



Universiteit  
Leiden  
The Netherlands

## **Ideological Framing in U.S. News Coverage of China**

Cai, Shitao

### **Citation**

Cai, S. (2022). *Ideological Framing in U.S. News Coverage of China*.

Version: Not Applicable (or Unknown)

License: [License to inclusion and publication of a Bachelor or Master thesis in the Leiden University Student Repository](#)

Downloaded from: <https://hdl.handle.net/1887/3487255>

**Note:** To cite this publication please use the final published version (if applicable).

LEIDEN UNIVERSITY



---

# Ideological Framing in U.S. News Coverage of China

---

International Relations and Diplomacy (MSc.)



MAY 20, 2022

Shitao Cai

Thesis

Shitao Cai - Thesis

Presented to the Faculty of  
Governance and Global Affairs

Leiden University

In partial fulfillment of the requirements for  
the Advanced Master's degree of  
International Relations and Diplomacy

By

Shitao Cai

Supervisor: Dr. Jaroslaw Kantorowicz

2<sup>nd</sup> Reader: Dr. M.S. Spirova

17692 words

The Hague, Netherlands

## ACKNOWLEDGEMENTS

I would like to first give a wholehearted thank-you to my supervisor *Dr. Jaroslaw Kantorowicz*, without whose guidance and inspiration I would never have reached where I am.

I would also like to show my gratitude to *Julian Ritter*, who has been a true comrade and companion throughout the entire journey. The daily lunchbreak with you had truly turned the days more pleasant. I shall never take it for granted.

To my parents – *Chundong* and *Wenhong* – Thank you for the unconditional support and encouragements, which had repetitively brought me back to focus. To my dear *Ruby*, all would not be the same without you. To my friends in Switzerland, France, U.S., Spain, Germany, and China – Distance never seems to sever true friendships, that is particularly true in my case thanks to you all.

*L'existence précède l'essence*. One exists to incessantly seek ways to define his/her identity, as such, every project encountered, big or small, would in a way influence the course of one's destiny. In Kundra's words, one can choose to live "lightly" – as light as air – with nothing to be weighed down. One can also choose to live heavily – engaging with responsibilities – making the life seem more real as you are tied to the ground.

During the elapsed four months, I often see in retrospect. At first, I casted upon myself tremendous amount of stress and doubt, for I was ever self-doubting in embarking on a journey that requires unprecedented willpower and strenuous efforts. By actively adapting into the situation, however, I started building up a fixed daily routing for the months to come. Energized and determined, a changed person I was. I could see the visible shift in the mentality as I became one of the earliest visitors to Wijnhaven. I can proudly say in the end: "Have you seen The Hague at 5:30? I see it often because that's when I start my day."

As a result, the last stage of writing is marked by an immense sense of purpose and pure self-indulgence. Excitements gripped me as all the puzzle pieces drawing closer and closer to a splendid whole. In hindsight, the thesis has been a spiritual journey through which I have rediscovered myself – knowing the limits and breaking through the barriers.

*Le vent se lève, il faut tenter de vivre*. The thesis marks the end of an era in my life and a commencement as new challenges await – each being seemingly as insurmountable as mountain. But there is always an ocean hidden behind that steep mountain – so climb it, if you will!

## **ABSTRACT**

This study is dedicated to discovering whether and how news framing styles can vary in the coverage of China from U.S. news media which hold different political ideologies. The *New York Times* and the *Wall Street Journal* are selected as the main samples based on their juxtaposing ideological positions. China-themed articles during Trump's Presidency are collected from the two newspapers. The study proposes sentiment analysis, Keynes analysis and Structural Topic Model to test the valence framing and the issue framing in the articles. The result shows both valence and issue framing on news coverage of China can vary depending on the ideology upheld by the news media.

## TABLE OF CONTENTS

<b>1. Introduction .....</b>	<b>6</b>
<b>2. Literature Review and Theory.....</b>	<b>8</b>
2.1 News Framing.....	8
2.2 Political Framing .....	10
2.3 Ideological Framing.....	12
2.4 Framing China.....	13
2.5 Issue Framing and Valence Framing.....	15
2.6 Hypotheses.....	15
<b>3. Methods and Data Description .....</b>	<b>16</b>
3.1 Case selection .....	16
3.2 Sample selection.....	18
3.3 Methods Description.....	21
3.3.1 Corpus .....	21
3.3.2 Document-Feature Matrix.....	23
3.3.3 Sentiment analysis .....	24
3.3.4 Keyness analysis .....	25
3.3.5 Structural Topic Model.....	26
<b>4. Descriptive statistics.....</b>	<b>28</b>
4.1 The New York Times.....	28
4.2 The Wall Street Journal .....	30
<b>5. Results.....</b>	<b>31</b>
5.1 Sentiment Analysis.....	31
5.2 Keyness Analysis.....	34
5.3 Structural Topic Model.....	37
5.4 Robustness Check .....	40
<b>6. Limitation.....</b>	<b>43</b>

**7. Conclusion.....44**  
**References.....47**  
**Appendix A.....52**

## 1. Introduction

China, as a rising superpower, has become prominent in the international arena in recent years. The image of China, however, has long been perceived by the U.S.-led West in a negative fashion, associating the country with various problems ranging from its domestic political issues (Guo, 2020), international geopolitical disputes (Fravel, 2011) and human rights contentions (Svensson, 2002). Most countries in North America and Europe, according to a 2019 survey, perceive the People's Republic of China (PRC) in a negative way. The U.S. in particular had only twenty-six percent of its people holding a favourable attitude toward China (Silver et al., 2020b). In fact, the negative perception of China had prevailed in the U.S. for an extensive period due to a series of ideological conflicts and historical rivalry between the two superpowers (Sutter, 2017). Despite of the historical aspects, the U.S. President Donald Trump, with his tough stance on China, has further turned the U.S. public opinion sharply negative against the country. Study shows, during Trump's Presidency, the percentage of U.S. citizens, who held an unfavourable attitude of China, has soared from forty-seven percent in 2017 to sixty percent in 2019 hitting a record-high in fourteen years (Silver et al., 2019). Meanwhile, the number of Americans, who perceive China as the greatest threat to the U.S, doubled from that of ten years ago (Silver et al., 2019).

Amid the prevailing negativity against China during Trump's era, however, contrasting opinions have been detected among U.S. citizens marked by their diverging political stances - most notably between the Democrats, who uphold the liberal values, and the Republicans, who endorse conservative principles. In research from Silver et al. (2020a), Republicans have historically held a more negative view towards China than Democrats, and this trend had intensified under Trump's administration. Compared to Democrats, Republicans tend to support a tougher policy stance against China and have a bigger likelihood to describe the country as an "enemy" rather than a "competitor" (Silver et al., 2020a).

On that regard, studies have established that news serve as a major informational source for citizens to gain relevant knowledge regarding foreign nations (Karlsson, 2011; Norris, 2009). In other words, public opinions on foreign countries are largely depended on the portrayal in the news coverage. An increasing number of Americans, however, believe news coverage tends to contain biases, and news media have done poorly in separating facts from opinions (Foundation, 2018). Indeed, when news media happen to be ideologically oriented, their readers are also subjected to biased (Hamborg et al., 2019). In fact, liberal and



conservative readers consciously choose news outlets based on their respective political values, as such, it leads to separate readerships in news media (El Baff et al., 2020). News outlets associating with different political ideologies - liberal or conservative – would in turn display distinct reporting styles in their news articles to shape the public opinion in a particular fashion (Islyenteva, 2020; Pan et al., 2010; Ross et al., 2021). A process named framing is involved in this phenomenon. News frames project subjective values onto an issue by presenting the coverage in a certain style, while casting focus away from aspects that did not blend in with the perceived reality that they attempt to promote (De Vreese, 2005). With the power of framing, news media can lead the public to associate, with a certain attitude, an issue, an event or even an entire country (Tewksbury & Scheufele, 2009). Furthermore, the frame adopted in news differs across news media depending on various factors - one of the most prominent being the ideological values endorsed by the news media (Pan et al., 2010).

Therefore, two preliminary assumptions can be drawn based on the existing information: 1) The public relies mainly on news as the primary source to gain knowledge of foreign countries - China in this case - and news media have their separate readerships based on the ideological divide; 2) Built upon the previous observation, if the conservatives in turn think more negatively and critically of China than their liberal counterparts, then it can be associated that conservative news media would depict China in a more negative and critical way than liberal news media. Hence, the key research question for this study remains: **Whether and how the portrayal of China can differ between the liberal and conservative news media in the U.S.?**

The originality of this study lies mainly on the relevance of the topic, the novelty of the data collection, and the proposed automated methods. In terms of the topic, the study concerns mainly the field of political communication, which studies media communication in the political process. The research proposes an innovative angle to address the China - U.S. relations from the perspectives of news articles. The result of the study shall contribute to the scholarly coverage on the ideological framing of China. In addition, contrary to the traditional qualitative methods, the automated text analysis methods enable the study to include a much larger sample size to produce a more robust and generalisable result. Regarding the overall research applicability, the design of the study can further be replicated in the future research on news framing and its influence in the perception of foreign countries.

Regarding the structure, this study first reviews relevant literature on news framing theory, its interaction with political ideology, and the framing of China. The study then selects

news articles on China from the *New York Times* and the *Wall Street Journal* based upon their juxtaposing ideological divide. The study, then, quantitatively analyses the articles from both newspapers by performing sentiment analysis, Keyness analysis and Structural Topic Model. The purpose of this study is to discover whether and how sentiments and topics can vary between *New York Times* (NYT) and *Wall Street Journal* (WSJ) regarding the coverage of China.

## **2. Literature Review and Theory**

### **2.1 News Framing**

Contrary to the common belief, studies over decades have revealed that news often does not reflect the objective reality but serves as a “socially constructed product” (Shoemaker & Reese, 1996, p.782) or a “cognitive schema” (Scheufele, 2004, p.404). In other words, the purpose of news is not merely to objectively document events, but to disseminate ideas to the general public, and to shape public opinions (Lewis & Wahl-Jorgensen, 2005). As a social construct, news is subject to influences from political, economic and ideological factors, and involves itself in a process of “cognitive simplification” named framing (Akhavan-Majid & Ramaprasad, 1998, p.132). Entman (1993, p.391) points out several key “locations” during the process of framing: To begin with, a communicator creates a framed message; This message is then carried along to the targeted receivers, and the reaction of the receivers would potentially contribute to shaping the collective culture of the target group. Such being the case, news framing, at its base, functions as a cognitive tool employed by journalists and news media to “structure the narrative flow” of events (Norris, 1995, p.357).

The origins of framing can be traced back to a broader perspective relating to social constructivism (Entman, 1993; Tewksbury & Scheufele, 2009). Goffman (1974, p.10) defines framing as a facilitator in offering characteristics and meanings for events that embed the “will, aim and controlling effort of an intelligence”. In a sense, frames can be considered as “guided doings” in subjecting events, which have occurred naturally without inherent meanings, to intentions or standards, generating a constructed “causality” in between, based on the will of the frame-maker (Goffman, 1974, p.11). According to this definition, framing in news, which serves as a major public channel for documentation, extends to all forms of natural events and human experiences (Akhavan-Majid & Ramaprasad, 1998; D’angelo, 2002; De Vreese, 2005;

Entman et al., 2009; Tewksbury & Scheufele, 2009). Mediated messages generated through purposive framing by media can exert influences on human consciousness and recognition through news articles (Entman & Rojecki, 1993). As a result, news framing has the power to influence readers' interpretation of the reality by linking an event to certain beliefs, convictions, causes, and definitions relating to that event (Entman, 1993). In other words, news frames are capable of shaping readers' opinions and attitudes and leading them to perceive the reality in a given fashion (Tewksbury & Scheufele, 2009).

