common to improve the forecasting accuracy since the well-known M-competition (Makridakis et al., 1982). Empirical findings suggest that combining different methods can be an effective and efficient way to improve forecasts (Makridakis, 1989; Newbold & Granger, 1974; Palm & Zellner, 1992 and Winkler, 1989). Research in time series forecasting argues that predictive performance improves the combined models (Bishop, 2013; Clemen, 1989; Hippert, Pedreira, & Souza, 2000; Terui & Van Dijk, 2002; Tseng, Yu, & Tzeng, 2002; Zhang & Qi, 2005 and Zhang, 2003).

As discussed in supra, Dunis et al. (2012) also included RNN in their research for the superior forecasting model. Especially in a financial application, Kamijo & Tanigawa (1990) applied RNN successfully to the recognition of stock patterns of the Tokyo stock exchange. Not only for forecasting stock exchanges RNN was successfully applied, but Tenti (1996) was also able to forecast the exchange rate of the Deutsche Mark successfully. Other remarkable and successful forecasts with an RNN-approach are from Tiňo, Schittenkopf, & Dorffner (2001) with the volatility of the DAX and the FTSE 100 using straddles and Dunis & Huang (2005) with the volatility from the currency options market on GBP/USD and USD/JPY.

However, in the study of Dunis et al. (2012), the authors concluded that a HONN was superior to both MLP and RNN. HONN was first introduced by Giles & Maxwell (1987) as a fast learning network with increased learning capabilities. Although HONN has a superior forecasting ability, the use of the model so far has been limited. Some examples of studies where HONN was successfully used for forecasting are Dunis, Laws, & Evans (2016) for the gasoline crack spread and Fulcher, Zhang, & Xu (2006) for the AUD/USD exchange rate.

2 Model

Since we have 3 different hypotheses, we cannot use one general approach to accept or reject the null hypothesis. For the first hypothesis, we'll need a totally different model than for hypothesis 2 & 3. In this chapter, we'll explain what our approach will be according to the different hypotheses.

2.1 Hypothesis 1

2.1.1 Event Studies

To accept or to reject the first null hypothesis we can use an event study. Now we'll compare the returns of the Eurostoxx 50 with the returns of MSCI. The idea is that MSCI represents the global market since the index includes 1,649 stocks from companies throughout the world. Therefore we can test if the returns of the Eurostoxx 50 are significantly different from 0 in the period of the terrorist attack.

2.1.2 Model Specification

To test this hypothesis, we first calculate the (daily) returns of the market indexes:

$R_{it} = [(P_{it} - P_{it-1})/ P_{it-1}]$

Where R_{it} stands for the return on share i during period t, P_{it} for the price of share i at the end of period (adjusted for capitalization changes).

Then we'll need to estimate the normal returns and abnormal returns during estimation window (e.g. [-240,-40]) and event window (e.g. [0, +5]) by using the market model:

$R_{i,t}= \alpha_t + \beta_i R_{mt}+\epsilon_{i,t}$

Whereby $\alpha_t + \beta_i R_{mt}$ is the predicted normal returns and $\epsilon_{i,t}$ are $AR_{i,t}$ (abnormal returns).

By using a T-test, we can test whether we need to accept or reject the null hypothesis (that the event had no effect on the stock price).

2.2 Hypotheses 2 & 3

2.2.1 Usage of GARCH-Model for Financial Data

An empirically justified approach to analyze the variance in financial data is with a GARCH-model (see examples as Alberg, Shalit, & Yosef, 2008; Dhankar & Chakraborty, 2007; Mohammadi & Su, 2010 and Ahmed & Shabri, 2014). Like most indexes, the BSE30 contains high degrees of volatility that tent to cluster rather than follow a random distribution (as with the TA100 in Yonamine, 2013). Whereas a classic autoregressive conditional heteroscedasticity (ARCH) (Engle, 1982) model tends to require a higher order autoregressive error term, generalized ARCH (Bollerslev, 1986) performs better with fewer parameters.

