More Trade, More Support: Evidence from Attitudes towards China in UN General Debates

Shuyi Yu

1 Introduction

China has surged to a prominent economic power in recent decades. Its trade has expanded at a breakneck pace. In 1995, the value of China's imports and exports of goods totaled $280.9 billion or 3 percent of global trade. By 2017, its total trade in goods had jumped to $4.1 trillion or 12.4 percent of global trade, ranking first and followed by the US (Comtrade, 2010). There is prevalent skepticism on whether China's economic links to the world have been translated into political influence globally, especially in East Asia, Africa or Latin America (Kang, 2007; Alden, 2005; Flores-Macías and Kreps, 2013; Chin and Helleiner, 2008). There is no scholarly consensus on how much trading with China influences a country's attitude toward this raising power.

This study will use text analysis methods and cross-national regressions to test the effect of trading with China on the China image. Use the UN speeches of 193 countries from 1993 to 2015 and Text as Data methods, this paper provides a new measure for the national image of China. Coupled with trade statistics from World Bank, I find empirical evidence that indicates an optimistic effect of trading with China – growing trade would encourage states worldwide to adopt more favorable attitudes towards China and welcome China to be more active in global order and international sustainable growth.

2 Literature Review

There has long been debates about the connections between commercial relations and foreign policy. The seminal work by Hirschman (1980) hypothesizes that trade dependence between states produce foreign policy convergence. Both sides would not risk damaging trade relations by holding conflicting positions on foreign affairs. Rosecrance (1986) extends Hirschman's argument on trade volume by theorizing that the more states trade, the more incentives they would have to maintain smooth trade relations. Keohane and Nye (1977) further elaborates on the "asymmetrical interdependence as a source of power". The states that are more dependent on bilateral trade would be more vulnerable and more likely to compromise their political standpoints. However, the assertion that foreign policy convergence follows growing commercial ties has not gone without criticism. Ross (2006) argues that states concede to political and military hegemony but not economic power alone, thus trade volume itself is not sufficiently influential. In addition, states dependent on trade could also leverage their political or intellectual power to offset their economic disadvantage (Wagner, 1988; Holsti, 1978), thus I would be likely to observe little or even adverse foreign policy consequences following increasingly robust or asymmetric trade.

The empirical findings on the foreign policy consequences of trade with China are quite mixed. There are some qualitative research done with inconclusive results. Kirshner (2008) suggests that increasing trade would improve the salience of the constituencies who gain from trade, and they would advocate for more foreign policy coordination. Medeiros et al. (2008) backs this view by detailing how business interests in Japan supports closer cooperation with China on non-trade issues. However, Ross (2006) reaches different conclusion with the evidence from East Asia. Small players like Japan, Korea and Taiwan are nervous about China's military rise due to historical legacies. They are more likely to balance against China on foreign policies than to accommodate it, despite their growing trade relationship with China. Furthermore, the two only cross-national studies using regression methods also give conflicting results. By measuring foreign policy convergence using states' reactions to the 2008 Taiwan UN referendum and to China's 2008 crackdown on protest in Tibet, Kastner (2016) shows no systematic relationship between states' inclination to side with China and their economic integration with China, a finding consistent with Ross's argument. By measuring foreign policy convergence using United Nations General Assembly (UNGA) votes on country-specific human rights resolutions, Flores-Macías and Kreps (2013) offers strong evidence that Africa and Latin America states that trade with China are more likely to side with it on human rights issues, which corroborates Kirshner's story. So the question of whether the increasing trade volume with China makes other countries converge with it in international politics arena remains unanswered.

3 Theory and Hypotheses

Based on Keohane and Nye (1977), there are two possible theoretical outcomes of growing trade volume between two states. One is following the theoretical argument on sensitivity of interdependence. As both states are more sensitive to each other's moves on trade policies, they would be more likely to cooperate on other foreign policies to maintain a good trade relationship. The other is following the theoretical argument on vulnerability of interdependence. Both states are also more vulnerable to each other's moves on trade policies, thus they would balance the vulnerability by being confrontational on other foreign policies, e.g. if a small state gains bargaining power on Taiwan issue, then it is less vulnerable to the trade threats by China.