Framing is not a static concept, but rather a dynamic two-step process necessitating both frame-building and frame-setting, which, although equally important in the study of news framing, often associate with different research interests and directions (De Vreese, 2005). Scholars, who study frame-building, often treat news framing as the dependent variable, whereas, news frames, in the literature for frame-setting, are often regarded as the independent variable (De Vreese, 2005, p.52; Tewksbury & Scheufele, 2009, p.22). Frame-building centres on the reason and rationale under which journalists and news media choose to adopt specific frames in the news. Multiple factors that are internal to the journalistic organization and news media, during the frame-building process, can play an important role in the final selection and adoption of frames (De Vreese, 2005; Shoemaker & Reese, 1996; Tewksbury & Scheufele, 2009; Tuchman, 1978). To begin with, news media are subject to societal norms, given that the generated news content must suit the dominant values in a society (Tewksbury & Scheufele, 2009). Meanwhile, there are, within the journalist organization, routines and regulations, which would prevent the media from adopting frames deviating to the ones established by the organization (Tewksbury & Scheufele, 2009). Political elites including various interest groups, national and local government and other state or non-state actors, on the other hand, can also exert impact to the process of frame-building (Scheufele, 1999). The dominant cultural values within a given society furthermore would contribute to the building of news frames. Frame-building, as Goffman (1974) argues, associate closely with the cultural perspective in a social constructivist point of view, because frames are not made in isolation from all the external influences but often with implicit cultural origin. More specifically, Goffman (1974) suggests that a frame links events to certain cultural aspects and leads audiences to apply the information within such context. There would often be existing news frames that are available. According to Van Gorp (2007), news media will not create sufficient effect in the news content if they adopt frames that do not fit closely with the main cultural context. Such phenomenon largely attributes to the fact news media are under the influence of the society in which they reside in,

and as a result, they would rely naturally on commonly shared frames (Tewksbury & Scheufele, 2009, p.24). By contrast, the concept of frame-setting relates closely to the socio-psychological perspective of framing - on the various psychological effects a frame can have on individuals through the exposure of a framed message, concerning the link between news framing and readers' perception and interpretation of the portrayed reality (Tewksbury & Scheufele, 2009). On an individual level, the effect of frame-setting manifest mainly in the interaction between media and audience's "prior knowledge and predispositions" (De Vreese, 2005, p.52). Entman (1993, p.52) stresses the importance of the aspect of selection in the process of frame-setting in which news media are able to make connection of definition, implication, and attitude to a perceived reality.

## **2.2 Political Framing**

News framing, with the potential to reconstruct audiences' perceived reality, has not been novel in the political world. The practice of framing, for an extensive period of time, has been openly explored by elites and news media in a political theme (Johnson-Cartee, 2004). Gamson and Modigliani (1989, p.3) suggest, elites and media often describe a particular issue using "packages" which contain essential pieces of information and linguistic devices. Those "packages" are key in shaping the audiences' belief and attitude toward political events, speeches, discussions, or announcements (Gamson & Modigliani, 1989). In short, the "packages" insert leading idea or backstory to an "unfolding strip of events" using rhetorical techniques, and endow the story with the intended meanings, which would eventually shape audiences' beliefs on the matters (Gamson & Modigliani, 1989, p.143; Tewksbury & Scheufele, 2009, p.19). This news framing technique has been repetitively exploited in the political field.

Government officials, interest groups and political institutions across the globe, in fact, have been actively engaging in multiple types of frame-building for various purposes (Scheufele, 1999). For example, Luntz (2007) discovers that political campaigns devote more resources on the delivery of political speeches and messages to manipulate the presentation of such in the news. Aside from political campaigns, government institutions also employ framing techniques to shape public opinions toward a favoured outcome. Schuck and de Vreese (2009) discovers that, during the 2005 Dutch European Constitution referendum, there was a generally positive framing towards the European Constitution in Dutch news media. As a result, some voters, who were exposed to such news frame, changed their decision in the final voting process (Schuck & de Vreese, 2009). Similarly, news framing was used in the process of electing the

political leader of a country. Strömbäck and Van Aelst (2010), compared the news coverage of political election in Belgium and Sweden, suggest a significant correlation between news framing and election, and reveal that news framing was used, during political elections, to create a façade of the electoral candidates to the voters. Similarly, Miller, Andsager and Riechert (1998) find that news framing on presidential candidates is generally detected during elections, and the framing of the candidates' political stances often deviates from those claimed by candidates themselves during press releases. As a result, news framing, regardless of its underlying purpose or its alignment to the objective reality, has the capability to swing opinions on political issues. Because it is the frame, rather than the information, on which news rely to draw association between an event and a certain attitude (Semetko & Valkenburg, 2000).

Regarding portraying political events, news media in general adopt various types of frames depending on the news objects. While subjecting to cultural differences and regional nuances, as De Vreese et al. (2001, p.109) posits, journalists and news media have a universal tendency to adopt “conflict frame”, which emphasizes on the conflict part of political events. During the launch of the Euro currency, moreover, news media had larger emphasis on the economic consequences of the new currency – in this case, a frame named “economic consequences frame” was generally adopted across news media to describe the positive and negative aspects and the potential outcomes of the introduction of the Euro (De Vreese, 2001, p.110). In a similar vein, the type of news frames adopted by the media would have consequences to the political events in the coverage. More specifically, the framed news content can potentially shape the course of events in political election. Walter and Ophir (2021, p.708) discovers news media, in terms of electoral news, would adopt the type of frames named “strategy frame”, which, instead of emphasizing on the political opinions and statements made by the electoral candidates, focuses on each candidate's personal information including motivation and character. It was revealed that the adoption of the strategy frame combined with the frequent exposure of the candidates in the news would produce a harmful effect to the final voting result to the candidate (Walter & Ophir, 2021, p.723). As a result, candidates, who are frequently portrayed by new media under the “strategy frame”, would receive less votes in the end, meanwhile, the candidate, whose opponents receive a larger coverage of strategically framed news, would gain an advantage over the final votes (Walter & Ophir, 2021, p.721). On the other hand, the power of news frames can also show, not only in the news concerning political elections, but also in the course of shaping the public opinion on the decisions made by a nation's judicial power. In their research, Hitt and Searles (2018) reveal the U.S. Supreme

Court, as the highest judicial power in the country, is perceived by the public mainly through relevant news coverage. As citizens rely mostly on news to gain relevant knowledge of the court's decisions on various matters, the framed news could potentially determine the public's attitude towards the Supreme Court's decision depicted in the news. Similar to the above-mentioned "strategy frame", Hitt and Searles (2018, p.567) underpin the adoption of the "game frame" by arguing news media, instead of focusing on reporting the key issues and policies, emphasize on the conflict, which resembles a game situation in which there are winner and loser. The study discovers that the increasing tendency over the years for the news media to adopt the game frame in depicting the Court and its decisions, as a result, would undermine the public's perception of the Court's "sincerity and principle" (Hitt & Searles, 2018, p.580).

### **2.3 Ideological Framing**

Various factors can affect news media to employ specific frames in portraying events including editorial routines of journalists (Van Dijk, 1985), external pressures from stakeholders and interest groups (Gamson & Modigliani, 1989) and ideological influences (Shoemaker & Reese, 1996; Tewksbury & Scheufele, 2009; Tuchman, 1978). Ideology stands out as being one of the most influential factors as it is predominantly "transmitted through news" (Akhavan-Majid & Ramaprasad, 1998, p.134). The concept of news framing and ideology share a large degree of resemblance and connection. Becker (1984, p.68) describes ideology as "an integrated set of frames" in which people perceive the reality and behave accordingly. News media, in turn, serve as a tool to "express", "amplify" and "disseminate" the main values carried by the ideology (Garyantes & Murphy, 2010, p.153). Three types of ideologies can exert major impact on the news frame-building process including *dominant ideology*, *elite ideology* and *journalistic ideology* (Akhavan-Majid & Ramaprasad, 1998, p.134; Becker, 1984, p73; Murdock & Golding, 1977, p35; Shoemaker and Reese, 1996, p.243).

Dominant ideology, in this context, represents the mainstream beliefs or views of a given country or the majority of citizens in the country. In the case of U.S, individualism, democracy, capitalism, and anti-communism can all be counted as parts of the country's dominant ideology (Ceplair, 2011; Eisenberg, 1999). Elite ideology, in comparison, stands for the major beliefs and ways of thinking held by the main government and the elites in power. Akhavan-Majid and Ramaprasad (2000) believe dominant ideology and elite ideology are often closely related in the context of news framing. For example, news coverage from the Cold War period shows that government would collaborate with news media to promote the dominant

ideology by portraying a common enemy (Hackett, 1984; Robinson, 2005). Study confirms, on the other hand, when there is a discrepancy between the dominant and the elite ideology, the elite ideology would exert the major influence on the framing of news (Wang, 1991).

The journalistic ideology of news media, furthermore, can create crucial impact on news framing along with the forementioned dominant and elite ideology (Akhavan-Majid & Ramaprasad, 1998). For instance, Pan et al. (2010) investigated the highlights within the coverage of gay marriage issue from the *New York Times* and the *Chicago Tribune* setting the Massachusetts legitimization of gay marriage as the key control variable. The result shows a significant journalistic divide in portraying the gay marriage issue between the two newspapers, despite both dominant ideology and elite ideology – the Massachusetts legitimization in the U.S. – remain the same in this case (Pan et al., 2010). The study thus sheds light on the importance to include journalistic ideology in the study of news framing. On that regard, studies, conducted over journalistic ideology, have defined it as “newsmen’s occupational ideology” (Schlesinger, 1978), “journalism’s occupational ideology” (Elliott, 1979) and “journalists’ professional ideology” (Zelizer, 2004). Despite of the bulk of research, most of them focused on discovering the nature of journalistic ideology and the potential factors that contribute to its formation (Deuze, 2005). Framing, as a reflection of journalistic ideology, therefore, remains largely unexplored (Akhavan-Majid & Ramaprasad, 2000; Garyantes & Murphy, 2010; Pan et al., 2010; Wang, 1991). The existing literature is, therefore, not sufficient in addressing one scenario in which journalistic ideology plays a dominant role in framing a certain country while the rest of the context remains the largely similar. A salient case proposed by the study is the overall portrayal of China in the news coverage between the liberal and conservative U.S. news media.

## **2.4 Framing China**

Historically, China has been negatively portrayed by the U.S. (Lams, 2016) largely due to the anti-communist sentiment reflected in the mainstream U.S. ideology resulting from the Cold War (Akhavan-Majid & Ramaprasad, 2000). Stone and Xiao (2007) argue that the negative portrayal of China from the U.S. news media might reflect the need for the U.S. to create a common adversary, as in the case of USSR during the Cold War. The Soviet Union, as Stone and Xiao (2007, p.93) explain, had been the primary news object for criticism in the U.S. news media associating the communist state with various negative perspectives. In the meantime, China, as a news object, was largely ignored in the U.S. newspapers until Nixon’s visit to

Beijing in 1972. Nixon's visit to China, followed by the formalization of the Sino-U.S. diplomatic relations and China's determination in conducting the economic reforms, attracted the initial interest from the U.S. press (Wang, 1992). Initially, the depiction of China was much different than that of now, as the country was described as mainly a non-threatening socialist state with a long history and remarkable cultural heritages (Xu & Parsons, 1997). Towards the end of the twentieth century, however, China started to be associated with the anti-Communist element prevailed in the contemporary U.S. dominant ideology (Akhavan-Majid & Ramaprasad, 2000). China, soon after the dissolution of the USSR, became the centre of concern for the U.S. news coverage because the country remained as the largest existing Communist country (Tefft, 1993). Based on the increased adoption of the anti-Communist frame in China's news coverage in the U.S. media after the demise of the Soviet Union, Stone and Xiao (2007, p.105) argue that the U.S. media, in the case of portraying China, "mirror" their government's policy toward the country.

Under the dominant ideology, the U.S. news media tend to negatively frame China in the news coverage for the purpose of "demotion and condemnation" (Huang & Leung, 2005). The negative sentiment towards China was further developed and transformed into dedicated news frames named "anti-communist frame" (Entman, 1991), or more the "China frame" (Akhavan-Majid & Ramaprasad, 2000; Stone & Xiao, 2007). Lee (2002) asserts that features within the "China frame" evolve over time depending on the contemporary U.S. - China relations, and the policy stance of the government in power. During his Presidency, Trump's harsh policy stance on China and his negative rhetoric against the country after the start of the U.S. - China trade war (Kapustina et al., 2020), and the outbreak of the Covid-19 (Christensen, 2020) brought more tension to the bilateral relations.

Three observations, therefore, can be drawn from the information presented above: 1) All three ideologies - dominant, elite, and journalistic ideology - can exert impact on news framing; 2) Both the dominant and elite ideology in the U.S. align closely in the negative perception of China during Trump's Presidency; 3) Journalistic values vary across the U.S. news media, and each of them have their separate readerships based on the political stances. It can then be assumed that news media endorsing different political values - liberal or conservative - portray China in their respective news coverage with nuances in the narratives, and thus, lead to the contrasting opinions between the liberals and the conservatives.