2.2.2 Specifying the GARCH-Model

A GARCH (p,q) model consists of a conditional mean and conditional variance equation, both of which allow for the inclusion of exogenous variables. For consistency, we follow and adopt the notation from Leblang & Mukherjee (2005) (which was also applied for TA100 in Yonamine, 2013). When applied, the conditional mean is as following:

$$\Delta(\ln(BSE30_t) = \lambda + \psi Z_t + \epsilon_t$$

Where $\Delta(\ln(BSE30_t) = \ln(BSE30_t) - \ln(BSE30_{t-1})$, λ is a constant that is approximately 0 due to the first differencing, Z_t is a vector of exogenous variables, ψ is a vector of estimated coefficients, and ϵ_t is the error term distributed $(0,\sigma^2{}_t)$.

The conditional variance is:

$$\sigma^2{}_t = \omega + \sum_{i=1}^{q} \alpha_t \epsilon^2{}_{t-1} + \sum_{i=1}^{p} \beta_i \sigma^2{}_{t-1} + \delta_i I_{i,t}$$

Where ω is a constant, ϵ_{t-i} is the lagged error, σ_{t-i} is the lagged variance, $I_{i,t}$ is a matrix of exogenous variables, and α_i, β_i, and δ_i are estimated parameters.

2.2.3 GDELT Dataset

As discussed above, Google's GDELT dataset is very promising to predict conflicts and social unrest. This is partly because it's the first database that also includes geolocation of an incident. Another advantage is that the dataset is machine coded and not manually coded, which ensures that the speed is adequate. Also, its sustainability is sustained since it includes publically available information (in contrast to Wikileaks with most of its information uploaded illegally). This makes a perfect fit for the requirements set for this analysis. To use the GDELT dataset as an independent variable, we calculate the frequency of conflicts that happened on one day. In this logic, when there are more entries in the GDELT of conflict during one day, it's more likely the market is going to react. Therefore, the independent variable has been labeled 'conflict count' in our model.

2.2.4 Control Variables

Due to a lot of researches on the financial markets, we know that many factors influence financial markets. Even though it is hard to discover a causal relationship, we know that statistically significant relationships exist between the equity markets and factors such as commodity prices, inflation rates, trade, other financial markets and domestic regime type. Interestingly, most studies looking at the relationship between terrorist attacks and financial markets fail to include such control variables. This, therefore, most likely leads to considerably underspecified models thereby systematically overestimating the effects of the terrorist events. Studies that do include control variables in their models are Leblang & Mukherjee (2005) and Schneider & Troeger (2006).

With keeping this in mind, we include 2 different control variables in our GARCH-models, namely the performance of the global market and commodity prices.

For the performance of the global markets, we take the first difference of the logged MSCI-index. The oil prices can perfectly reflect the commodity prices. We, therefore, use the first difference of the logged daily Brent oil prices.

3 Results

3.1 Hypothesis 1

As described above in the methodology, an event study was used to compare the daily returns between MSCI and Eurostoxx50 during a terrorist attack. During the event studies, two time windows were selected, namely 1 and 5. Also, seven different terrorist attacks were analyzed, namely all the attacks ISIS has claimed responsibility for in Europe during 2015 and 2017. Hereunder a table summarizes all important statistics needed for the event studies (which are CAR, standard error of regression or SER and t-test). The critical value of the t-test for 198 degrees of freedom (length of the estimation window minus 2 df) and for a coincidence interval of 95% is 1.65258578.

With event window of 5 [0;5]:

Event	CAR	SER	T-test
November 2015 Paris attacks	.00951208169834	.013685734734637	.283747359
Brussels bombings	-.00643689698093	.013893125414659	-.18914766
Nice truck attack	.01571177269177	.013879326222491	.46214811
Berlin Christmas market attack	.01676964811547	.012183105298678	.56194049
Manchester Arena bombing	-.01351567470163	.010007605609235	-.55135577
2017 London Bridge attack	-.01754740229521	.007746367666503	-.92478143
2017 Barcelona attacks	-.0158856927288	.007381030541713	-.87864518

Table 1

With event window of 1 [0;1]:

Event	CAR	SER	T-test
November 2015 Paris attacks	-.01627526254267	.013685734734637	-.84090104
Brussels bombings	-.00100156955742	.013893125414659	-.05097605
Nice truck attack	.01360978768003	.013879326222491	.69337466

Berlin Christmas market attack	.00066595004877	.012183105298678	.03865171
Manchester Arena bombing	.00128887705793	.010007605609235	.09106811
2017 London Bridge attack	-.01182152179202	.007746367666503	-.0790965
2017 Barcelona attacks	-.01161557607818	.007381030541713	-.11277857

Table 2

None of the above t-tests exceeds the proposed critical value of 1.65258578, which means that they all are insignificant. In other words, we cannot reject the first null hypothesis. This finding would suggest that during a terrorist attack, the returns of the Eurostoxx50 are not significantly different from the returns of the MSCI. There can be three non-exhaustive explanations.

It could be that companies in the whole world react to a terrorist attack committed in Europe; however, this is less likely. Another explanation could be the opposite, which is that on the European level, investors don't take a large interest in a terrorist attack when it's not committed in their own country. The last explanation could lay in the nature of the index. Since the Eurostoxx50, in general, is a diversified stock index, it might lever away this specific risk. Say, when a French investor hears about a domestic terrorist attack, it could sell its domestic stock and therefore invest in another European stock which most likely is also listed in the Eurostoxx50. When this happens, the index doesn't need to react per se.

The existing literature (such as Rigobon & Sack, 2005; Schneider & Troeger, 2006 and Zussman, Zussman, & Nielsen, 2008) suggests that terrorist attacks significantly influence the financial markets. However, we couldn't find any differences between the returns of the Eurostoxx50 and those of the MSCI, hereby proving that we cannot find any significance during a terrorist attack with this methodology on this (European) level.

3.2 Hypothesis 2

```
ARCH family regression

Sample: 2018w2 - 2019w3                         Number of obs     =         54
Distribution: Gaussian                          Wald chi2(3)      =       6.84
Log likelihood =  148.4462                      Prob > chi2       =     0.0773

------------------------------------------------------------------------------
                |                 OPG
        BSE30_c |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
BSE30_c         |
          OIL_c |   .0009499   .0051194     0.19   0.853    -.0090839    .0109836
         MSCI_c |   .1265323   .0610094     2.07   0.038     .0069561    .2461085
ConflictCount_c |  -9.28e-06   .0000119    -0.78   0.436    -.0000326     .000014
          _cons |  -.0026798    .002966    -0.90   0.366    -.0084931    .0031336
----------------+-------------------------------------------------------------
ARCH            |
           arch |
             L1. |   .2900565   .2097251     1.38   0.167    -.1209972    .7011102
                |
          garch |
             L1. |   .5327397   .2636204     2.02   0.043     .0160533    1.049426
                |
          _cons |   .0000476   .0000444     1.07   0.284    -.0000394    .0001346
------------------------------------------------------------------------------
```

Table 3

In table 3, a GARCH-regression can be found with the coefficients and significance intervals of the variables included. The variables included are oil, MSCI and conflict count which already were defined theoretically in 2.2.4. The data in this regression were on a weekly basis and is thereby a model to test hypothesis 2, namely whether the variation in the level of political violence in India will have a statistically significant impact on the variance in returns of the BSE30 *on a weekly basis*. To aggregate this data, the returns and conflict counts were *cumulated* per week.

As we can see in the table, the variable conflict count (=frequency of a conflict recorded in the GDELT-dataset in one week) was not significant in the model, and therefore we need to reject hypothesis 2. This might suggest that, in line with the research of Yonamine (2013), the investors of the BSE30 already discounted their shares for a possible conflict that might happen in the foreseeable future. However, another explanation might be that investors do react on conflicts but that the index recovers faster (and thereby making the abnormal returns insignificant). To test for this, we calculated another GARCH-regression model but with data on a *daily basis* (see hypothesis 3).