Accordingly, I could raise the following three hypotheses: if sensitivity argument works, I should observe increasing trade volume has a positive effect on foreign policy towards China; if vulnerability argument works, I should observe a negative impact; if both work, I should observe little or no impact. Thus it comes to an empirical question that, as China has become a dominant trading partner of many countries, whether these countries become more favorable towards China on foreign policies or not.

We would make use of a new data set on UN General Debate and employ text analysis methods to analyze it, such that I could present new evidence to this empirical question. The traditional measurement of foreign policy convergence is the UNGA voting. But there are serious drawbacks with this measure. First, it includes procedural votes that are meaningless for analysis. Second, it does not show the direction of convergence regarding whether other countries are moving towards China or the opposite, which China is switching its own position. My methods could more directly and better capture other countries' policy attitudes towards China on meaningful issues, thus eliminating concerns on these drawbacks.

4 Research Design

4.1 Data

The first source of data is the UN General Debate corpus, which is introduced by Baruro and colleagues (Baturo et al., 2017). Each year, member states of the United Nations General Assembly (UNGA) deliver speeches at the General Debate session on major issues in world politics. There is one General Debate session in each year, and each country will give one speech at each session. The UNGA speech data is a great source to measure the political attitude toward China for the following reasons. First, the UNGA speeches provide a full set of longitudinal information on countries' foreign policy and holistically reflect political attitudes toward other countries. This corpus includes over 7500 UNGA General Debate speeches for 193 countries from 1970 to 2015, but given the time coverage of my bilateral trade data, I will use speeches from 1993-2015 to construct my measurement for country's foreign policy towards China. Second, all the countries are using the same podium and same language to express their foreign policies at UN, so it makes the cross-country comparison possible. Finally, these speeches provide invaluable information on governments' perspectives and preferences on a wide range of policy areas, but they have been largely overlooked in the study of international relations. For instance, previous studies such as Flores-Macías and Kreps (2013) make use of country-specific human rights votes in the UNGA as the measurement for foreign policy convergence, but this has the limitation of restricting to a narrow policy dimension, whereas using text data from country's original speeches may allow us to better capture a country's general political standpoint and sentiment towards China on various important global issues.

My second data source is bilateral trade statistics reported by the World Bank. I collect data on each country's exports to and imports from China during 1992-2014. The main explanatory variables of interest are ***total trade volume*** with China and the ***balance of trade***, both measured in thousand US dollars. These two measures provide a relatively comprehensive bilateral trade relationship between countries. On the one hand, the total trade volume reflects the overall quantity of bilateral economic relationship. On the other hand, balance of trade – difference between exports and imports in specific year – demonstrates the quality and sustainability of countries economic relationship with China.

4.2 Empirical Strategy 1 – Regression Analysis

I use Text as Data method and construct the dependent variable, country's ***attitude towards China***, from the UN General Debate speeches by sentiment analysis. It is conceivable that these speeches focus on country's perception on a wide range of important global affairs, among which many would be irrelevant to China per se. Therefore, I first sort out the documents relevant to China by identifying keywords in context. I managed to extract

779 speeches which directly mentioned the word "China" between 1993-2015, and these documents constitute my main sample. In Appendix Table A1, I listed the top 25 countries which mentioned China most frequently in their speeches.

Then I adopt a dictionary-based method to capture the sentiment of each document. Applying Hu and Liu (2004)'s dictionary of positive and negative features, I create the document-feature matrix and obtain the number of positive and negative words in each speech. The sentiment score is calculated using the difference between the number of positive and negative words, divided by the total number of sentiment features in each document. Figure 1 below shows the distribution of sentiment scores for all speeches related to China between 1993 and 2015, which varies from -0.583 to 0.889, and is approximately normally distributed around the mean of 0.262.