## 2.5 Issue Framing and Valence Framing

In terms of journalistic nuances, two specific framing styles - valence framing and issue framing - are crucial in examining the above-mentioned assumptions. Valence framing, on one hand, looks at the affective dimension of news framing, specifically, how news framing can affect the readers' perception and attitude of certain news topics and objects (Golan & Lukito, 2017). Valence framing theory proposes that news media highlight positive or negative aspects of the certain news objects. Such positive or negative aspects, in turn, transform into sentiments that can influence the audiences' attitudes and interpretation of the presented news topics (Schuck & De Vreese, 2006). Studies found that valence framing in news coverage can be correlated directly and significantly with audience' perception of foreign nations (Han & Wang, 2012).

Issue framing, on the other hand, is frequently studied regarding news coverage of current and salient topics (Golan & Lukito, 2017). The issue frame helps shaping the readers' understanding on issues by highlighting certain aspects about specific topics (Druckman, 2001) and providing judgements for the issues discussed (Golan & Lukito, 2017). It is especially relevant to examine the issue frames used in the news portrayal of China, because it identifies the difference in the highlights of topics across the U.S. news media as well as the different arguments and judgements imposed on the issues.

## 2.6 Hypotheses

The principal hypothesis for the study, based on the literature review, posits that the divergence in the journalistic ideology would lead the news media to adopt different frames in their news coverage of China. The dependent variable (DV) is the **framing style in the news coverage of China**, which can be operationalised in detecting the differences in the valence frame and issue frame. Meanwhile, the independent variable (IV) is the **journalistic ideology**, which can be operationalized by the liberal – conservative ideological spectrum. For the first hypothesis, liberal and conservative news media, differing in their valence framing, are expected to exhibit different overall sentiments - negative or positive - compared to each other regarding China-oriented news. Specifically, the journalistic ideology of a given news media, whether it is left leaning or right leaning in its political stance, would have correlation with its valence frame and issue frame on the topic of China. In terms of the overall news sentiment, the politically left leaning news media would portray China in a more negative way while the right leaning news media would portray the country in a more positive manner. In the second hypothesis,

the study also expects to see a divergence in terms of their respective issue framing style. Liberal and conservative news media, differing in their issue framing, are expected to highlight different set of topics related to the coverage of China. In other words, news media with different journalistic ideology would focus on different news objects. Therefore, the study predicts the following:

**H1:** Liberal news media would frame China more positively than conservative news media. Specifically, liberal news media would produce a more positive result in the sentiment of its news coverage. On the other hand, conservative news media would frame China more negatively resulting in a lower sentiment score of its coverage.

**H2:** Liberal and conservative news media in the U.S. would differ in their selection of issues and topics on China. The study predicts that liberal and conservative news media would pick different news objects of China. The divergence in the selection of news object would manifest itself in their different focus of topics as well as the key words included in the news articles. The specific topics and key words, however, cannot be predicted before the analysis as they vary from case to case.

### **3. Methods and Data Description**

#### **3.1 Case selection**

For the research, the study chooses a most-similar-systems design, which entails comparing mostly similar cases that differ mainly in the independent variable. The key to this design is to select the scenario in which the context remains fixed to a large extent, whereas the journalistic ideology varies across samples in order to accentuate the correlational effect between the independent variable - journalistic ideology, and the dependent variable - news framing of China. It is beneficial, therefore, for the study to focus on the news coverage during the presidency of a single U.S. President. Because, in the case in which the selected articles are published across the timespan of multiple Presidents, the different political emphasis of each President could potentially lead to deviating results. Furthermore, the study chooses to further limit the timeframe of articles to that of Donald Trump's Presidency from January 20, 2017 to January 20, 2021. This period is selected because the anti-China sentiment was peaked in the U.S. (Silver et al., 2020a). By focusing on this period, the research can reflect the most up-to-

date patterns of U.S. news frames regarding China. Therefore, the selection of news articles published during Trump's presidency has its scholarly novelty and salience.

Under the proposed time frame, the study will compare the news coverage of China in two mainstream U.S. newspapers - *The New York Times* (NYT) and *The Wall Street Journal* (WSJ). They are selected based on their juxtaposing ideological stances and their importance in setting the national agenda. Both newspapers are regarded as mainstream U.S. daily newspapers with a large base of readership. Ample studies have established that NYT is considered a liberal news media as it employs a "liberal tone" to frame political and social issues (Powers & Fico, 1994; Lichter et al., 1986; Ju, 2005). The WSJ, on the other hand, was more conservative in its news approach especially in its editorial section, in which opinions of the authors and the organization are reflected and manifested (Groseclose & Milyo, 2005). Although both newspapers score high in their news value and reliability, as shown in Figure 1, NYT and WSJ have an opposing stance in terms of political orientation. The vertical axis of Figure 1 examines the news value and the reliability of the news source -- the lower the position a news media has on the graph, the less reliability and value it has on its news content, which, in other words, can be understood as the trustworthiness and authenticity of the news media. The lower a news media scores in the reliability continuum, the more inaccurate information it contains in its news report. By contrast, a highly reliable news media can be seen as one that reports mainly the objective facts.

On the other hand, the horizontal axis entails the political orientation of the news media based on their news content and framing style. The more towards the left, the more liberal-leaning a news media is in its political orientation, and vice versa. A general pattern from the figure can be observed, as a result, the lower a news media scores in its news reliability, the more politically oriented it is on its news content, as it either skews extremely to the left or to the right. It can also be observed both NYT and WSJ score high in their news reliability marked by their high position of reliability in the figure. It shows that both news media mostly report facts with a large degree of analysis. It should be noted, regardless of their high reliability score, NYT skews slightly to the left and the WSJ skews slightly to the right. The relative position of two news media on the horizontal axis further confirms their divergence in journalistic ideology - NYT skews to the left, whereas WSJ leans to the right. It should also be noted both news media obtain a similar level on news reliability, meanwhile, their respective political orientations, despite of going towards an opposite direction, have about an equal distance from the centre. It demonstrates, first, the result of the analysis will be reliable because the news

articles contain mostly reliable information. And second, the independent variable tested here - the journalistic ideology - are comparably different between the two news media. As a consequence, this setup would enable a credible test of the hypotheses.

### 3.2 Sample selection

The main data for the study is therefore the corresponding news articles on the topic of China from *The New York Times* and *The Wall Street Journal*. The theme and the coverage of the

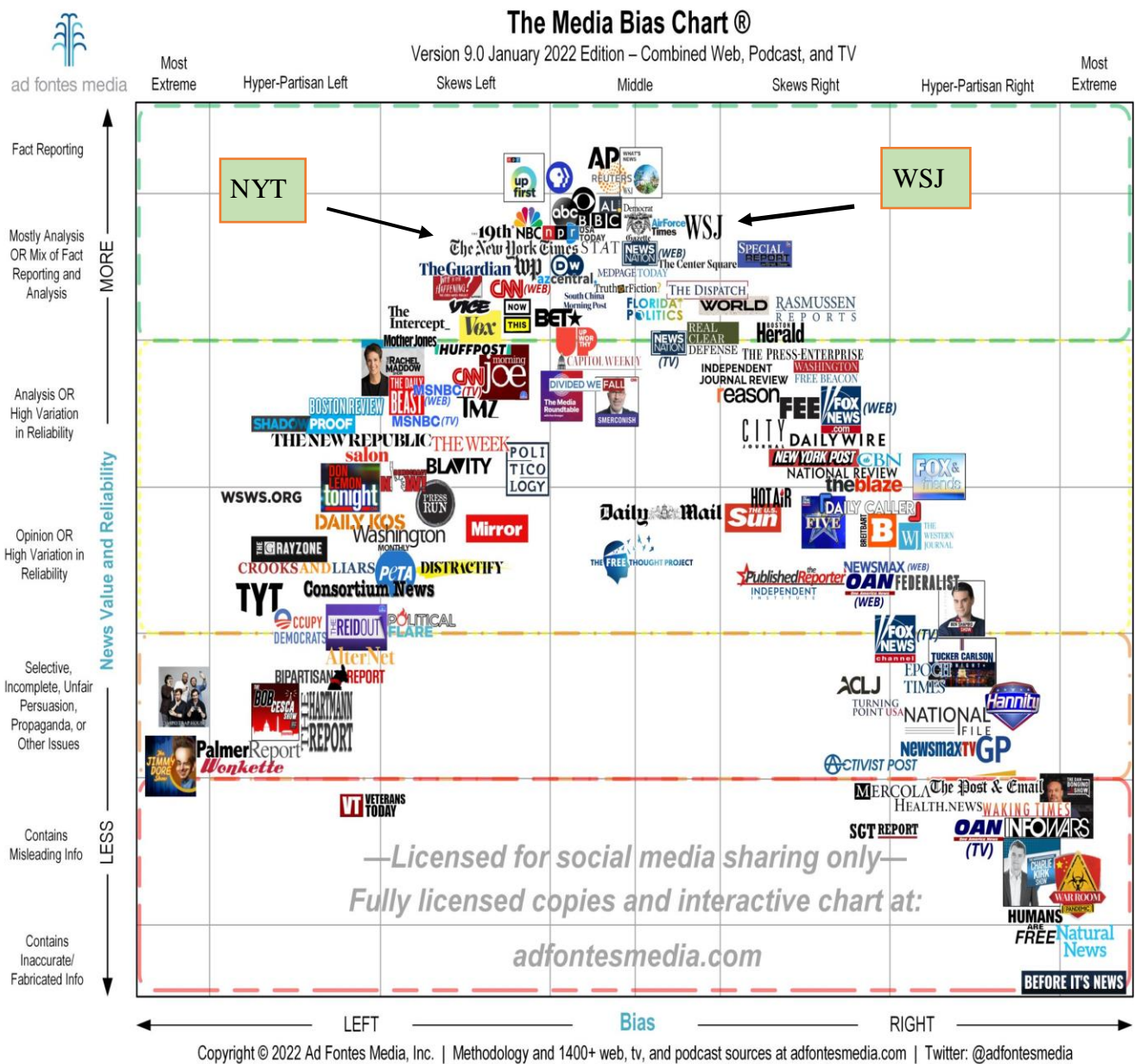


Figure 1

Muller, B. (2022), *Ad Fontes Media*, retrieved from <https://adfontesmedia.com/> Copyright © 2022 Ad Fontes Media, Inc.

articles should be specifically China-focused. In terms of being China-focused, here it specifically refers to the People's Republic of China (PRC). The reason is because the selection of the Great China region would likely bring variances to the article selection and complicate the following text analysis. Another reason, which has been discussed in the previous literature review section, is that news media tend to adopt issue-specific frames on different topics (Golan & Lukito, 2017). Therefore, different results might be generated if regions other than China are included in the article search. On top of the region selection, the articles should be strictly politics-oriented in order to reduce the noise from other non-targeted fields such as entertainment, travel and cuisine. For example, the news articles can cover famous touristic sites in China as well as critiques and reviews for Chinese films. In those cases, the ideological effect in news framing might significantly deviate from framing politics-oriented topics. It is for this reason that the study narrows its scope in its selection of news articles to solely the ones related to the topic of international relations for a more relevant and robust result.

The news articles selected would include both the article section and the editorial section of the newspapers in order to reflect the full perspective of newspaper. As mentioned previously, most News media in the U.S. are owned privately by individuals or groups. As a result, those news media, which are subject to the influence of their owner, would be more or less politically biased in terms of their news content. On that note, the article section of a newspaper concerns more about fact-reporting, which mostly contains the coverage of real events without excessive remarks or judgements made by the author or editor. As a result, the article section mostly does not manifest the opinions of the news media. Additionally, considering the high reliability score both NYT and WSJ have on their news content, it can further prove the article section provide largely reliable facts. On the other hand, the editorial section reflects the opinions of the news media on various issues as it contains mostly the opinions and judgements of the editors, who in turn reflect the political orientation of the owner of the newspaper. Therefore, it is to reflect the full perspective of the newspaper that the study picks both the article section and the editorial section of the selected news media.

In terms of data collection, the study will extract news articles from Dow Jones Factiva - an online database archiving over 32,000 major global newspapers, newswires, industry publications, magazines, reports, and other sources. Factiva can be used to locate information on a variety of topics, including politics, current events, trends, and government data, and it is a comprehensive and reliable source to extract news articles from specific newspapers, while controlling for the time frame, topics and article length. Once assessed to the Dow Jones

Factiva website, news articles can be retrieved through the “search function”, which allows the user to apply for filters to refine the search. Several filters are necessary in searching for qualified articles on Factiva:

<b>Newspapers</b>	<b>NYT</b>	<b>WSJ</b>
<b>Text</b>	China	China
<b>Date</b>	20/01/2017 to 20/01/2021	20/01/2017 to 20/01/2021
<b>Source</b>	The New York Times	The Wall Street Journal
<b>Author</b>	All	All
<b>Company</b>	All	All
<b>Subject</b>	International Relations	International Relations
<b>Industry</b>	All	All
<b>Region</b>	China	China
<b>Language</b>	English	English
<b>Results Found</b>	1276	1400

**Figure 2. Factiva Search Summary**

1. The articles must contain the key word ‘China’ either in the headline or the leading paragraph.
2. The articles need to be the traditional news articles (in contrast with online articles) published in the newspapers from either *The New York Times* or *The Wall Street Journal*.
3. The date of publication should be limited to President Trump’s Presidency from January 20, 2017 to January 20, 2021.