3.3 Hypothesis 3

```
ARCH family regression

Sample: 20180220 - 20190220, but with gaps         Number of obs    =        298
Distribution: Gaussian                             Wald chi2(3)     =      24.33
Log likelihood =  1039.751                         Prob > chi2      =     0.0000
```

BSE30	Coef.	OPG Std. Err.	z	P>|z|	[95% Conf.	Interval]
BSE30						
OIL	.0001096	.0008517	0.13	0.898	-.0015598	.0017789
MSCI	.0918342	.0220544	4.16	0.000	.0486084	.13506
GDELT_Conflict_Count	-.0000134	5.54e-06	-2.43	0.015	-.0000243	-2.59e-06
_cons	-.0004023	.0004736	-0.85	0.396	-.0013305	.000526
ARCH						
arch						
L1.	.1314816	.0797883	1.65	0.099	-.0249006	.2878638
garch						
L1.	.7665619	.2436475	3.15	0.002	.2890215	1.244102
_cons	6.30e-06	.0000119	0.53	0.596	-.000017	.0000296

Table 4

In table 4, we can see the results of the GARCH-regression with data on a daily basis. Interestingly, however, is that now the variable 'conflict count' is significant. This suggests that investors, in fact, do react on conflicts on short term (=daily) and we, therefore, need to accept hypothesis 3. This means that the second explanation of hypothesis 2 is true, namely that investors do react on conflicts in India, but that the index recovers quickly. This might explain why the variable 'conflict count' is not significant on a weekly basis (thereby rejecting hypothesis 2) but significant on a daily basis (thereby accepting hypothesis 3).

While retrieving the previous GARCH-regression models (see table 3 and 4), one might suggest economic insignificance due to the small estimated coefficients. The reason for the small estimates are because of the specification of the GARCH-model itself (see 2.2.2). Since the dependent variable is logged, it represents an exponential response model. This needs to be interpreted as follows; when one conflict accrues, the change in the BSE30 will decrease with 0,00134% the next day.

We confirm literature such as Elshendy et al. (2018) that impressively forecasted crude oil by using four different online media sources. In this research, they also proved that using the GDELT-dataset significantly increases the forecast quality for crude oil (see 1.5).

3.4 What about Israel?

On the other hand, other literature (Yonamine, 2013) suggests that there is, in fact, no significance of political conflicts on markets that enjoy higher variances of violence. Whereas Yonamine (2013) only tested his hypotheses on a weekly basis, we also checked on a daily basis in our model. Here, however, we *do* find significance on a daily basis for India (see 3.3). We thereby give proof that using political conflicts as a variable in a forecasting model might, in fact, be significant for *some* markets that enjoy a high level of violence.

```
ARCH family regression

Sample: 20180220 - 20190220, but with gaps        Number of obs   =       298
Distribution: Gaussian                            Wald chi2(3)    =      8.39
Log likelihood = -1067.432                        Prob > chi2     =    0.0385
```

closeta125	Coef.	OPG Std. Err.	z	P>|z|	[95% Conf.	Interval]
closeta125						
OIL	-.386211	.7046875	-0.55	0.584	-1.767373	.9949511
MSCI	-100.1724	37.36371	-2.68	0.007	-173.4039	-26.94088
conflictcount	.0107404	.0185143	0.58	0.562	-.0255469	.0470277
_cons	2.460085	.7188691	3.42	0.001	1.051127	3.869042
ARCH						
arch						
L1.	.0048244	.010143	0.48	0.634	-.0150555	.0247042
garch						
L1.	-1.27121	.3241043	-3.92	0.000	-1.906443	-.6359773
_cons	174.3766	20.52528	8.50	0.000	134.1478	214.6055

Table 5

However, if we apply the same methodology as seen for hypothesis 3 (see 2.2.2) but for Israel, we retrieve the results presented in table 5. Indeed, if we calculate a GARCH-model for the Israeli market (represented by the 125 biggest Israeli companies included in the TA-125), we can confirm and strengthen the conclusion made by Yonamine (2013). Namely, Israeli investors do *not* react significantly on political conflicts on a daily basis on their market index. This leads us to interesting conclusions. Where Indian investors react significantly on political conflicts on a daily basis, Israeli investors do not. Thereby, we need to conclude that using political conflicts might be a significant variable in countries where violence is common, but not for all of them.