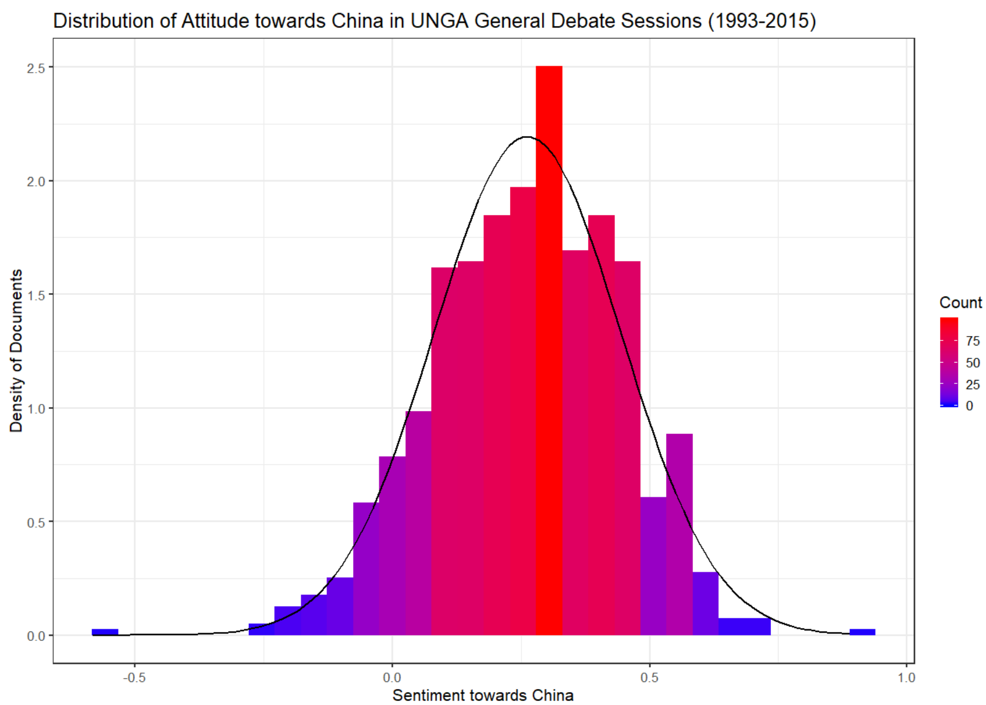

Figure 1: **The Distribution of Sentiment in Country's UNGA Speech (1993-2015)**
The X-axis is the sentiment towards China (higher value means more positive attitude) and Y-axis is the density of documents, with the black curve showing normal density. The grey scale of each bar represents the count of documents.

To test my hypotheses on the relationship between China's trade policy and partner states' sentiment towards China, I merged the UN General Debate speeches with my bilateral trade data to conduct regression analysis. The outcome variable is ***Sentiment score*** for each country in each year, based on the sentiment analysis of their speeches. I construct two main explanatory variables: (1) ***Total trade volume with China***, measured by the sum of exports and imports and taken *log* form to satisfy the normality assumption. (2) ***Balance of trade*** (or trade surplus), measured by subtracting a country's imports from China from its exports to China. The summary statistics of variables are presented in Table 1 below.

Table 1: Summary statistics of variables for regression analysis

	N	Mean	SD	Min	Max
Country	150				
General Debate speech (text)	779				
Year	779			1993	2015
Sentiment score	779	0.262	0.182	-0.583	0.889
Exports to China (*log*)	674	9.372	4.758	-4.711	19.025
Imports from China (*log*)	771	11.074	3.280	0.954	19.800
Total volume of trade (*log*)	673	12.120	3.090	3.533	20.138
Balance of trade (*log*)	673	8.361	0.326	0	8.517

Notes: I have a total of 779 speeches from 150 countries in my sample. Each year, every country delivers one speech at the General Debate session. Speeches which did not include the word "China" are dropped from my sample. All trade-related variables are measured in thousand US$ and with *log* transformation in the analysis.