4. The subject should be limited to ‘International Relations’.
5. The region of the coverage should be limited to mainland China.
6. The language of the articles is limited to English.

After applying the filters, as shown in Figure 2, Factiva returns 1276 relevant articles from the NYT, compared to 1400 articles from the WSJ. It should be noted that Factiva has a function which allows to automatically detect and exclude duplicate articles. Those articles serve as duplicates to the previously published articles, and do not contain any original or unique information. As a result, the final number of articles will be less than the search result shown depending on the number of duplicates.

### **3.3 Methods Description**

#### **3.3.1 Corpus**

The study adopts a quantitative text analysis to assess the overall news frames in the selected articles. It should be noted that the traditional qualitative methods entail treating text as it is and interpreting the meanings behind it while keeping and considering all the linguistic features of the text. On the other hand, automated text analysis treats text as quantitative data reducing a large part of its linguistic features. The automated method adopts the so called “bag-of-words” approach: It treats documents as a “vector containing the count of each word type within the document, disregarding the order in which the words appear” (Lucas et al., 2015, p.257). The advantages of the quantitative method manifest in its capacity of processing a large amount of text data as well as a reduced likeliness in human-related errors in the interpretation process, generating, in the end, a more robust result compared to that of qualitative methods. To begin with, the proposed method requires collecting all the news articles and transform the raw texts in a machine-recognizable form named ‘corpus’. The corpus will combine the collected documents into one file while defining the raw texts. The corpus will provide a preliminary structure for the follow-up analysis.

In practice, the study uses ‘Readtext’ R-packages to facilitate the corpus-building process. To begin with, it should be noted that one disadvantage of obtaining large bulk of articles from Factiva is that the online database only allows, in one load, a maximum of one hundred articles to be extracted. Consequently, it requires multiple times of extraction to collect all the articles, and the resulting form of the articles will be in several files with each of them containing one hundred articles. Each file will be named in the order of its newspaper, year of

publication and month of publication separated by underscore. For example, it would be stored under the name of “NYT\_2017\_Jan.-Jun” for the first one hundred articles extracted from NYT in a chronological order. Afterwards, the files will be read into R-studio, and several columns of metadata including ‘Newspaper’, ‘Year’ and ‘Month’ will be generated based on the attached name for each file. The advantage for this is that there will be a small number of files generated and each of them will be clearly labelled differentiating the newspaper and the date of publication. This form of article storage, however, would create technical difficulty for the text analysis, because large amount of articles are compiled into one file, which will complicate the following word-trimming process. This situation is particularly disadvantageous for the Keyness analysis and the Structural Topic Model, because both methods require the corpus to be precisely trimmed in order to achieve a stable performance. Therefore, the study proposes a way to create a cleaner corpus through coding in R-studio to facilitate the text analysis.

For the cleaner version of corpus, a specific format is required when extracting articles from Factiva. There are multiple separate components consisting of a news articles including but not limited to its news type, word count, authors, source, publication date, page number, copyright information, language, headline, leading paragraph, main body, artists information and region. Therefore, it is required to select the format of articles in Factiva that includes the index for each forementioned components. For example, the index characters “HD” will be placed in front of the headline of an article, as well as “LP” being placed in front of the leading paragraph. As a result, R-studio will be able to distinguish the boundary of each article within the one-hundred-article file and sort them accordingly by setting the separation indicator to “HD”, which stands for the headline of a news article. In this way, R-studio will identify and store, instead of multiple data entries with each containing up to one hundred articles, the number of data entries matching that of articles extracted from Factiva - 1979 data entries in this case – corresponding to the same number of articles extracted from Factiva. It is important for the performance of the proposed methods that only the headline, the leading paragraph, and the main body of text of the articles are kept. Therefore, the study would need to manually remove the content headed by all other indexes such as the authors’ information indexed by “BY” and the source of the news indexed by “SN”. The exclusion of certain indexed content will preliminarily filter out most of the irrelevant words and leave a relatively cleaner corpus for further use.



### 3.3.2 Document-Feature Matrix

The second step entails converting the established corpus into a more structured form named Document-Feature Matrix (DFM). The DFM extracts words from the corpus and tabulate their counts. This matrix will become the foundation for the following text analysis methods. It is also at this step further abstraction and selection criteria can be imposed on the text data – a process named preprocessing. Preprocessing helps with further refining the selection of the unique words from the documents, because, even after the bag-of-words approach, the resulted Document-Feature Matrix may still contain numerous unique words with a large part of them consisting of the ones that are not useful to the overall research interest (Lucas et al., 2015). Therefore, by preprocessing the textual data, a large portion of the unnecessary or uninteresting words shall be removed from the collection while only the ones that are important to the analysis shall remain. In this study, the steps of preprocessing entails stemming and removing numbers, symbols, punctuations and stop words. Stemming removes the endings of conjugated verbs or plural nouns, leaving just the root of the vocabulary. For example, words such as “President” and “Presidents”, after stemming, will both become as “Presid” despite of the difference in the linguistic sense. Stop words refer to those words that are extremely common in serving as function word in a sentence but are unhelpful to the analysis such as “and” and “the”. The study will employ the built-in list from the ‘Quanteda’ package<sup>1</sup> to filter out the English stop words within the sample texts. It should be noted that after prepossessing the Document-Feature Matrix, there are in total 25410 unique features remained. These features should contain specifically the stemmed version of unique words, while all the numbers, symbols and punctuations are moved from the tabulated count. The R-function “topfeature” enables user to view the selected number of top-counted words. Thus, the top-ten most counted words in the lists, in a descending order, are “china”, “mr”, “said”, “the”, “chines”, “trade”, “trump”, “us”, “countri” and “state”. Among those top-counted words, “china” appears in total 28138 times, “trade” appears 10662 times, “trump” appears 10577 times and “us” appears 10212 times in the corpus. The list of the top-counted words, therefore, proves the face value of the study in the sense that the focus of the selected articles is on China, U.S., and trade. After

---

<sup>1</sup> The ‘Quanteda’ package in R-studio helps manipulating the texts in a corpus by performing common natural language processing tasks such as tokenizing, stemming, or stop words removing.

achieving a trimmed version of the Document-Feature Matrix, the study will start the official analysis with three proposed methods – sentiment analysis, keyness analysis and structural topic. All three methods require the trimmed version of the Document-Feature Matrix.

### 3.3.3 Sentiment analysis

For the first hypothesis, the study employs sentiment analysis to compare the valence framing between the two newspapers. Sentiment analysis reveals the affective tones lying behind the articles determining whether a body of text conveys in general a sentiment that is more positive or more negative. In general, there are two ways within the sentiment analysis that can be used to test the hypothesis – dictionary techniques and supervised machine learning techniques (Curini & Fahey, 2020, p.2). This study applies the dictionary approach due to its speed and simplicity compared to the supervised machine learning method. This approach first creates a dictionary in which a list of words is collected. This list of words will serve later as a reference in indicating the overall sentiment of a body of text. The method, then, adopts this dictionary in identifying specific words, within the collected text data, as positive or negative and calculate the sentiment of those texts based on the scores (Curini & Fahey, 2020). The key to this approach is the selection of an appropriate dictionary which would, to a large extent, include most of the sentiment-indicating words within the context of political science. It is for this purpose that the study adopts the Lexicoder Sentiment Dictionary (LSD)<sup>2</sup>, which is a broad lexicon scored for positive and negative tone and tailored primarily to political texts (Young & Soroka, 2012). Specifically, the study will apply the above-mentioned LSD in analyzing the sentiment score of the collected news articles. The result should clearly entail the sentiment scores from both the NYT and WSJ in the topic of China. Such sentiment scores shall, in turn, test the valence framing of the selected news media.

In practice, the study will load the pre-coded Lexicoder Sentiment Dictionary. After that, the study will first apply the dictionary to the corpus which contains only several, article-compiled files. The reason for the study to test this corpus is because, although uncategorized and untrimmed when comparing with the other cleaner corpus, this one is clearly labelled and

---

<sup>2</sup> The LSD is a bag-of-words dictionary designed for the automated coding of sentiment in news coverage, legislative speech and other text. It is discussed and tested in detail in Young and Soroka 2012. It is a part of the forementioned ‘Quanteda’ package.

separates the articles by the time span of roughly six months. Therefore, the result of the test can bring insight into the sentiment of the articles across time. On the other hand, the study will run the same process of the analysis on the cleaner corpus. By grouping the tokens based on the newspaper of their belonging, the study can simply categorize all the unique words into two group – one belongs to NYT and the other WSJ. The study will then run the dictionary on these two group of words in order to discover the overall sentiment of the two news media.

### **3.3.4 Keyness analysis**

For the second hypothesis, the study aims to test the issue framing of the news articles between the NYT and the WSJ with both keyness analysis and structural topic model. Keyness analysis can be understood as a tool to identify, within a given corpus, words which contain a larger frequency when comparing to a reference corpus (Gabrielatos, 2018). In other words, the method, in this case, helps recognising key words that discriminate between the articles of NYT and WSJ regarding the topic of China. As mentioned previously, issue framing highlights certain aspects about specific topics within a news coverage (Druckman, 2001). Keyness analysis, therefore, can help detecting the extent to which both newspapers differ from each other regarding their topic and news objects when portraying China, by focusing on the difference in the frequency of the key words used in the articles. In simple words, there are various ways for news media to approach the topic of China, in the sense that news media with different issue framing style might pick different aspects of China in their news stories. The potential divergence in the issue framing can be partly detected by identifying the frequently used words which discriminate between NYT and WSJ through Keyness analysis.

Regarding the method, the study will use the ‘`quanteda.textstats`’ R package. To begin with, the method requires a trimmed Document-Feature Matrix made from the cleaner corpus, which has removed numbers, punctuations, symbols, and English stop words. The second step entails a further filtering of the remaining features by assigning both a frequency threshold and a cap as well as manually excluding certain features. Regarding the limitation of frequency, the study will set the maximum document appearance ratio to be ninety percent for every unique word. This means that each unique word can only appear across no more than ninety percent of the documents. Similarly, the minimum appearance ratio is set to be one percent of the documents. That means every single unique word included in the DFM can only appear in no less than one percent of all documents. The meaning behind this step is to filter out two type of words: The first type marks words, such as “news”, “China” and “Trump”, that prevail the

majority of the documents but are rather too generic to provide any fruitful results in the interpretation process; The second type of words represents those that are rarely used across documents or irrelevant to be included in the testing such as the name of the authors. After limiting the document appearance ratio, the total number of unique words included within the list reduced significantly from the previous 25410 to 4110 unique words. It should be noted, however, it is difficult to have the exact appearance ratio to filter out all the irrelevant words. Therefore, extra filtering methods may be required afterwards. In this case, the study will manually inspect and select a group of words that are irrelevant to the research interest, and exclude them from the trimmed DFM. As a result, certain words including “factiva”, “redit” and “facebook” are manually reviewed and removed. After the filtering, the method will proceed to generate two lists of key words – one for the test group and the other for the reference group - corresponding to the totality of news articles from NYT and WSJ. The study then will consult the lists of key words in assessing and comparing the respective issue framing styles between NYT and WSJ.

### **3.3.5 Structural Topic Model**

On top of Keyness analysis, the Structural Topic Model (STM) allows the study to further test the issue frames between NYT and WSJ. The goal of the method is to allow the study to discover the most prevalent topics within the corpus, and estimate their relationship to the news media – whether certain topics are more prevalent in NYT or in WSJ. The STM is a “generative model of word counts” (Roberts et al., 2019, p.2). It adopts automatic methods to inspect each document and then use the result to locate the most likely values to match the parameters within the model (Roberts et al., 2019). The logic of the methods proceeds as such: Each topic is defined as a mixture of words, and each word has a “probability” of belonging to a topic, and each documents contains a mixture of topics with different proportions (Roberts et al., 2019, p.2). As such, the key of STM is to discover the topical prevalence, which, in its most basic level, captures to what extent each detected topic contributes to a given document. In a case-specific language, STM will shed light on the prevalent topics of the entire corpus. At the same time, the method is also able to reveal to which news media – NYT or WSJ – the detected prevalent topics associate.