4 Conclusion

After the terrorist attack on the WTC towers on 9/11 in 2001, the New York Stock Exchange (NYSE) and the Nasdaq did not open for trading the following day. This, they did to prevent a stock market meltdown. Quoting Davis (2017) “anticipating market chaos, panic selling and a disastrous loss of value in the wake of the attacks, the NYSE and the Nasdaq remained closed until September 17, the longest shutdown since 1933. […] On the first day of NYSE trading after 9/11, the market fell 684 points, a 7.1% decline, setting a record for the biggest loss in exchange history for one trading day.” Even other researches proved that this had abundant aftermath effects on international markets (Brounen & Derwall, 2010; Charles & Darné, 2006).

Terrorist attacks and other political conflicts, thus, could have substantial effects on the stock markets. However, the question arises if this is still the case for a terrorist attack in an EU-country on a European index (such as the Eurostoxx50).

We could see in our research that there was no statistical difference in the returns of the Eurostoxx50 with those of the MSCI during the event of a terrorist attack. The reason for the insignificance could lay, however, in the diversifying nature of the index itself and that it levers away this specific risk.

Also, after the nuanced study of Yonamine (2013) on the effect of political conflicts in countries where a high level of violence is common, a new question arose. Is the common thinking, that political conflicts affect the stock markets, also true in other countries that enjoy a high level of violence? According to Yonamine (2013), Israeli investors already discount their prices on domestic shares. However, is this true for all dangerous countries? In this research, we took a closer look on India since it is rated 18 on places most risky to do business in according the Conflict and Political Violence Index for 2014 (Dhoot, 2014).

As our research has demonstrated, investors of the Indian BSE30 *do* react significantly on political conflicts, but the index recovers quickly. Therefore the abnormal returns are significant on a *daily* basis but insignificant on a *weekly* basis. Conflicts recorded in the GDELT-dataset, thus, *might* remain a significant variable in predicting the future (or analyzing the variance of) returns of stock markets, even for an index in a country that enjoys a high level of violence. In other words, it remains possible for some countries with high levels of violence to use material conflicts to predict future returns. It, therefore, might be possible to make a forecast based on these variables and make a profit on it.

However, we also recalculated the model of Yonamine (2013) with our methodology and on a daily basis. Now we confirm the insignificance of political conflicts on the TA-125 on a daily basis. Including political violence with the GDELT-dataset, thus, might be an important variable for *some* countries with a high level of violence, but not for all of them.

By coming to the previous conclusions, we came across some possible limitations and improvements. Firstly, while calculating the GARCH-models, only data for one year was used. However, this can be easily improved by extending the dataset. Secondly, for the variable ‘conflict count’ we only accumulated material conflicts. This can also be extended with verbal conflicts which might improve the significance of the model even further.

References

Ahmed, R. A., & Shabri, A. Bin. 2014. Daily crude oil price forecasting model using arima, generalized autoregressive conditional heteroscedastic and Support Vector Machines. ***American Journal of Applied Sciences***. https://doi.org/10.3844/ajassp.2014.425.432.

Alberg, D., Shalit, H., & Yosef, R. 2008. Estimating stock market volatility using asymmetric GARCH models. ***Applied Financial Economics***. https://doi.org/10.1080/09603100701604225.

Apergis, E., & Apergis, N. 2016. The 11/13 Paris terrorist attacks and stock prices: The case of the international defense industry. ***Finance Research Letters***. https://doi.org/10.1016/j.frl.2016.03.002.

Bishop, C. 2013. Mixture Density Networks. ***Journal of Chemical Information and Modeling***. https://doi.org/10.1017/CBO9781107415324.004.

Bollerslev, T. 1986. Generalized Autorregresive Conditional Hetersoskedasticity. ***Journal of Econometrics***. https://doi.org/10.1016/0304-4076(86)90063-1.

Brounen, D., & Derwall, J. 2010. The impact of terrorist attacks on international stock markets. ***European Financial Management***. https://doi.org/10.1111/j.1468-036X.2009.00502.x.