In order to measure the relationship of a country's bilateral trade with China and its political attitude toward China, I formally specify model as:

$$Sentiment_{i,t} = \alpha_i + \delta_{t-1} + \beta_1 Trade\ with\ China_{i,t-1} + \epsilon_{i,t-1}.$$

The main variable of interest, *Trade with China*, is measured by the log transformation of total trade volume and trade balance. Given the concern for potential reverse causality, which is that a speech signaling positive attitude towards China might also boost trade relations in later years, I estimate the effect of the trade variables in year $t-1$ on the sentiment toward China in year t. Furthermore, the inclusion of country fixed effects, α_i controls for unobserved time-invariant and country-specific factors that may affect countries' attitudes towards China. And the year fixed effects, δ, control for the time trend of the favorability toward China. x

I expect the key parameter of interest (β_1) to be positive if the "sensitivity of interdependence" theory is true; negative if the "vulnerability of interdependence" theory is true; or insignificant if the effect of trade interdependence is null or has been cancelled out by mechanisms in opposite directions.

4.3 Empirical Strategy 2 – Topic Models

In addition to how trade relations influence countries' foreign position towards China, I also want to explore what are the main themes countries care about China. In specific, whether countries with positive and negative sentiment towards China would discuss China differently? Since we have a rich corpus of UN speeches from countries all over the world, diving into the text data and reading between the lines can provide a qualitative and even explanatory information why exactly trade with China can boost or block the favorability toward China.

First, I labeled the UN speeches dichotomously, where speeches with a sentiment score above zero are classified as "positive", and speeches with sentiment score below or equal to zero are classified as "negative". I then apply a Latent Dirichlet Allocation (LDA) model on both the positive set and the negative set of UN speeches. Before running the topic model, I follow the standard in the text analysis literature and perform the following pre–processing steps on the raw speeches: (i) delete all punctuation; (ii) remove capitalization; (iii) drop stop words and numbers; and (iv) eliminate words with low frequency. I also identify the optimal number of topics with respect to the trade–off that balances the granularity and the generality of the topics. Here, I fit a model for 20 topics.

The LDA model returns prevalence of the 20 topics for each individual speech. I then extract the most likely topic for each speech and rank the topics based on the number of speeches for which they show up as the most likely topic. Using this approach, I am able to identify the most widely discussed topics for "positive" and "negative" class respectively, and make comparison what are the most distinctive topics between the positive speeches and negative speeches toward China.

5 Results

In this section, I first present results of my regression analysis on the relationship between trade and country's sentiment towards China, followed by an exploration of the major themes countries tend to portray China positively and negatively.

5.1 Does trade increase country's support to China in the UN?

The main results are shown in Table 2 below. In column (1), I find strong evidence that increasing bilateral trade with China does help boost support for China in the UN. Specifically, one percent increase in total trade volume with China will increase the country's sentiment score towards China by 0.038 in the following year, which is around 14.5% given the average sentiment score of 0.262. I find similar results for the balance of trade in column (2). The coefficient is positively significant at the 5% level. The magnitude of effect is even greater than that of total trade volume: one percent increase in trade surplus will boost the sentiment score by 0.072, or 27.5%.

These findings lend support to the sensitivity of interdependence theory, which predicts trade relations to have a positive impact on country's foreign policy alignment. Indeed, I discover that more intimate economic ties will advance the positive image of a country's trade partner. Furthermore, not only total trade volume but also how much countries gain in terms of bilateral trade surplus matters for their attitudes towards China in the international arena. This finding on the balance of trade extends the earlier debate from the magnitude of interconnectedness between states to the symmetry of this relationship. Practically, it also helps account for the United States's continued dissatisfaction with China, as well as the escalation of the trade war recently. Finally, my findings also suggest that bilateral trade can be an effective policy instrument for China to "buy" support from other countries on its foreign policy position.