In practice, the study will utilize the ‘stm’ R package to analyse the corpus and generate the major topic distribution. To begin with, the study continues with the filtered version of the Document-Feature Matrix established by the previous Keyness analysis. For the research

purpose, the study chooses to generate twenty most prevalent topics from the DFM. Although there are, in theory, no strict cut-off number for the number of topics to be generated (Roberts et al., 2019, p.12), too few topics will lead to ambiguous result. For example, if only two topics are chosen to be generated out of a large corpus, it might be difficult in the end to give a clear label to each of the two topics, because the machine is forced to compress a large quantity of key words into two topical categories – the two topics generated are likely to be extremely vague and broad in its content. Therefore, it is for the purpose of a clear categorization in the final interpretation process while considering the size of the corpus, the study chooses to generate twenty prevalent topics. Each topic will contain a set of key words that are considered, by computer, to be most likely categorized under the same topical group based on their proportion within the documents. The set of topical words contained within each topic will, in turn, help the study obtaining relevant information about its topic. The next step then entails manually inspecting the key words within each prevalent topics detected, and the study shall label them accordingly based on the key words. This step can only be done manually as the machine can only generate the groups of topical words. As such, certain degree of human knowledge and judgement is required for each topic to have a clear title. It can be expected, then, most of the twenty topics generated shall be given a clear-cut label to differentiate between each other.

On top of topical words, STM can also reveal the weight for each topic in terms of the entire corpus. This corresponds to the number of topics chosen. A corpus, as mentioned earlier, can contain multiple topics, and each topic has their own proportion for the corpus with the overall sum across all topics for a corpus being one (Roberts et al., 2019). Therefore, the weight of each topic associates with the number of topics in the sense that each topic has different proportion in a given corpus, and all topics will always add up to one regardless of the number of topics. In other words, the more topics there are, the less weight or the less representativeness each topic will have of the corpus. As such, the next step entails finding the correlation between each topic to the documents, as estimating the relationship between metadata and topics is a core feature of the ‘stm’ package (Roberts et al., 2019, p.16). Here, the study will establish a covariate which differentiates, within the corpus, the articles from NYT and from WSJ. As a result, STM will enable the study to discover to which new media, NYT or WSJ, each topic has a stronger association. After identifying their association, the study will pick the topics that are the most representative for either news media, and juxtapose them in order to find out the topical differences between NYT and WSJ. By comparing the topics combined with the result

from Keynes analysis, the study will be able to conclude whether there is a significant divergence in the issue framing of NYT and WSJ.

## 4. Descriptive statistics

In this section, the study will first examine the summary statistics which quantitatively describes and summarizes features from collected news articles from NYT and WSJ. These descriptive statistics comes from the summary function of Factiva. After returning the search result on Factiva, summary statistics of the search result can be accessed automatically regarding the distribution of articles across the selected period, the most mentioned subjects, the most mentioned regions, the most mentioned industries, the most mentioned companies and the most mentioned executives. The study will then explain and compare the statistics to give preliminary insights into the differences between the corpus of the NYT and the WSJ.

### 4.1 The New York Times

At a first glance, the study collects, in total, 1276 articles from the NYT between the period of 20<sup>th</sup> January, 2017 to 20<sup>th</sup> January, 2021. One observation is the distribution of the articles vary across the years in terms of their publication date. According to *figure 3* presented in *Appendix A*, during the period of 2017, which measures from 20<sup>th</sup> January to 31<sup>st</sup> December, 223 articles are collected. During the years of 2018 and 2019, however, the number of published articles reaches 390 and 432 respectively. For the year of 2020, however, the number drops back to 224, which is about the same level as 2017. Lastly, seven articles are included corresponding to January 2021. This finding suggests that the NYT had an increased coverage on China between 2018 and 2019, and this increase in volume has dropped in 2020. On the other hand, *figure 3* shows the top ten most mentioned subjects from the articles of NYT. Although this study applies the filter, mentioned in the previous section, which only collects the articles by the subject of ‘international relations’, the statistical result shows otherwise. According to the chart, despite of the subject of international relations being mentioned 915 times on the list, other subjects take nevertheless a large portion of the overall mentions. Noticeably, the subject of “Trade Barriers/Restrictions” scores 576 times; “Domestic Politics” reaches 469 times of mentioning. It should also be noted that “Physical Trade” which has 214 times of mentioning, along with the subject of “Trade Barriers/Restriction”, suggest that trade-related issues may be

one of the key focuses in the NYT articles on China during the targeted period. Another important observation is that there are respectively 84 and 83 times of mentioning of “Commentaries/Opinions” and “Editorials”. This finding, along with the previously mentioned fact that editorial section of a newspaper is the main source for a news media to exhibit its political orientation (Golan & Lukito, 2017), might suggest that some portion of the articles might be subject to NYT’s political stance.

Among the most mentioned regions, shown in *figure 3* (see *Appendix A*), besides from the region of China, United States, and their respective capital cities - Beijing and Washington D.C, other regions have also played a role in NYT’s China coverage including the region Eastern Asia, which mainly consist of Great China region, Japan, North Korea and South Korea. According to the chart, China and U.S. have a dominant number of mentioning of 1271 and 999 times. Hong Kong, North Korea, South Korea and Taiwan also score respectively 112, 88, 69 and 69 times, indicating their share as a subject in the overall articles. In terms of the most mentioned industry, the number of mentioning for each industry is relatively low compared to the previous charts. Among the industries, “Banking” and “Banking/Credit” sectors have their dominant proportion, while “Telecom Service”, “Agriculture”, “Healthcare” and “Automotive” also appear in the list. Among the most mentioned institutions, China Ministry of Foreign Affairs has the most mentions scoring 29 times. This finding suggests that some portion of the articles may centre on the Chinese foreign policies. On top of that, other governmental entities or institutions such as the European Union, United States Department of State, United Nations and United Nations Security Council correspond to the main targeted subject of international relations. In addition to the above-mentioned governmental entities, telecommunication companies such as Huawei and ZTE Corp also appear in the list.

Regarding the most mentioned executives, Xi Jinping, the leader of the People’s Republic of China, scores the most mentions of 650 times among the articles marking a great majority of portion compared to the other nine persons included in the list. After him, there are Steven Terner Mnuchin, the former United States Secretary of the Treasury, scores 167 times; Barack Obama, the former President of the US; scores 140 times; Kim Jong Un, the current Supreme Leader of North Korea, scoring 138 times; Michael Richard Pompeo, the former U.S. Secretary of State, scores 114 times; Lawrence Kudlow, the Director of the National Economic Council during the Trump Administration from 2018 to 2021, scores 78 times; Peter Navarro, served as the Assistant to the President and Director of Trade and Manufacturing under the Trump administration, scores 78 mentions; Wilbur Ross, the former U.S. Secretary of

Commerce, scores 78 times; Joseph Biden Jr., the current U.S. President, scores 60 times; Rex Tillerson, the form U.S. Secretary of State, scores 57 times of mentioning. An observation from the top ten most mentioned executives is that, besides Xi Jinping and Kim Jong-un, the rest of the most mentioned executives all served as U.S. government officials. This find suggests that a considerable portion of the focus might be on the U.S. governmental policy making or response towards China. Regardless of the majority being U.S. government officials, however, the U.S. President Donald Trump does not appear in the NYT's most mentioned executives list.

#### **4.2 The Wall Street Journal**

The study collects 1400 articles in total from the WSJ. Similar to the article distribution of NYT, the WSJ has the most articles on China between the year of 2018 and 2019, having respectively 448 and 452 articles according to *figure 4* presented in *Appendix A*. By contrast, only 190 articles are collected from the year of 2017, and 10 are collected from January 2021. From the article distribution of NYT and WSJ, a trend can be observed that both new media's coverage on China increases from 2017 to 2018 and peaks during 2019. The number of articles from both sides, then, drops again in 2020, marking a reduction in China-related news articles. On the other hand, *figure 4* returns the similar result of most mentioned subjects comparing with that of NYT. The subject of 'International Relations' ranks the highest, scoring 927 times, while 'Trade Barriers/Restrictions' and 'Domestic Politics' rank the second and third on the list, scoring 661 and 462 times of mentioning respectively. It should be noted that despite the majority of the most mentioned subjects being similar to those of NYT, the WSJ most noticeably contains more economics or trade-related subjects such as 'Economic News' and 'Equities Asset Class News' compared to its counterpart. Regarding the most mentioned regions, China and U.S. rank the first and second from the list. The rest of the list remains similar to the result of NYT. Certain regions such as Europe and Russia, however, while both do not appear in the NYT's list, score high in the list of WSJ, having respectively 158 and 104 times of mentioning. Two preliminary assumptions can be drawn from this divergence in the most mentioned region: 1. Both newspapers have different set of topics which, in terms of portraying, concerns difference regions or 2. Both newspapers have similar set of topics when covering China, but the concerning regions are different. Both assumptions might contribute to the hypothesis testing of the issuing framing style.

Regarding the most mentioned industries, 'Nuclear Power Generation' scores 54 times, marking the top. After it, 'Energy' sector has 50 times of mentioning; 'Passenger Cars' and



‘Industrial Good’ both score 38 times of mentioning; The majority of the rest is of heavy industry such as ‘Ferrous Metals’, ‘Aerospace Products’ and ‘automotive’, scoring respectively 29, 27 and 25 times. One remarkable difference between the two newspapers is that WSJ has more mentions on all most mentioned industries compared to the result of NYT. This might be due to the fact that the WSJ is more economically orientated, while NYT is more politically oriented. Another observation is that both newspapers have different focus in industries with NYT covering a larger variety in contrast to WSJ being more focused on manufacturing industry. In terms of the most mentioned companies, Huawei ranks the top of the list scoring 56 times; Aside from that, this list has a stronger focus on international organisations compared to NYT: United Nations scores 50 times; World Trade Organisation scores 46 times; The Boeing company appears in the list scoring 14 times of mentioning. The most noticeable difference is that the WSJ has less mentions of the U.S. government institutions. Similar to NYT, Xi Jinping is the most mentioned executive in WSJ, having mentioned 480 times. The second to the list is Donald Trump, who scores 273 times in the list. It should be noted that Vladimir Putin, the current President of Russia, appears in the WSJ’s list, having mentioned 56 times. This observation possibly corresponds to the appearance of Russia in WSJ’s most mentioned regions list.

## 5. Results

### 5.1 Sentiment Analysis

*Table 1* shows the result of the sentiment analysis for the corpus of NYT and WSJ on the topic of China. There are two parts to the result – time-specific sentiment score and total sentiment score. For the first part, the first three columns in the table correspond to metadata of the data entries including the newspaper, the year of publication and the month. The last three columns represent the sentiment scores of each data entry. The ‘positive’ column represents how many positive words are detected from the corpus that match those in the forementioned Lexicoder Sentiment Dictionary. The ‘negative’, by contrast, represents the negative ones that are found in the corpus. For every word detected, whether positive or negative, it will add one point to the corresponding column. The ‘sentiment’ column, in turn, is calculated based on the formula:

$$\textit{Sentiment score} = \textit{positive score} - \textit{negative score}$$

According to the formula, the ‘sentiment’ column measures the net sentiment score after subtracting the ‘positive’ sentiment with the ‘negative’ score. The bigger the score is in the ‘sentiment’ column, therefore, the more positive the data entry is in terms of its news sentiment. The table is, therefore, arranged on a descending order based on the ‘sentiment’ column. From

**Table 1. Sentiment Scores**

Newspaper	Year	Month	sentiment	positive	negative
WSJ	2020	Aug-Jan	531	2539	2008
WSJ	2017	Jan-July	478	2286	1808
NYT	2017	Jan-Jun	448	3513	3065
WSJ	2017	Aug-Dec	420	2398	1978
WSJ	2019	Feb-May	386	2552	2166
WSJ	2019	Sept-Jan	348	2361	2013
WSJ	2018	July-Oct	278	2223	1945
NYT	2018	Jan-May	192	3773	3581
WSJ	2018	Feb-June	164	2259	2095
NYT	2017	Jul-Nov	152	3413	3261
WSJ	2018	Nov-Feb	89	2289	2200
WSJ	2021	Jan	61	227	166
NYT	2020	Aug-Jan	57	2943	2886
WSJ	2019	June-Sept	53	2344	2291
NYT	2018	Jun-Oct	-21	3343	3364
NYT	2019	Mar-July	-40	3814	3854
NYT	2019	Sept-Feb	-84	3707	3791
NYT	2018	Nov-Feb	-119	3465	3584
WSJ	2020	Jan-June	-222	2625	2847
NYT	2019	July-Sept	-422	3463	3885
NYT	2020	Feb-July	-977	4070	5047
WSJ	Total		-267	18450	18717
NYT	Total		-4165	27049	31214

the top, it shows that most of the WSJ news entries have an overall positive sentiment score. The highest score achieved by WSJ is by articles published between August of 2020 and January of 2021, having 2539 positive score and 2008 negative score. The lowest sentiment score for WSJ is during January 2020 to June 2020 having a positive sentiment score of 2625 and a negative score of 2847. The sentiment analysis also tests on the overall score of the

articles from both NYT and WSJ within the total period selected, which associate with the last two rows in the bottom of the table. From the result, it shows that articles on the topic of China from WSJ during the exact period of Trump's Presidency has an overall positive sentiment score of 18450 and an overall negative score of 18717. After the formula, the WSJ overall sentiment remains negative as -267. On the other hand, NYT is more negative in its sentiment towards China as most of its net sentiment score stays negative. Out of the ten time-controlled data entries, there are four of them have a net positive sentiment. Among the entries, the highest one achieved by NYT represents the news articles published during the period of January 2017 to June 2017 with a net score of 448. The lowest one is during the period of February 2020 to July 2020 with an overall score of -977.

