

FIN3210 Week 4 Assignment Report

Ma Kexuan 120090651

Abstract

This report constructs two indexes for different purposes, and presents the indexes into several graphs and interpretations.

Data Preprocessing

The preprocessing procedures and some interpretations of the code are described in each code blocks in the appendix, please check.

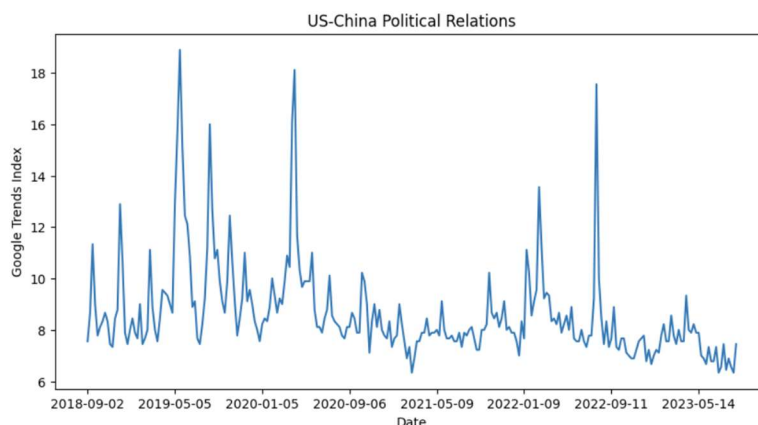
Questions

1) Using Google Trends (<https://trends.google.com/trends/?geo=US>), construct a weekly index to capture political relations between U.S. and China from the US perspective, draw the variable in a graph, and discuss its time-series variation.

For this question, I selected 10 distinct words to describe the relationship between U.S. and China from the US perspective, which are Tariffs, South China Sea, Huawei, Trade War, Made in China, Tibet, Hong Kong, Taiwan, U.S.-China, 5G. Most of which are relevant to name of “controversial” area of China, and some of them are relevant to the new technology developed by China, which is somewhat more advanced than the same kind of techs in the US, and the rest are relevant to the trade between these two countries. It can be necessary to analyze the political relation using index relevant to these keywords. Below is a correlation map of these words, one thing to be mentioned is that the U.S-China keyword is dropped since there’s not much information retracted by the Google Trend. By the correlation coefficient, we can see that these words are not so correlated, which means they can describe the relationship from quite different perspective.

	Tariffs	South China Sea	Huawei	Trade War	Made in China	Tibet	Hong Kong	Taiwan	5G
Tariffs	1.000000	0.064757	0.508916	0.887880	-0.133332	0.059853	0.293850	-0.151932	-0.372672
South China Sea	0.064757	1.000000	0.031378	0.046260	0.167426	0.068504	-0.009325	0.073614	-0.024228
Huawei	0.508916	0.031378	1.000000	0.528077	-0.212102	0.045272	0.220016	-0.214053	-0.350924
Trade War	0.887880	0.046260	0.528077	1.000000	-0.156281	0.088927	0.408141	-0.150847	-0.433250
Made in China	-0.133332	0.167426	-0.212102	-0.156281	1.000000	0.095936	-0.071522	0.113407	0.405090
Tibet	0.059853	0.068504	0.045272	0.088927	0.095936	1.000000	-0.080955	0.191257	-0.056895
Hong Kong	0.293850	-0.009325	0.220016	0.408141	-0.071522	-0.080955	1.000000	-0.097992	-0.289373
Taiwan	-0.151932	0.073614	-0.214053	-0.150847	0.113407	0.191257	-0.097992	1.000000	0.196832
5G	-0.372672	-0.024228	-0.350924	-0.433250	0.405090	-0.056895	-0.289373	0.196832	1.000000

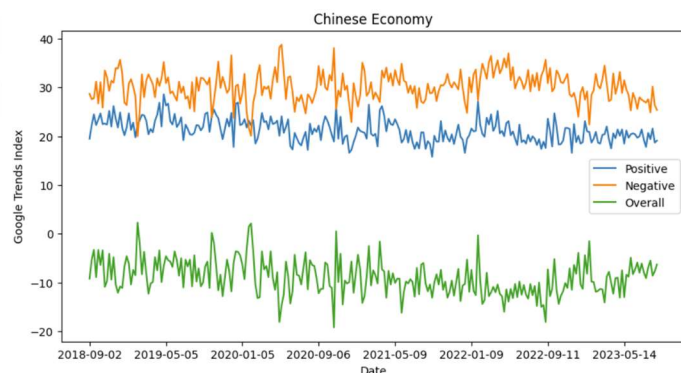
The time series of the overall constructed index is shown below. We can see that the political relation between these two countries are at the pinnacle in 2019, 2020 and 2022. During 2019 and 2020, there were trade wars and the burst of the pandemic, which made the index relatively high. In the late 2022, the pinnacle happened because of the Taiwan issue.



2) Using Baidu Index (<http://index.baidu.com/>) or Google Trends (<https://trends.google.com/trends/?geo=US>), construct an index to capture investor sentiment in the Chinese market, draw the variable in a graph, and discuss its time-series variation.

For this question, I provide 10 positive words and 10 negative words to describe the financial market sentiment in China. The positives are boom, buy, credit, gain, profit, reward, surge, rise, boost, win. The negatives are bankrupt, capital, decline, default, fall, inflation, liability, loss, recession, short. All of them have quite evident pos or neg sentiments. Then I use these words to aggregate a word-level search sentiment index. First, I perform an OLS to discover the linear relationship between the return and the sentiment index from last week. The result is shown below. Unfortunately, the p-values are so large that we are inconclusive about this relationship. This maybe can be interpreted as the sentiments are with some delay, since it is constructed based on weekly trends, and within the week, the sentiments have been already digested by the market, leading to no predictive power to the next week. Another possibility is that the words are not enough to cover the whole sentiment, which maybe improved by adding more words to construct the index.

Dep. Variable:	week_ret	R-squared:	0.002			
Model:	OLS	Adj. R-squared:	-0.002			
Method:	Least Squares	F-statistic:	0.5385			
Date:	Sat, 14 Oct 2023	Prob (F-statistic):	0.464			
Time:	18:18:13	Log-Likelihood:	601.44			
No. Observations:	254	AIC:	-1199.			
Df Residuals:	252	BIC:	-1192.			
Df Model:	1					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	0.0035	0.004	0.862	0.389	-0.005	0.012
total_index	0.0003	0.000	0.734	0.464	-0.001	0.001
Omnibus:	4.268	Durbin-Watson:	2.123			
Prob(Omnibus):	0.118	Jarque-Bera (JB):	4.990			
Skew:	-0.118	Prob(JB):	0.0825			
Kurtosis:	3.645	Cond. No.	28.0			



The time series is also attached above. We can see that the Chinese Economy in the recent five years was quite fatigue. This was corroborated by the Shanghai main market index, which remains at 3000 points over the years. There're some of the points that the positive sentiment transcends the negative one. We can observe that during the whole year 2019, there was a quite bull trend in the A share market, so the positive market somewhat went beyond the negative one. Nevertheless, after the COVID-19 pandemic occurred, since the manufacturing process are retarded, the economy remained gloomy, which reflects the comparative numerical value of the positive sentiment and the negative sentiment of the market.