Table 2: Does bilateral trade affect countries' sentiment towards China?

	Dependent variable: Sentiment Score	
	(1)	(2)
Trade volume with China$_{(t-1)}$	0.038*	
	[0.022]	
Trade balance with China$_{(t-1)}$		0.072**
		[0.030]
Country Fixed Effects	✓	✓
Year Fixed Effects	✓	✓
Observations	673	673
R^2	0.510	0.519

Notes: The explanatory variables are measured in *log* form. Standard errors clustered at the country-level are reported in brackets. Significance levels: ** p<0.05, * p<0.1.

5.2 What topics are widely discussed about China in UN speeches?

After showing that China's leading role in international trade does matter for its ability to attract policy support among its trade partners, I extend my analysis to explore the more general issue topics where countries tend to comment on China. I present the top five topics in the "positive" class of UN speeches in Table 3, and those from the "negative" class of speeches in Table 4. After extracting the top ten keywords for each topic, I labeled the topics manually. I find some interesting patterns from these topics, and evidence that countries with positive sentiment towards China indeed tend to discuss issues different from countries with more negative sentiment towards China.

To begin with, there are a few language expressions which are very commonly used in UN speeches, such as "country", "international", or "organization" in positive Topic 1 of Table 3, and therefore it is not surprising that they appear the most frequently. Nevertheless, I consider these UN common expressions having fewer meaningful implications than themes which contain more substantive contents. Besides these UN common expressions, I find that speeches with more positive sentiment towards China are more likely to discuss about the global order and cooperation, Pacific islands, and sustainability and development. These topics suggest that China's international role in pursuing cooperation and boosting global economic development has been largely acknowledged and praised by the international community. One tricky finding is that, positive Topic 4 on small open economies is frequently discussed by both positive and negative speeches, which may indicate either China's influence on the development of these economies are still debated, or that these countries tend to have ambiguous perception of China's global image.

Switching to topics on which China has often been perceived more negatively (Table 4 below), I find that the United States (negative Topic 2), international relations (negative Topic 4) , and security issues in Africa (negative Topic 5) appear among the top 5. These

Table 3: Keywords for Top 5 topics among speeches with positive sentiment on China

+ Topic 1 UN Expressions	+ Topic 2 Global Order	+ Topic 3 Small Open Economies	+ Topic 4 Pacific Islands	+ Topic 5 Sustainable Growth
republic	international	states	island	development
country	social	small	pacific	sustainable
community	central	caribbean	islands	agenda
international	peace	saint	solomon	goals
organization	country	economic	states	international
must	us	developing	small	national
efforts	order	community	nations	support
particular	organization	must	united	security
situation	also	island	sustainable	global
like	cooperation	trade	forum	millennium

Table 4: Keywords for Top 5 topics among speeches with negative sentiment on China

− Topic 1 Common Terms	− Topic 2 U.S.	− Topic 3 Small Open Economies	− Topic 4 International Relations	− Topic 5 Africa Security
one	world	states	international	africa
us	people	small	security	african
can	peoples	caribbean	european	peace
now	america	saint	cooperation	security
time	countries	economic	economic	liberia
even	united	developing	important	community
also	president	community	relations	support
years	american	must	states	continent
first	states	island	regional	delegation
many	nations	trade	problems	conflict

results also make much sense intuitively. Firstly, given the divergence in core interests and major foreign policy positions between the U.S. and China, countries with closer alliance to the U.S. are expected to be loosely connected to China or hold more divergent views from China's foreign policy position, and thus convey more negative sentiment towards China in their speeches. Secondly, "China's economic assistance to African countries can be detrimental to the public welfare and security of these states by empowering authoritarian rulers and state repression" has been a mainstream criticism from many liberal democracies. Thirdly, the rising power and influence of China on the global platform has been perceived by the U.S. and some European countries as a new threat to the existing world order, which suggests that speeches made by these countries who view China as a strong competitor, or speeches related to the stability of current geopolitical structure and international relations may express more negative attitudes toward China.