In a holistic perspective, one primary observation is that both news media have an overall negative net sentiment score, despite the fact that the net scores for both newspapers fluctuate across different periods within the targeted time span. The WSJ, however, has a net score of -267, which is relatively much more positive than that of NYT. It should also be noted that NYT, compared to WSJ, has, in both time-specific scores and the overall scores, more words detected for positive and negative sentiments. Therefore, the conclusion should be carefully phrased as: NYT, compared to WSJ, has overall a more negative sentiment towards China in relations to the calculation of all the matching sentiment words detected based on the Lexicoder Sentiment Dictionary. Another important observation is that the lowest time-specific net sentiment score for both newspapers are around the period of January to June 2020 – NYT has a net score of -977 while WSJ has a net score of -222. This period coincides with the time where the Covid-19 virus reportedly originated in China in January 2020. The outbreak of the virus, in a few months, developed into a pandemic. As the country where the coronavirus was first spotted, China was thus associated with a rather negative image by the U.S. news media during that time (Lu & Atadil, 2021). As a result, the low sentiment scores for both WSJ and NYT towards the topic of China during the initial outbreak of a pandemic further raises the robustness of the results, and demonstrate the overall credibility of the dataset. It should also be noted that the findings contradict to the hypothesis which posits that, in terms of their comparative valence framing, liberal news media would portray China more positively while conservative news media would frame China more negatively. As a consequence, it points out that liberal news media, represented in this case by NYT, would have a more negative sentiment in portraying China comparing with WSJ - its conservative counterpart.

### 5.2 Keyness Analysis

Keyness analysis, as mentioned previously in the method description section, is employed by the study, combined with the following Structural Topic Model, to test the issue framing of NYT and WSJ. In principle, Keyness analysis, as the first step in testing the issue framing,

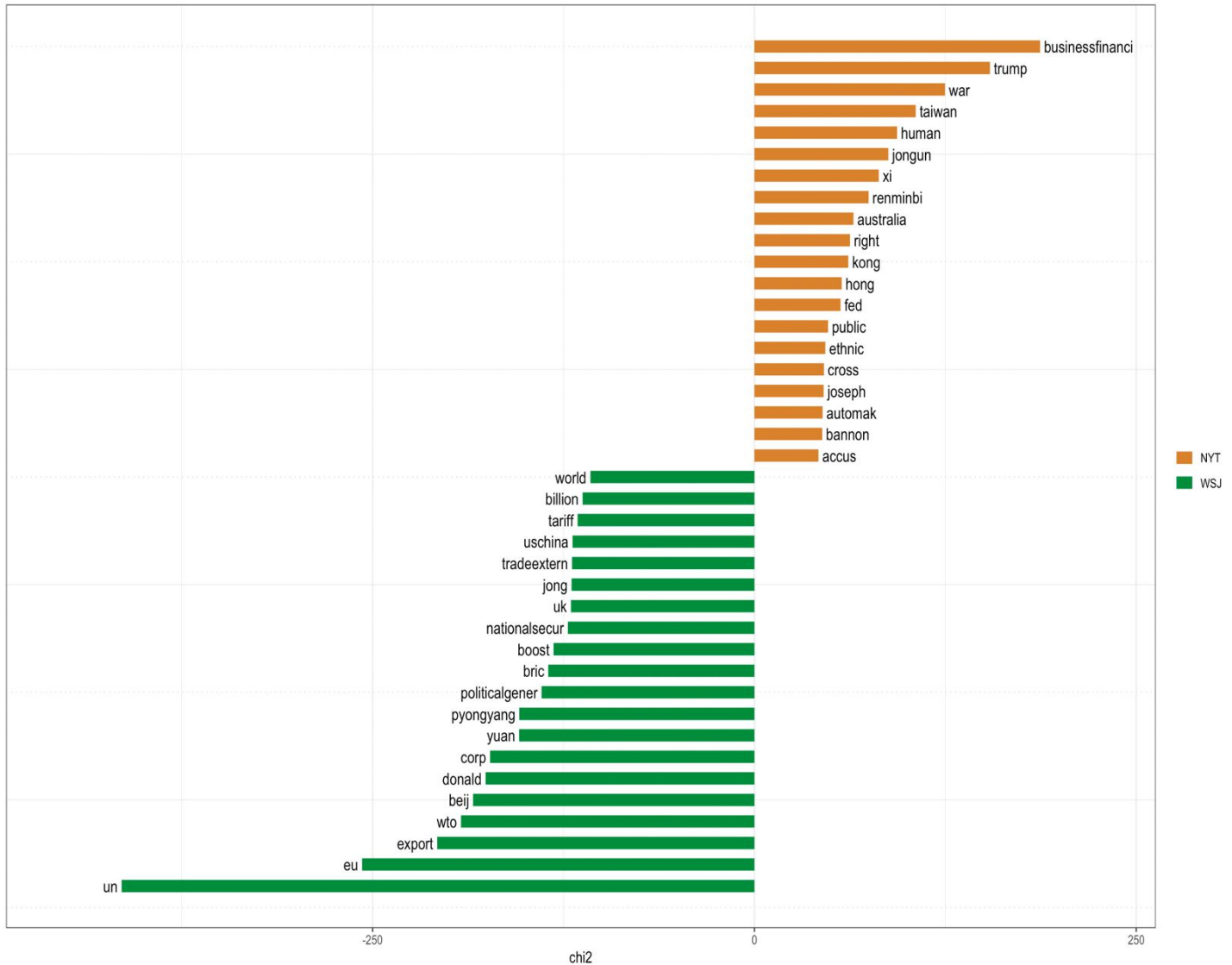


Figure 5. Keyness Analysis Plot

contrasts and showcases the key words that are used in two bodies of text. In simple words, if document A mentions the word “cake” one hundred times and document B mentions the same word only fifty times. Keyness analysis, then, will detect this difference and returns the outcome stating that the word “cake” appears fifty times more in document A than in document B. The only difference between the example and the study is that the collected sample texts contain a large number of words, therefore, Keyness analysis will detect the key words based

on their counts in the documents. To begin with, Keyness analysis requires a trimmed Document-Feature Matrix, which needs to be much cleaner than the previous one used by the sentiment analysis, because junk words<sup>3</sup>, especially the ones with large frequency within the document, may severely interfere with the outcome. Afterwards, a score for features that occur differentially across the documents will be generated.

The result, shown in *figure 5*, contains the twenty-twenty, most-counted key words differentiating across NYT and WSJ. According to the plot, the bottom-left side represents the top-counted words in WSJ that have a higher frequency compared to those in NYT. From the list, it can be observed that the word “un” appears significantly more in NYT than its journalistic counterpart. Since all words included in the analysis are stemmed and transformed into the lower case, it is necessary to double-check the words within the original articles to raise the result’s robustness. After the inspection, the study discovers the word “un” stands for UN – the abbreviated term for United Nations. Similarly, the next word in the line “eu” stand for EU – the shorter term for European Union. Thereafter, the following words include: “export”, “wto” short for World Trade Organization, “beij” short for Beijing, “donald”, “corp” short for corporate or corporation, “yuan” the Chinese currency, “pyongyang” the capital city of North Korea, “politicalgener” short for “political general news”, “bric” short for BRICS countries, “boost”, “nationalsecur” short for national security, “uk”, “jong” short for Kim Jongun – the Supreme Leader of North Korea, “tradeextern” short for trade externality, “uschina” short for U.S-China, “tariff”, “billion” and “world”. On the other hand, the top-right side represents the higher-frequency words in NYT in relation with WSJ. From the top, it shows that “businessfinanci” is the most distinct word in NYT compared to its counterpart. Other than that, there are “trump”, “war”, “taiwan”, “human”, “jongun” short for Kim Jongun, “xi” short for Xi Jinping, “renminbi” the official Chinese currency, “australia”, “right”, “Kong”, “Hong”, “fed” short for the Federal Reserve – the central banking system of the U.S, “public”, “ethnic”,

---

<sup>3</sup> Junk words referred here are the ones in a corpus or a Document-Feature Matrix that do not carry any real meanings valuable to the research purpose.

“cross” short for the Red Cross, “Joseph” short for Joseph Biden Jr. – the current U.S. president, “automak” short for automaker, “bannon” short for Steve Bannon – the former U.S.

	Topic Label	Highest probability
Topic 1	North Korea Nuclear	north, korea, korean, south, nuclear, kim, beij, missil, trump, pyongyang, offici, test, xi, presid, weapon, meet, will, militari, sanction, secur
Topic 2	U.S politics	trump, presid, administr, trade, polici, biden, hous, xi, white, will, he, leader, advis, democrat, republican, america, obama, meet, elect, nation
Topic 3	Taiwan issue	taiwan, beij, tsai, island, presid, taiwanes, elect, taipei, diplomat, govern, tie, peopl, one, polit, nation, offici, support, militari, will, mainland
Topic 4	Intelligence	offici, report, depart, foreign, govern, media, beij, diplomat, accord, work, intellig, in, research, includ, agenc, organ, washington, year, investig, nation
Topic 5	Finance	trade, year, growth, econom, stock, bank, rate, investor, war, global, currenc, month, fell, financi, quarter, cut, dollar, also, economist, last
Topic 6	India-Pakistan dispute	india, indian, border, militari, two, troop, disput, modi, side, offici, soldier, pakistan, delhi, armi, recent, along, area, territori, say, nation
Topic 7	Huawei extradition	canada, arrest, canadian, charg, case, meng, huawei, two, foreign, detain, secur, govern, detent, beij, author, extradit, offici, justic., indict, execut
Topic 8	Trade	tariff, trade, product, import, compani, good, export, billion, trump, will, year, manufactur, price, industri, steel, farmer, consum, cost, car, soybean
Topic 9	U.S-China power war	xi, world, will, america, war, power, econom, beij, global, polici, one, polit, nation, can, now, trade, intern, leader, foreign, communist
Topic 10	Covid	coronavirus, health, pandem, virus, medic, outbreak, diseas, world, govern, vaccin, global, beij, spread, in, wuhan, peopl, travel, offici, foreign, respons
Topic 11	Russian force	russia, militari, russian, nuclear, secur, europ, forc, moscow, arm, missil, weapon, treati, defens, alli, will, putin, offici, america, nation, also
Topic 12	Sanction Iran	sanction, oil, iran, north, un, energi, ship, saudi, nation, coal, export, council, iranian, offici, arabia, intern, accord, compani, east, gulf
Topic 13	Technology security	compani, technolog, huawei, industri, firm, invest, secur, govern, in, administr, busi, equip, network, appl, electron, use, includ, execut, also, chip
Topic 14	EU cooperation	european, europ, union, world, trade, global, wto, eu, climat, germani, presid, britain, agreement, leader, year, econom, western, will, german, chang
Topic 15	South China sea	sea, south, militari, island, claim, beij, philippin, vietnam, dutert, disput, territori, defens, ship, navi, water, oper, intern, region, also, naval
Topic 16	Topic 16	year, compani, busi, peopl, one, million, in, mani, work, citi, now, factori, local, show, govern, make, day, say, can, worker
Topic 17	U.S-China trade war	trade, trump, tariff, deal, negoti, beij, talk, offici, administr, econom, presid, agreement, will, two, good, billion, side, washington, week, meet
Topic 18	Belt and Road Initiative	project, road, invest, beij, belt, billion, infrastructur, govern, initi, econom, port, africa, pakistan, offici, loan, year, in, plan, one, will
Topic 19	China human-rights	hong, kong, protest, beij, right, human, govern, law, secur, peopl, xinjiang, polic, citi, mainland, foreign, uighur, nation, offici, group, muslim
Topic 20	Asia economic partnership	japan, australia, australian, region, trade, econom, minist, nation, will, japanes, foreign, beij, asian, world, partnership, prime, america, alli, abe, partner

**Figure 6. Topic Distribution**

Whitehouse Chief Strategist, and “accus” short for accuse or accusation.