Brown, S. J., & Warner, J. B. 1985. Using daily stock returns. The case of event studies. ***Journal of Financial Economics***. https://doi.org/10.1016/0304-405X(85)90042-X.

Charles, A., & Darné, O. 2006. Large shocks and the September 11th terrorist attacks on international stock markets. ***Economic Modelling***. https://doi.org/10.1016/j.econmod.2006.03.008.

Chojnacki, S., Ickler, C., Spies, M., & Wiesel, J. 2012. Event Data on Armed Conflict and Security: New Perspectives, Old Challenges, and Some Solutions. ***International Interactions***. https://doi.org/10.1080/03050629.2012.696981.

Clemen, R. T. 1989. Combining forecasts: A review and annotated bibliography. ***International Journal of Forecasting***. https://doi.org/10.1016/0169-2070(89)90012-5.

Davis, M. 2017. How September 11 Affected The U.S. Stock Market. ***Investopedia***. https://www.investopedia.com/financial-edge/0911/how-september-11-affected-the-u.s.-stock-market.aspx.

Dhankar, R. S., & Chakraborty, M. 2007. Non-linearities and GARCH effects in the emerging stock markets of South Asia. ***Vikalpa***. https://doi.org/10.1177/0256090920070303.

Dhoot, V. 2014. India does better on global conflict, political violence index. ***The Economic Times***. https://economictimes.indiatimes.com/news/politics-and-nation/india-does-better-on-global-conflict-political-violence-index/articleshow/34775107.cms.

Dunis, C. L., & Huang, X. 2005. Forecasting and Trading Currency Volatility: An Application of Recurrent Neural Regression and Model Combination. ***Applied Quantitative Methods for Trading and Investment***. https://doi.org/10.1002/0470013265.ch4.

Dunis, C. L., Laws, J., & Evans, B. 2016. Modelling and trading the gasoline crack spread: A non-linear story. ***Derivatives and Hedge Funds***. https://doi.org/10.1057/9781137554178.

Dunis, C. L., Laws, J., & Karathanasopoulos, A. 2012. Modelling and trading the Greek stock market with hybrid ARMA-neural network models. ***Springer Optimization and Its Applications***. https://doi.org/10.1007/978-1-4614-3773-4_4.

Dunis, C. L., Laws, J., & Sermpinis, G. 2010. Modelling and trading the EUR/USD exchange rate at the ECB fixing. ***European Journal of Finance***. https://doi.org/10.1080/13518470903037771.

Eck, K. 2012. In data we trust? A comparison of UCDP GED and ACLED conflict events datasets. ***Cooperation and Conflict***. https://doi.org/10.1177/0010836711434463.

Elshendy, M., Fronzetti Colladon, A., Battistoni, E., & Gloor, P. A. 2018. Using four different online media sources to forecast the crude oil price. ***Journal of Information Science***. https://doi.org/10.1177/0165551517698298.

Engle, R. F. 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. ***Econometrica***. https://doi.org/10.2307/1912773.

Fulcher, J., Zhang, M., & Xu, S. 2006. Application of Higher-Order Neural Networks to Financial Time-Series Prediction. ***Artificial Neural Networks in Finance and Manufacturing***. https://doi.org/10.4018/978-1-59140-670-9.

Galla, D., & Burke, J. 2018. Predicting social unrest using GDELT. ***Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)***. https://doi.org/10.1007/978-3-319-96133-0_8.

Galtung, J., & Ruge, M. H. 1965. The structure of foreign news : the presentation of the Congo, Cuba and Cyprus crisis in four Norwegian Newspapers. ***Journal of Peace Reasearch***. https://doi.org/10.1159/0000102124.

Gilbert, E., & Karahalios, K. 2010. Widespread Worry and the Stock Market. ***ICWSM***. https://doi.org/10.1.1.220.9231.

Giles, C. L., & Maxwell, T. 1987. Learning, invariance, and generalization in high-order neural networks. ***Applied Optics***. https://doi.org/10.1364/AO.26.004972.