In general, my exploration through the application of topic models provides some preliminary evidence that China's global image based on the perception of other countries tends to be a mixture of both positive role on economic development, and negative role on impeding democratization and challenging the power of existing leaders in international system.

References

Alden, C. (2005). China in africa. *Survival 47*(3), 147–164.

Baturo, A., N. Dasandi, and S. J. Mikhaylov (2017). Understanding state preferences with text as data: Introducing the un general debate corpus. *Research & Politics 4*(2), 2053168017712821.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent dirichlet allocation. *Journal of Machine Learning 3*, 993–1022.

Chin, G. and E. Helleiner (2008). China as a creditor: a rising financial power? *Journal of International Affairs*, 87–102.

Comtrade, U. (2010). United nations commodity trade statistics database. *URL: http://comtrade. un. org*.

Flores-Macías, G. A. and S. E. Kreps (2013). The foreign policy consequences of trade: China's commercial relations with africa and latin america, 1992–2006. *The Journal of Politics 75*(2), 357–371.

Hirschman, A. O. (1980). *National power and the structure of foreign trade*, Volume 105. Univ of California Press.

Holsti, K. J. (1978). A new international politics? diplomacy in complex interdependence. *International Organization 32*(2), 513–530.

Hu, M. and B. Liu (2004). Mining opinion features in customer reviews. In *AAAI*, Volume 4, pp. 755–760.

Kang, D. C.-o. (2007). *China rising: peace, power, and order in East Asia.* Columbia University Press.

Kastner, S. L. (2016). Buying influence? assessing the political effects of china's international trade. *Journal of Conflict Resolution 60*(6), 980–1007.

Keohane, R. O. and J. S. Nye (1977). *Power and interdependence.* Boston.

Kirshner, J. (2008). The consequences of china's economic rise for sino-us relations: Rivalry, political conflict, and (not) war. *China's Ascent: Power, Security, and the Future of*

International Politics, 238–259.

Medeiros, E. S., K. Crane, E. Heginbotham, and N. D. Levin (2008). *Pacific Currents: The Responses of US Allies and Security Partners in East Asia to China's Rise*, Volume 736. Rand Corporation.

Rosecrance, R. N. (1986). *The rise of the trading state: Commerce and conquest in the modern world*. New York: Basic Books.

Ross, R. S. (2006). Balance of power politics and the rise of china: Accommodation and balancing in east asia. *Security Studies 15*(3), 355–395.

Wagner, R. H. (1988). Economic interdependence, bargaining power, and political influence. *International Organization 42*(3), 461–483.

Appendix

Table A1: Top 25 countries which mentioned China in UNGA speech most frequently (excluding China)

Country	Number of speeches
Solomon Islands	22
Burkina Faso	19
Dominica	19
Gambia, The	17
Nicaragua	17
St. Kitts and Nevis	17
Papua New Guinea	16
Paraguay	16
St. Lucia	16
Palau	15
St. Vincent and the Grenadines	15
Fiji	14
Grenada	14
Vietnam	14
Costa Rica	13
Malawi	13
Guatemala	12
Senegal	12
Belize	11
El Salvador	11
Liberia	11
Philippines	11
Sao Tome and Principe	11
Chad	10
Dominican Republic	10