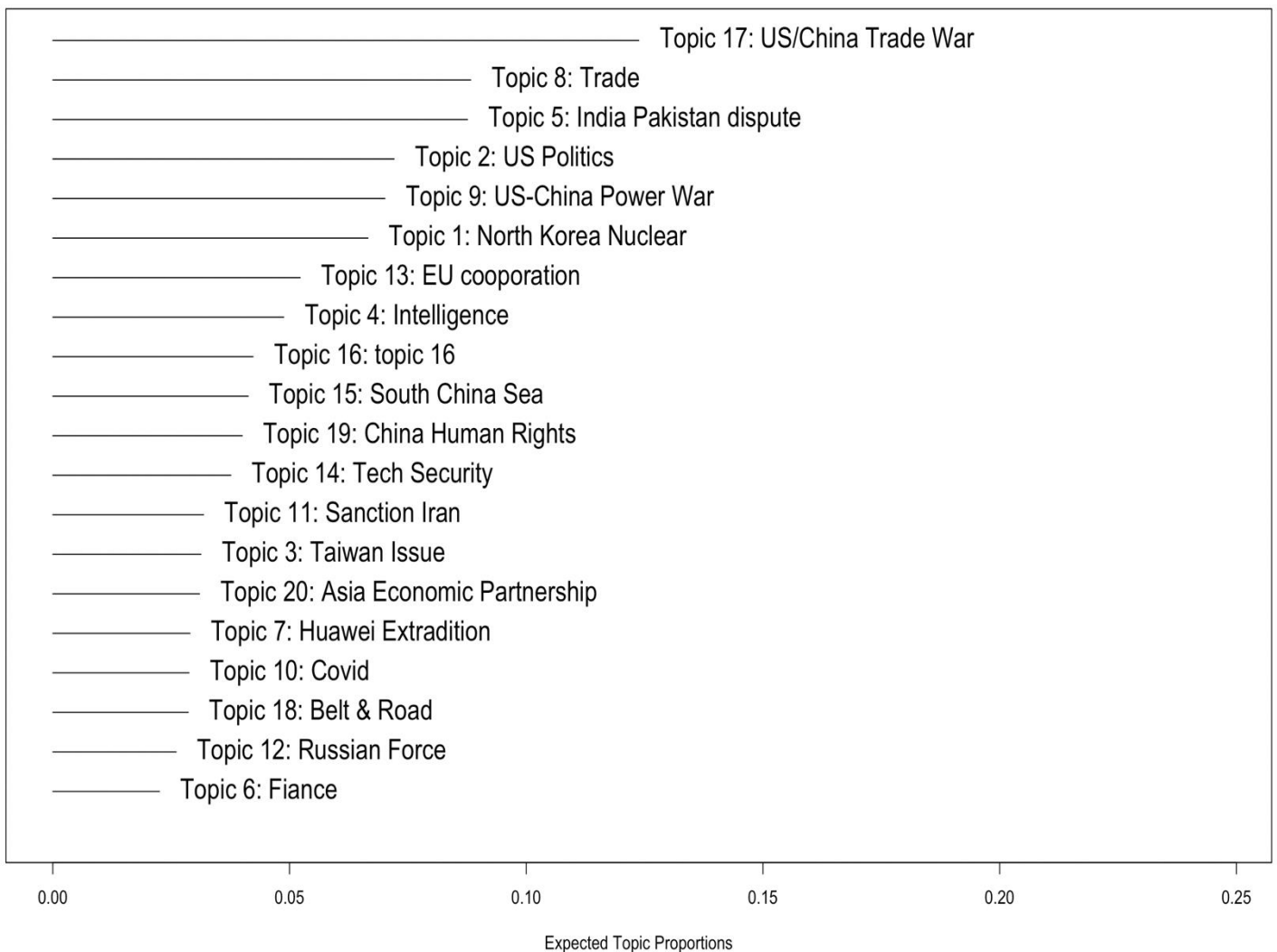
From the result, it is obvious that WSJ has a stronger focus on the economics or international finance. Most of the WSJ top-frequency words centred on this theme including “export”, “wto”, “tradeextern”, “tariff” and “billion”. On the other hand, NYT tends to focus more on the ongoing, prevalent, international-political issues. This tendency is marked by words such as “war”, “taiwan”, “human”, “xi”, “australia”, “right”, “hong”, “kong”, “ethnic” and “cross”. Each of the words included represent an ongoing political issue. Therefore, NYT, from the perspectives of Keynes analysis, is more comprehensive in its news topics and subjects, and tends to cover a larger variety of issues not limiting to finance and economics in terms of portraying China compared to WSJ.

### 5.3 Structural Topic Model

On top of the Keynes analysis, Structural Topic Model (STM) serves as a next step in testing the hypothesis regarding issue framing. Unlike Keynes, which mainly tests the differentiating key words between documents, STM reveals the prevalent topics among all the documents. The number of the topics are set manually. In this case, the study chooses to generate twenty initial topics. According to the *Figure 6*, the twenty topics are generated in the form of topical words based on their affinity with each other across the documents. Afterwards, specific labels will be given to each topic in order to differentiate and to give a clear notion for the study to compare among them. It should be noted, however, there is not a correct number of topics to be generated for any given corpus. Therefore, it is normal that some of the topical words within a topic seems not compatible with the rest majority. The bottom line is, therefore, the vast majority of topical words contained within a given topics should enable the study to clearly distinguish this topic from others for the follow-up analysis. In addition, it is also normal that a small minority of topics will be left undefined, as the topical words contained within do not suffice for this topic to be labeled. The first step, after the topics are generated, is to inspect the topical words within each topic. Each topic will then be named based on the majority of topical words included. For Topic 1, as shown in *figure 6*, it shall be titled as “North Korea Nuclear” as it contains many key words indicating the general affinity to the title such as “north”, “korea”, “nuclear”, “kim” as Kim Jongun, “pyongyang” and “weapon”. The principle is that a title does not necessarily need to reflect the full aspects of the topic, as long as it shows the relevant characters or dimension outlined by the topical words. Based on this principle, Topic 2 is named as “U.S. politics” for it centres on the theme of U.S. domestic politics and

includes “trump”, “biden”, “democrat”, “republican”, “elect” and “Obama”. Topic 3 is named as “Taiwan issue” for it contains many key words on such subjects including “taiwan”, “island”, “taiwanese”, “tsai” – President of Taiwan, “mainland”. Topic 4 is named as “Intelligence” as it contains words including “intellig” as intelligence, “agenc” as agency and “investig” as investigation. Topic 5 is named as “Finance” as it includes “trade”, “growth”, “econom”, “stock”, “bank”, “investor” and “financi” as in “financial”. Topic 6 is named as “India-Pakistan dispute” for it contains “india”, “pakistan”, “border”, “militari” and “territori”. Topic 7 is named as “Huawei extradition” as it includes “huawei”, “canada”, “arrest” and “meng” as in Meng Wanzhou – CFO of Huawei. Topic 8 is named as “Trade” as it includes “tariff”, “trade”, “product”, “import”, “compani” and “export”. Topic 9 is named as “U.S-China power war” for it contains “xi”, “america”, “power”, “war”, “communist” and “leader”. Topic 10 is named as

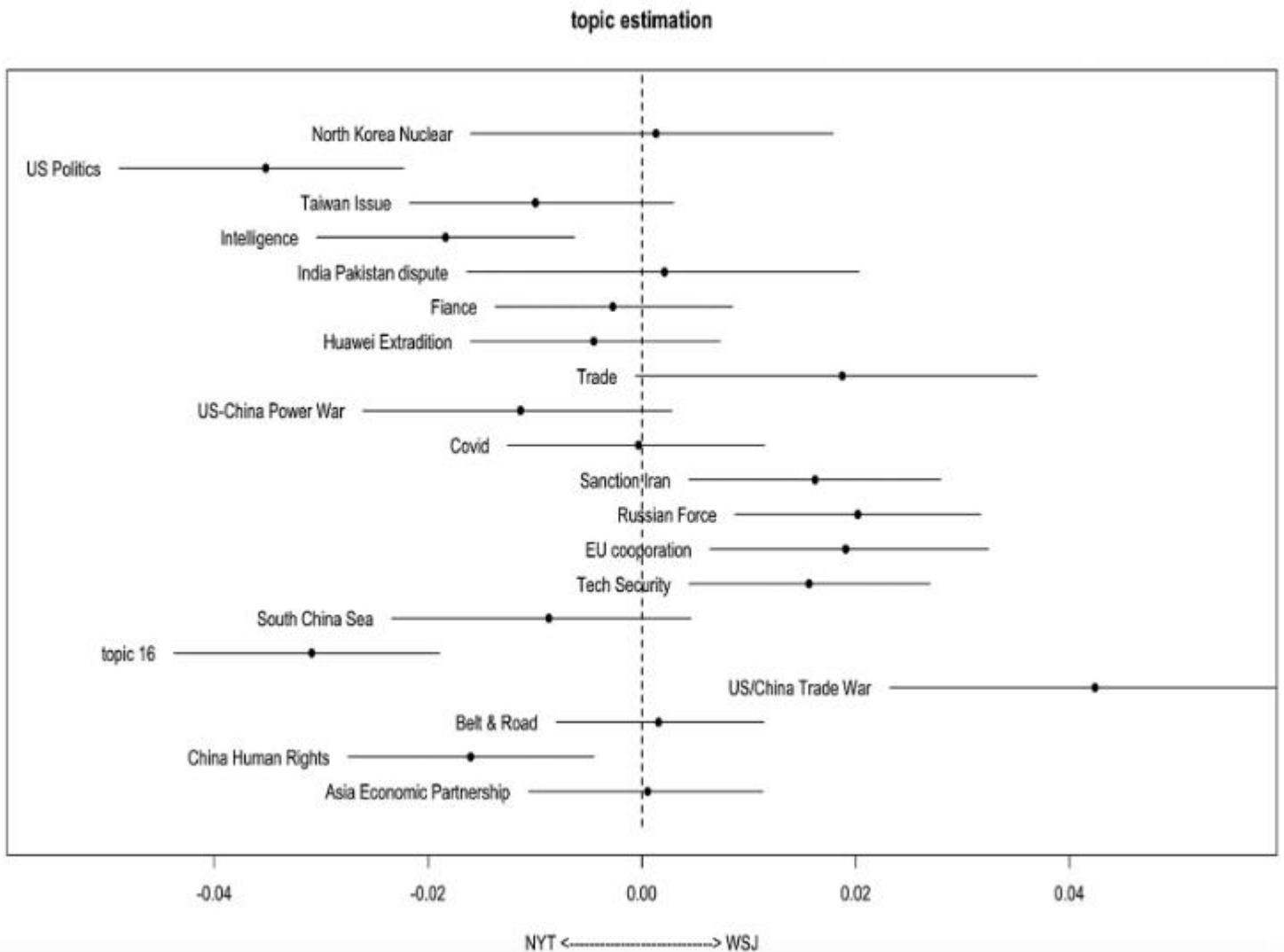
Top Topics



**Figure 7. Expected Topic Proportions**



“Covid” as it has “coronavirus”, “pandem” as in “pandemic”, “medic” as in “medical” and “spread”. Topic 11 is named as “Russian force” as it contains “russia”, “militari”, “forc” as in “force” and “putin”. Topic 12 is named as “Sanction Iran” because “iran”, “sanction”, “oil” and “energi” are included. Topic 13 is named as “Technology security” because it includes “technolog” as in “technology”, “equip” as in “equipment” and “electron” as in “electronic”. Topic 14 is named as “EU cooperation” as it contains “european”, “union”, “eu” as in “European Union”, “western” and “agreement”. Topic 15 is named as “South China sea” as “sea”, “south”, “island”, “beij” as in “Beijing”, “philippin” and “vietnam” are spotted. Topic 16 lacks a coherent topical logic to be given a specific title, thus it remains unchanged as “Topic 16”. Topic 17 is named as “U.S-China trade war” for the inclusion of “trade”, “trump”, “tariff”,



**Figure 8. Topic Association**

“negoti” as in “negotiation are “beij” as in “Beijing”. Topic 18 is named as “Belt and Road

Initiative” for it contains “project”, “road” and “belt”. Topic 19 is named as “China human-rights” for it has “hong”, “kong”, “protest”, “xinjiang” and “muslim”. Topic 20 is named as “Asia economic partnership” for it includes “japan”, “australia”, “trade”, “partnership” and “econom” as in “economy”.