Hippert, H. S., Pedreira, C. E., & Souza, R. C. 2000. Combining neural networks and ARIMA models for hourly temperature forecast. ***Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks, 2000 (IJCNN 2000) , Vol. IV***. https://doi.org/10.1136/emermed-2013-203373.

Kamijo, K., & Tanigawa, T. 1990. Stock price pattern recognition-a recurrent neural network approach. ***1990 IJCNN International Joint Conference on Neural Networks***. https://doi.org/10.1109/IJCNN.1990.137572.

King, G., & Lowe, W. 2003. An Automated Information Extraction Tool for International Conflict Data with Performance as Good as Human Coders: A Rare Events Evaluation Design. ***International Organization***. https://doi.org/10.1017/S0020818303573064.

Kolaric, S., & Schiereck, D. 2016. Are stock markets efficient in the face of fear? Evidence from the terrorist attacks in Paris and Brussels. ***Finance Research Letters***. https://doi.org/10.1016/j.frl.2016.05.003.

Kwak, H., & An, J. 2014. A First Look at Global News Coverage of Disasters by Using the GDELT Dataset. ***International Conference on Social Informatics***, 300–308.

Leblang, D., & Mukherjee, B. 2005. Government partisanship, elections, and the stock market: Examining American and British stock returns, 1930-2000. ***American Journal of Political Science***. https://doi.org/10.1111/j.1540-5907.2005.00155.x.

Leetaru, K., & Schrodt, P. A. 2012. GDELT: Global Data on Events, Location and Tone. ***International Studies Association***.

Lindemann, A., Dunis, C. L., & Lisboa, P. 2005. Level estimation, classification and probability distribution architectures for trading the EUR/USD exchange rate. ***Neural Computing and Applications***. https://doi.org/10.1007/s00521-004-0462-8.

Makridakis, S. 1989. Why combining works? ***International Journal of Forecasting***. https://doi.org/10.1016/0169-2070(89)90017-4.

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., et al. 1982. The accuracy of extrapolation (time series) methods: Results of a forecasting competition. ***Journal of Forecasting***. https://doi.org/10.1002/for.3980010202.

Mohammadi, H., & Su, L. 2010. International evidence on crude oil price dynamics: Applications of ARIMA-GARCH models. ***Energy Economics***. https://doi.org/10.1016/j.eneco.2010.04.009.

Newbold, P., & Granger, C. W. J. 1974. Experience with Forecasting Univariate Time Series and the Combination of Forecasts. ***Journal of the Royal Statistical Society. Series A (General)***. https://doi.org/10.2307/2344546.

O'Brien, S. P. 2010. Crisis early warning and decision support: Contemporary approaches and thoughts on future research. ***International Studies Review***. https://doi.org/10.1111/j.1468-2486.2009.00914.x.

O'Loughlin, J., Witmer, F. D. W., Linke, A. M., & Thorwardson, N. 2010. Peering into the Fog of War: The Geography of the WikiLeaks Afghanistan War Logs, 2004-2009. ***Eurasian Geography and Economics***. https://doi.org/10.2747/1539-7216.51.4.472.

Palm, F. C., & Zellner, A. 1992. To combine or not to combine? issues of combining forecasts. ***Journal of Forecasting***. https://doi.org/10.1002/for.3980110806.

Rigobon, R., & Sack, B. 2005. The effects of war risk on US financial markets. ***Journal of Banking and Finance***. https://doi.org/10.1016/j.jbankfin.2004.06.040.

Schneider, G., & Troeger, V. E. 2006. War and the world economy: Stock market reactions to international conflicts. ***Journal of Conflict Resolution***. https://doi.org/10.1177/0022002706290430.

Schrodt, P. A. 2012. Precedents, Progress, and Prospects in Political Event Data. ***International Interactions***. https://doi.org/10.1080/03050629.2012.697430.

Tenti, P. 1996. Forecasting foreign exchange rates using recurrent neural networks. ***Applied Artificial Intelligence***. https://doi.org/10.1080/088395196118434.

Terui, N., & Van Dijk, H. K. 2002. Combined forecasts from linear and nonlinear time series models. ***International Journal of Forecasting***. https://doi.org/10.1016/S0169-2070(01)00120-0.