Notes: Since I use data of UNGA annual speech from 1993 to 2015, the maximum number of speeches a country could possibly mention China is 23.

```
## 1 Load packages

library(quanteda)
library(dplyr)
library(stringi)
library(stringr)
library(ggplot2)
library(readstata13)
library(miceadds)

## 2 Data preparation
rm(list = ls())

# Load dataset of UN General Debates: 1970-2015 (session 25-70)
UNGA <-
get(load("/Users/ShuFu/Downloads/TextasDataProject/Data/UNGA_debate.Rdata"))

rm(undata)

# Extract sample from the period 1992-2015
UNGA_samp <- UNGA[which(UNGA$year>1991), ]

# Inspect distribution by year and country (relatively balanced)
table(UNGA_samp$year)
table(UNGA_samp$country)


# Load bilateral trade dataset
Trade <- read.dta13("Trade_with_China.dta")

# Lag trade by one year
UNGA_samp$year <- UNGA_samp$year-1  # This year's speech in response to
last year's trade.

# Merge two datasets by ISO country code and year
main_data <- merge(UNGA_samp, Trade, by=c("country","year"))

# Label document name
main_data$docname <- paste(main_data$country, main_data$year, sep = "_")

# Create corpus
corpus <- corpus(main_data)


## 3 Identify documents related to China

# Technique: Key Words in Context
China <- kwic(corpus, pattern = "China", valuetype = "fixed", window = 5)

# Get the document index
index <- unique(as.numeric(str_extract_all(China$docname, "[0-9]+")))
```

```
# Extract the corresponding documents
China_doc <- main_data[index, ]

# Which countries tend to mention China in UNGA annual speech more often?
China_freq <- China_doc %>% group_by(PartnerName) %>% tally() %>%
arrange(-n)
head(China_freq, 20)

# Developing countries mention China more often.


## 4 Sentiment Analysis (country's attitude towards China)
## Technique: Dictionary-based method

# Load dictionaries (Hu & Liu, 2014)
positive <- readLines("positive-words.txt")
negative <- readLines("negative-words.txt")

# Create DFM using dictionaries
dfm_p <- dfm(China_doc$text, select = positive) %>% convert("matrix")
dim(dfm_p)

dfm_n <- dfm(China_doc$text, select = negative) %>% convert("matrix")
dim(dfm_n)

# Compute the total number of positive and negative words in each speech
words_p <- rowSums(dfm_p)
words_n <- rowSums(dfm_n)

# Sentiment score: # of positive words minus negative words
sentiment <- (words_p - words_n)/(words_p + words_n)

# Label dichotomous sentiment categories
label <- rep(NA, nrow(China_doc))
measure <- cbind.data.frame(sentiment, label)
measure[measure$sentiment>0, "label"] <- "positive"
measure[measure$sentiment<=0, "label"] <- "negative"

# Proportion of documents in each category
prop.table(table(measure$label))  # UN General Debate tend to be positive
overall?

# Histogram by frequency
hist(measure$sentiment, xlab = "Sentiment towards China",
     ylab = "Number of documents",
     main = "Distribution of Attitude towards China in UNGA Annual
Speech", freq = T)
# Relatively normally distributed (though not mean-centered at zero)
# Use sentiment score instead of label of outcome variable.
```

```
# ggplot (better visualization)
ggplot(measure, aes(x=sentiment)) +
  geom_histogram(aes(y=..density.., fill=..count..)) +
  scale_fill_gradient("Count", low = "blue", high = "red") +
  scale_x_continuous(name = "Sentiment towards China") +
  scale_y_continuous(name = "Density of Documents") +
  ggtitle("Distribution of Attitude towards China in UNGA General Debate
Sessions (1993-2015)") +
  theme_bw(base_size = 16) +
  stat_function(fun = dnorm, args = list(mean = mean(measure$sentiment),
sd = sd(measure$sentiment)),
                color = "black", size = 1)


