

Who are the financial profiteers of war?

University of Applied Sciences Lucerne

Master of Science in Applied Information and Data Science

Time Series in Finance (TSA01)

Authors: Michèle Gerber and Damien Grossniklaus

Date of Submission: 22. December 2023

Contents

1. Context and Goal	3
2. Methods	3
3. Data	3
Visualize Ticker Data	5
Visualize Google Trends Data	6
Stationarity test and adjustment	7
Ticker Data	7
Google Trends Data	7
Create final dataframe	7
4. VAR & Causality Testing	7
5. Results & Discussion	7

1. Context and Goal

With recent outbreaks of war (e.g., in Israel or in the Ukraine), the question arises of how stock markets react to this and if there are sectors that profit from such crises. Specifically, this case study examines if several indices from specific economic sectors perform well in situations of crisis. The following six indices are included in the analysis:

- **XLK**: Technology Select Sector
- **XLF**: Financial Select Sector
- **XLV**: Health Care Select Sector
- **XLY**: Consumer Discretionary Select Sector
- **XLE**: Energy Select Sector
- **XLU**: Utilities Select Sector

Essentially our research question is the following: *Who are the financial profiteers of war?*. To answer this question, the performance of the sector-specific indices should be compared to an overall crisis measure. As an overall crisis measure the Google Trend data for the search word “war” is used. We postulate, that an increased search for this term points to a higher crisis. So in conclusion, we want to examine if there are positive correlations between our crisis measure and stock market profits in the different indices.

2. Methods

Our research question is of an exploratory nature. Therefore we don’t have a hypothesis going in, as to which indices will perform better under situations of crisis. The method we will use is vector autoregression (VAR). This allows us to measure the influence of our crisis measure on the performance of the different indices. For each index we will calculate a separate model. Furthermore, to address potential spurious correlations a Granger causality test will be conducted for every ticker. However, before a VAR and causality test can be carried out the data in form of different time series has to be imported, cleaned and pre-processed. Concretely, potential seasonal effects and trends in time series have to be identified and adjusted for to achieve stationarity, which characterizes a stabilized mean and variance for the time series over time and will ultimately allow more reliable and accurate modeling. To test the stationarity of a time series the “*Augmented Dickey-Fuller Test*” will be carried out. If a time series is non-stationary, lagged differences will be calculated. Lastly, to deal with right skewness and achieve normal distributions in the data, the time series will first be logarithmized before applying the difference.

3. Data

Financial Data

The financial data used for the paper at hand comes from Yahoo Finance (2023)(**SOURCE**). The data can be retrieved using the `getSymbols.yahoo()` function from the `quantmod` package. The data represents the weekly adjusted closing prices for the tickers introduced in chapter 1 for the time frame from December 2021 until December 2023. Mentioned time frame was chosen as it captures the start of the Russian invasion of Ukraine as well as the start of the Israel-Hamas war. Additionally, the time frame and periodicity is considered of good balance between having enough data points as well as keeping the data set manageable and focused. The decision to use the adjusted closing prices instead of the closing price can be justified that adjusted prices are adapted to address any splits, dividends or capital gain distributions. Lastly, the weekly periodicity of the ticker data was chosen as the Google Trends data comes in weekly measurements.

Google Trends Data

To import the Google Trends data the package `gtrendsR` is used. The package allows to query the data of the interest over time for a keyword. In the case of this study the keyword *war* in web searches was chosen,

as it is considered to capture all conflicts alike. When it comes to the geographical origin of the searches the whole world was considered as the study aims to give a generalized and non location specific view. Returned data from Google Trends is normalized and reflects the search volume for a keyword on a scale from 0-100, where 100 means very high interest and 0 no interest. (**SOURCE:** <https://newsinitiative.withgoogle.com/en-gb/resources/trainings/fundamentals/google-trends-understanding-the-data/#:~:text=The%20percentages%20are%20based%20on,for%20the%2030%20days%20prior.>)

<https://cran.r-project.org/web/packages/gtrendsR/gtrendsR.pdf>

Inflation Data

In order to handle potential confounding variables that could cause spurious associations due to influence on the dependent and independent variable, it is important to also include a control variable that represents the general economic state. Therefore, the worldwide inflation rate is included as well, which was retrieved from the OECD. To allow a simple aggregation the mean inflation rate per month for every available country in the data set was calculated.