Tiňo, P., Schittenkopf, C., & Dorffner, G. 2001. Financial volatility trading using recurrent neural networks. ***IEEE Transactions on Neural Networks***. https://doi.org/10.1109/72.935096.

Tseng, F. M., Yu, H. C., & Tzeng, G. H. 2002. Combining neural network model with seasonal time series ARIMA model. ***Technological Forecasting and Social Change***. https://doi.org/10.1016/S0040-1625(00)00113-X.

Weidmann, N. B., & Ward, M. D. 2010. Predicting conflict in space and time. ***Journal of Conflict Resolution***. https://doi.org/10.1177/0022002710371669.

Winkler, R. L. 1989. Combining forecasts: A philosophical basis and some current issues. ***International Journal of Forecasting***. https://doi.org/10.1016/0169-2070(89)90018-6.

Wu, H. D. 2000. Systemic determinants of international news coverage: A comparison of 38 countries. ***Journal of Communication***. https://doi.org/10.1111/j.1460-2466.2000.tb02844.x.

Yermack, D. 1997. Good timing: CEO stock option awards and company news announcements. ***Journal of Finance***. https://doi.org/10.1111/j.1540-6261.1997.tb04809.x.

Yonamine, J. E. 2013. A nuanced study of political conflict using the Global Datasets of Events Location and Tone (GDELT) dataset. ***ProQuest Dissertations and Theses***.

Zammit-Mangion, A., Dewar, M., Kadirkamanathan, V., & Sanguinetti, G. 2012. Point process modelling of the Afghan War Diary. ***Proceedings of the National Academy of Sciences***. https://doi.org/10.1073/pnas.1203177109.

Zhang, G. P., & Qi, M. 2005. Neural network forecasting for seasonal and trend time series. ***European Journal of Operational Research***. https://doi.org/10.1016/j.ejor.2003.08.037.

Zhang, P. G. 2003. Time series forecasting using a hybrid ARIMA and neural network model. ***Neurocomputing***. https://doi.org/10.1016/S0925-2312(01)00702-0.

Zhang, X., Fuehres, H., & Gloor, P. A. 2011. Predicting Stock Market Indicators Through Twitter "I hope it is not as bad as I fear." ***Procedia - Social and Behavioral Sciences***. https://doi.org/10.1016/j.sbspro.2011.10.562.

Zussman, A., Zussman, N., & Nielsen, M. Ø. 2008. Asset market perspectives on the Israeli-Palestinian conflict. ***Economica***. https://doi.org/10.1111/j.1468-0335.2007.00607.x.

Press release

GDELT-dataset remains a noteworthy variable to predict the change in index price for countries that enjoy a high level of violence.

With the attempt of Ballings, M (2019) to predict social unrest and financial market collapse, the author proves that using GDELT-dataset as a variable might still significantly improve the quality of models that try to forecast changes in index prices in countries that enjoy a high level of violence. On first sight, he gives proof that is inconsistent with other authors that suggest that using political violence as a variable is insignificant in countries where violence is common. What makes this research different is that he looks at the differences in the index prices of the Indian BSE30 whereas the other looks at the Israeli TA100. Also, there are differences in the dataset itself. The research of Ballings, M (2019) looks at both daily and weekly differences, where other only looks on a weekly basis.

Also, he recalculated the effect of political conflicts on the TA-125. For this index, however, he comes to the same conclusion, namely that conflicts do not have a noteworthy effect. Thus, using the GDELT-dataset to predict changes in index prices for countries where conflicts are common might have a significant effect, but this is not always the case.

Another incongruity pointed out in the research of Ballings, M (2019) with the existing literature is the fact that he couldn’t find any significance in returns during a terrorist attack for European stocks. Here, the existing literature proves negative significance in returns of the market indexes after a terrorist attack. However, this could be due to the fact that he uses the index Eurostoxx50 as an index for European stocks. Here, the main explanation for the insignificance might lay in the nature of the index itself, and that is levers away from the specific risk of a possible terrorist attack.