## 5 RQ: Does China use trade policy to buy support in UN?
UNGA_trade <- cbind(China_doc, measure)

# Log transformation
UNGA_trade$imports <- log(UNGA_trade$imports)  # all positive
UNGA_trade$exports <- log(UNGA_trade$exports)  # all positive
UNGA_trade$balance <- log10(UNGA_trade$balance+240034736)  # max. trade
deficit


# Balance of trade
reg_balance <- miceadds::lm.cluster(UNGA_trade,

sentiment~balance+factor(country)+factor(year),
                                    cluster="country" )
summary(reg_balance)



## 6 Topic models

# Load packages
libraries <- c("ldatuning", "topicmodels", "ggplot2", "dplyr", "rjson",
"quanteda", "lubridate", "parallel", "doParallel", "tidytext", "stringi",
"tidyr")
lapply(libraries, require, character.only = TRUE)

# Removes solitary letters
UNGA_trade$text <- gsub(" [A-z] ", " ", UNGA_trade$text)

# Create DFM
China_dfm <- dfm(UNGA_trade$text, stem = F, remove_punct = T, tolower = T,
remove_numbers = TRUE, remove = stopwords("english"))

# remove words with low frequency
China_dfm <- dfm_trim(China_dfm, min_termfreq = 30, termfreq_type =
"count",
                      min_docfreq = 30, docfreq_type = "count")
```

```
dim(China_dfm)


# Identify optimal number of topics K (takes too long to run...)
k_optimize <- FindTopicsNumber(
  China_dfm,
  topics = seq(from = 2, to = 30, by = 1),
  metrics = c("Griffiths2004", "CaoJuan2009", "Arun2010", "Deveaud2014"),
  method = "Gibbs",
  control = list(seed = 2019),
  mc.cores = detectCores(), # to usa all cores available
  verbose = TRUE
)

FindTopicsNumber_plot(k_optimize)


# Fit topic model using selected K
set.seed(2019)
China_topic <- LDA(China_dfm, k = 20, method = "Gibbs", iter = 1000,
                   control = list(seed = 2019))

# save the model for later retrieval
save(x = China_topic, file = 'topic2.RData')
# China_topic <- get(load('topic.RData'))


# Explore the results:
# Top 10 words for each topic
top10words <- get_terms(China_topic, 10)
top10words

# Most likely topic for each doc
top1topic <- topics(China_topic, 1)
top1topic <- cbind.data.frame(top1topic, UNGA_trade$label)
colnames(top1topic) <- c("top_topic", "category")

# Separate by positive vs. negative
China_positive <- top1topic[which(top1topic$category=="positive"), ]
China_negative <- top1topic[which(top1topic$category=="negative"), ]


# What are the main themes countries speak positively about China?
# Rank topics by the number of docs they are most likely
head(sort(table(China_positive$top_topic), decreasing = T ), 5)

# What are the main themes countries speak negatively about China?
head(sort(table(China_negative$top_topic), decreasing = T ), 5)
# Negative views have clearer implication.
```

```
## 7* Explore 1-D convergence to China's position (China vs. US)*

## Technique: Wordfish

# Use all speeches from 1992-2015 as sample (including China)
UNGA_samp <- UNGA_samp[order(UNGA_samp$country), ]

# Create DFM (with pre-processing)
UNGA_samp$text <- gsub(pattern = "'", "", UNGA_samp$text)  # replace
apostrophes
China_US_dfm <- dfm(UNGA_samp$text, stem = T, remove =
stopwords("english"),
                    remove_punct = T, remove_number = T)

# Add text labels
UNGA_samp$docname <- paste(UNGA_samp$country, UNGA_samp$year, sep = "_")
China_US_dfm@Dimnames$docs <- UNGA_samp$docname


# Fit Wordfish model: use the 2015 speech of China and US as index texts
China_US_WF <- textmodel_wordfish(China_US_dfm, dir = c(5460, 5425))
# This is time-consuming!

# 1-D visualization
textplot_scale1d(China_US_WF)

# Rank by document position
China_US_WF_rank <- China_US_WF[order(-China_US_WF$theta), ]
head(China_US_WF_rank, 10)
tail(China_US_WF_rank, 10)
```