<https://data.oecd.org/price/inflation-cpi.htm>

Import Time Series Data

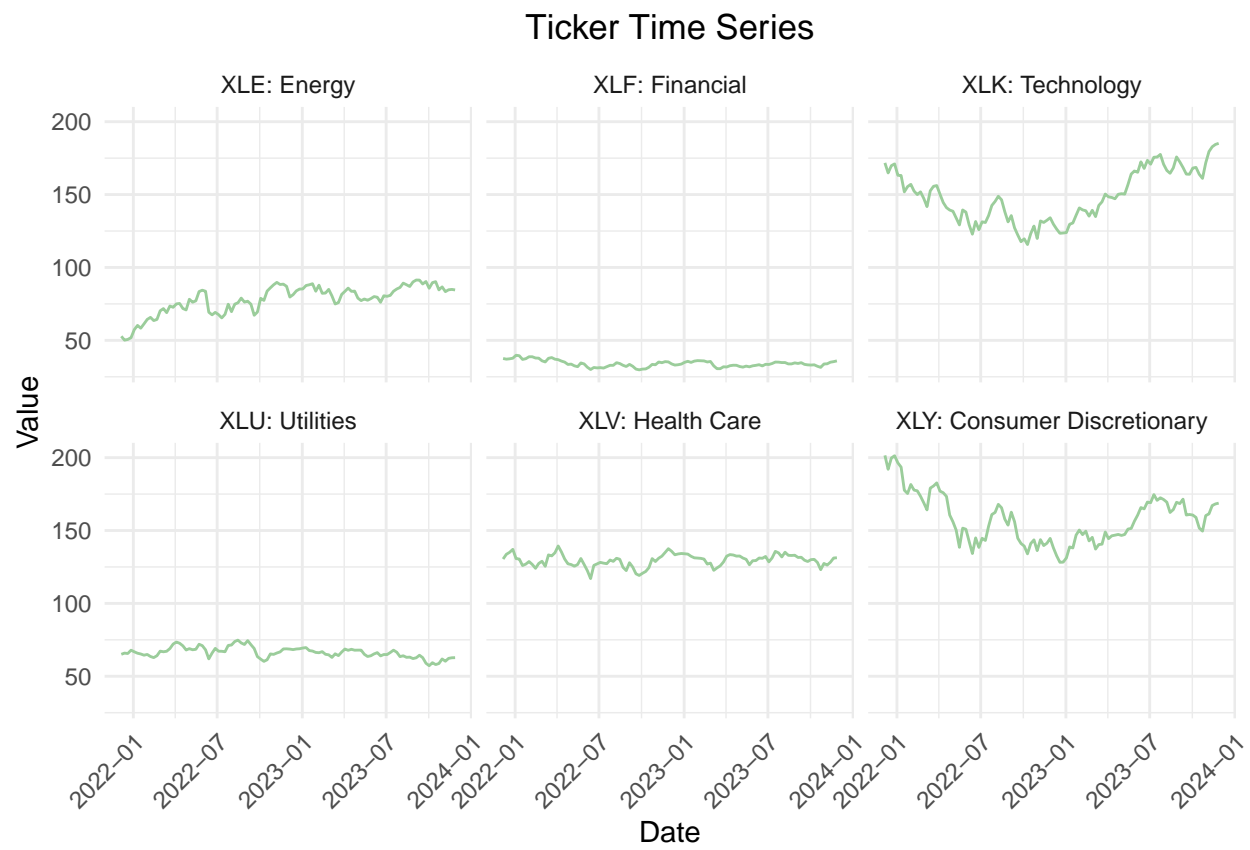
TODO: - Justify why Adjusted values and not close values - Timeframe: December 2021 until December 2023

Google Trends delivers data always from Sunday weekly. This has to be put to the weekly data from the financial tickers, which always starts at Monday.

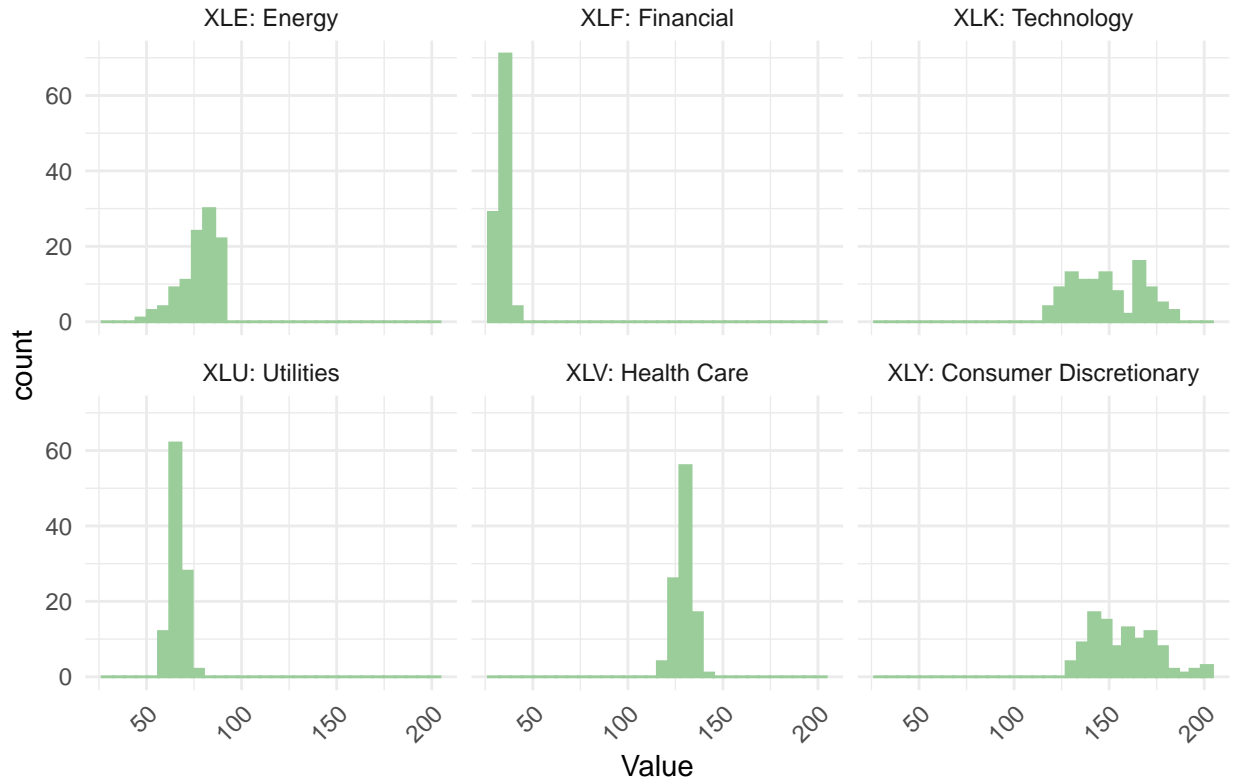
Inflation Data

Aggregation and preprocessing

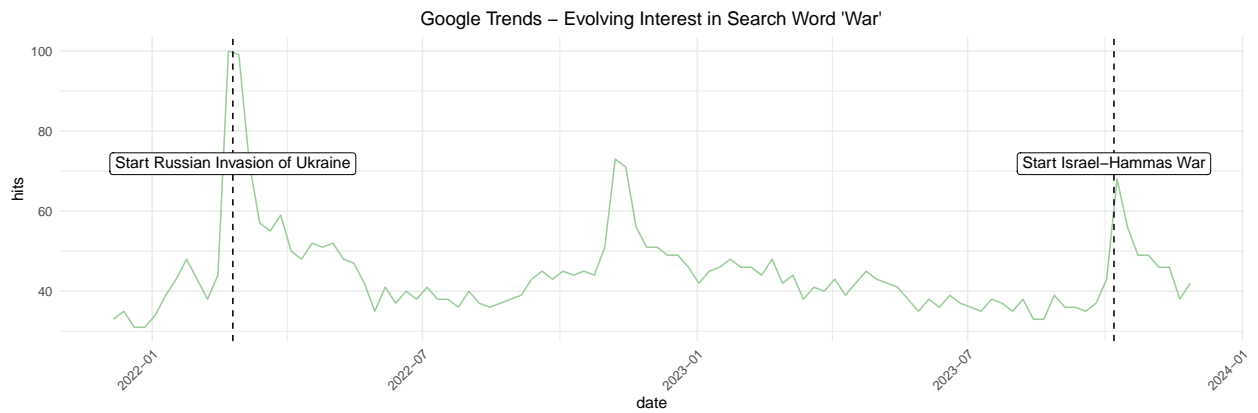
Visualize Ticker Data

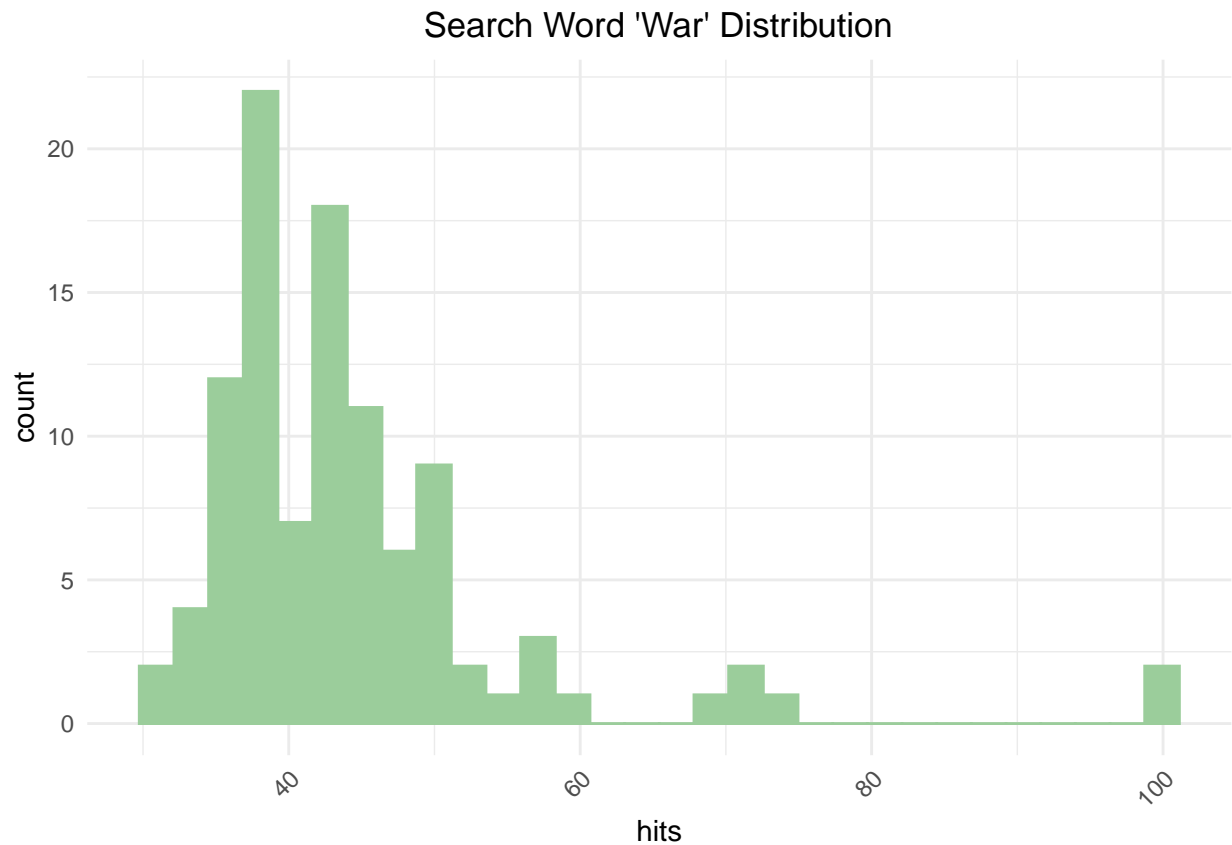


Ticker Distributions



Visualize Google Trends Data





Stationarity test and adjustment

Ticker Data Non-stationary are XLK, XLF, XLY, XLE. Make them stationary. Only log the others.

After the diff-log-transformation all the time series are stationary.

Google Trends Data Not stationary and distribution has long tail on right → Log Diff. Although p value is slightly below 0.05.

Check that differenced time series have same length.

Create final dataframe

Add inflation data

4. VAR & Causality Testing

TODO: - Check lag.max what to apply - Control variable?

5. Results & Discussion