The next step after properly naming each topic entails finding out their respective proportions or weights within the corpus. It should be noted that regardless of the number of topics generated, the total added proportion should equal to one. From the *Figure 7*, Topic 17 regarding U.S-China trade war has the highest proportion covering about ten to fifteen percent of the corpus content. Besides from that, topics including “Trade”, “India-Pakistan dispute”, “US politics”, “North Korea nuclear” and “EU cooperation” account for five to ten percent of the contents across the documents. The rest of the topics have generally a similar proportion of less than five percent. In addition to weighing each topic in relation to the corpus, it is important to pinpoint their correlations with the documents – whether any given topic is more correlated with NYT or WSJ.

As *Figure 8* shows, the overall result can be divided into three categories – topics that are more prevalent in the articles of NYT, topics that are more prevalent in WSJ, and topics that are mentioned with equal portions in both media. For topics that are more prevalent in the articles of NYT, they include “U.S. politics”, “Taiwan issue”, “intelligence”, “U.S-China power war”, “South China sea” and “China human rights”. Topic 16, although has a high prevalence towards NYT, is excluded from the consideration due to its lack of defining title. For topics that are more prevalent in WSJ, on the other hand, “U.S-China Trade War” has a largest proportion. The rest of the topics includes “Trade”, “Sanction Iran”, “Russian Force”, “EU cooperation” and “Tech. Security”. It should also be mentioned that both newspapers have a similar prevalence of certain topics including “North Korea Nuclear”, “India Pakistan dispute”, “Finance”, “Huawei Extradition”, “Covid”, “Belt and Road” and “Asia Economic Partnership”. It can be observed that, therefore, despite of selecting articles that are specifically international relations-oriented and China-focused, NYT tend to cover more on political topics, whereas WSJ has a strong focus on economic or finance-related news subjects.

#### **5.4 Robustness Check**

The study, in addition to the three main automated methods, proposes a qualitative inspection to test the robustness of the results from Keynes analysis and Structural Topic Model. For

such purpose, the study will randomly pick small portions of texts from all the articles differentiating on the selected news media based on the key words and topics generated. Such test, serving as a quality control, would further increase the robustness and incredibility of the automated results by adding a qualitative layer to the overall inspection.

Table 2 contains the randomly extracted sentences and their corresponding topics from the *New York Times* based on the Structural Topic Model results mentioned in the previous section. There are, on the left side of the table, the prevalent topics and, on the right side, the randomly selected text portions. According to the STM results, six main topics listed in the table are prevalent mainly in NYT articles. Thus, the study randomly searches for sentences which correspond to each prevalent topic indicated. The study picks the main key word within each topic and utilize the ‘find’ function to locate the matching sentences contained in the NYT corpus. All the key words used for the searching purpose are highlight in red. The following results prove that the categorization for the prevalent topics is accurate to a large extent as each text portion, despite of being randomly extracted from the overall NYT corpus, captures the essence of each corresponding topic.

**Table 2. Robustness Check for the New York Times**

<i>U.S. Politics</i>	The people with the most ability to limit the damage are <b>Republicans</b> who see themselves as advocates of a strong America.
<i>Taiwan Issue</i>	President of <b>Taiwan</b> Says China 's Military Expansion Could Destabilize the Region
<i>Intelligence</i>	The director of the F.B.I., Christopher A. Wray, testified recently that he believed universities were underestimating the ability of Chinese students to collect valuable national security <b>intelligence</b> .
<i>U.S.- China Power War</i>	In some respects, Mr. Xi's move to extend his rule in tandem with his drive to make China a dominant global <b>power</b> should not have surprised the United States, Chinese analysts said.

<i>South China Sea</i>	"The U.S. is virtually the last man standing to which Hanoi can look for support in the <b>South China Sea</b> dispute."
<i>China Human Rights</i>	Human rights advocates and legal scholars say the mass detentions in the northwest region of Xinjiang are the worst collective <b>human rights</b> abuse in China in decades.

On the other hand, *Table 3* shows similar result for the *Wall Street Journal* as each randomly extracted sentence through key word searching aligns with the main categorization of the topics. In other words, each sentence or paragraph extracted captures the essential message of its corresponding topic. Therefore, this robustness check proves the prevalent topics generated by the Structural Topic Model to a large extent capture the essence of the overall topic layout.

***Table 3. Robustness Check for the Wall Street Journal***

<i>U.S. - China Trade War</i>	Chinese misbehavior has thus brought the U.S. to the brink of a <b>trade war</b> with its own economic and strategic allies.
<i>Trade</i>	The administration's actions, including a threat of <b>tariffs</b> on \$60 billion of imports and tighter restrictions on acquisitions and technology transfers, came in response to what it said were Chinese efforts to obtain U.S. technology through intimidation, state-financed acquisition and subterfuge.
<i>Sanction Iran</i>	Following through on threats to impose tariffs would have escalated tensions between the U.S. and its European allies at a moment when other high-stakes discussions were under way on the <b>Iran</b> nuclear deal, U.S. sanctions against Russia, how to confront economic challenges from China and other issues.

<p><i>Russian Force</i></p>	<p>The U.S. ejected <b>Russia</b>, for example, after Moscow's incursion into Ukraine in 2014.</p>
<p><i>EU Cooperation</i></p>	<p>China courted the <b>European Union</b> as an ally in its trade conflict with the U.S., offering to improve access for foreign companies and work with the EU on overhauling the World Trade Organization .</p>
<p><i>Tech Security</i></p>	<p>Those relations have soured, especially after Mr. Trump's initial tariffs on Chinese imports, which he said were designed to punish Beijing for alleged intellectual-property violations and <b>technology</b> theft.</p>

## 6. Limitation

The overall validity and applicability of the study is subject to certain limitations. On the research level, this study is mainly interested in discovering to what extent and in what manner the portrayal of China can differ between the liberal and conservative news media in the United States based on their journalistic, ideological differences. The selected time frame is, however, limited to that of Trump's presidency. Therefore, the result of the study is limited to address only the news framing nuances in the U.S. during Trump's Presidency. Second, the sample selection is limited to the *New York Times* and the *Wall Street Journal*. Although both news media are mainstream, influential, and representative in the U.S. and their juxtaposing political orientations fit well into the research design, the results cannot reflect the full aspects of all the existing liberal and conservative new media in the country. Furthermore, both NYT and WSJ, as mentioned in the sample selection, score high in reliability, and only slightly drift way from being politically neutral. Therefore, the results may vary in the case when different news media, which, for instance, have a low score in news reliability or an extreme political orientation, were to be selected.

On the method level, the automated quantitative approaches, on one hand, enables the study to collect and process a larger number of data than the traditional qualitative methods. The result, on the other hand, compromises the overall accuracy due to the lack of controlled

human inspection. For example, the sentiment analysis, although provides a clear image on the news sentiment score for the included news media, utilize a pre-coded dictionary to detect sentiment-indicating words in each document. The result, therefore, might be subject to errors for two reasons: First, the pre-coded dictionary might not be able to exhaust all the words that can be used to mark sentiment in political articles; And second, the phrasing in political news may contain complicated rhetorical techniques in expressing the true sentiments. Thus, computer might have trouble deciphering and judging all the correct sentiments. Keyness analysis, on the other hand, is also prone to the similar type of errors. The method requires a more vigorous pre-processing of the corpus, and thus is more susceptible to lose key words during the process. For example, the study sets the maximum document appearance ratio to be ninety percent and the minimum appearance ratio to be one percent of the documents for every unique word. This cut-off ratio would, though bringing a clearer corpus with much less remaining words, potentially filter out key words which might contain valuable information to the research. The Structural Topic Model, finally, is also subject to loss of information in the sense that the method tends to simplify the corpus data by compressing all the potential topics from all the articles into a selected few - twenty in this case. The result, therefore, may likely fail to capture exhaustively all the possible topical dimensions contained in the corpus. The large sample size of the study and the proposed machine-assisted methods, however, still promise a credible overlook of the main body of texts regarding the final results despite of the forementioned limitations.

## **7. Conclusion**

To sum up, this study starts with an observation that people in the United States holds, in general, a negative view towards China, and that the opinions between the Democrats and the Republicans vary distinctly on such regard. In addition, news media are proven to be the main source for people to gain knowledge of foreign countries. As the news media in the U.S. generally diverge in their political orientations, the study thus proposes an assumption that news media may contribute to this split of public opinions as they can portray China in a diverging way. The main interest of this research, therefore, centres on finding out whether and how the portrayal of China can vary in the U.S. news media based on ideological differences. Based on the literature, examining the style of news framing can be operationalized into testing the valence framing and the issue framing - the former tests the overall news sentiments, while

the other focuses on the main topics contained in the news. Two hypotheses are thus derived. The first hypothesis concerning valence framing states that liberal news media would frame China more positively than conservative news media. The second hypothesis regarding issue framing posits that liberal and conservative news media would differ in their selection of issues and topics of China. The *New York Times* and the *Wall Street Journal* are thus introduced as the key sample for this study as the former is liberal-oriented and the other upholds conservative values. The study limits its scope of data selection to the China-oriented news articles during the Trump's Presidency. The study proposes three automated methods to test the valence framing and the issue framing of the samples. The result from the sentiment analysis, on one hand, proves that there is a difference in the news sentiments between *New York Times* and *Wall Street Journal*. The results, however, negates part of the first hypothesis because the news sentiment of the NYT is more negative in the coverage of China than that of the WSJ. Such conclusion is further proved by the following sentiment scores differentiating the time. Although the sentiment scores of NYT fluctuate across different periods during Trump's Presidency as well as those of WSJ, the general trend still proves that NYT has in general a more negative sentiment on China than WSJ. The results from Keyness analysis and Structural Topic Model, on the other hand, proves the divergence in the topic selection between the articles of NYT and WSJ. Even though all the articles in the topic of China are selected based on a series of criteria, the articles from the New York Times tend to focus more on political topics. Whereas the *Wall Street Journal* has a stronger focus on economic or finance-related news subjects.

In conclusion, ideological difference plays a role in the news framing of China among the U.S. news media. The difference, on one side, manifests in their respective news sentiments – liberal news media represented by the *New York Times* tend to be more negative, while conservative media tends to be more positive in the coverage of China. On the other side, liberal and conservative news media tend to include and cover different set of topics on China as NYT focuses more on politically oriented issues, while WSJ devotes more emphasis on economy-related topics. There are two potential directions to which future studies can further explore. On the first direction, the external validity of this research can be further tested: As political orientation is a continuous spectrum, it would be important to find out whether other news media, which score differently in their political ideology spectrum than WSJ and NYT, would produce similar results regarding their news sentiments as well as their selection of issues. The

other direction entails exploring whether in other countries, which also hold a generally negative view towards China, would produce similar results.



## References

- Akhavan-Majid, R., & Ramaprasad, J. (1998). Framing and ideology: A comparative analysis of US and Chinese newspaper coverage of the fourth United Nations conference on women and the NGO forum. *Mass Communication and Society*, 1(3-4), 131-152.
- Akhavan-Majid, R., & Ramaprasad, J. (2000). Framing Beijing: Dominant ideological influences on the American press coverage of the fourth UN Conference on Women and the NGO Forum. *Gazette (Leiden, Netherlands)*, 62(1), 45-59.
- Becker, S. L. (1984). Marxist approaches to media studies: The British experience. *Critical Studies in Media Communication*, 1(1), 66-80.
- Ceplair, L. (2011). *Anti-communism in twentieth-century America: a critical history*. ABC-CLIO.
- Christensen, T. J. (2020). A modern tragedy? COVID-19 and US-China relations. *Brookings Institution*.
- Curini, L., & Fahey, R. A. (2020). *Sentiment Analysis and Social Media*.
- D'angelo, P. (2002). News framing as a multiparadigmatic research program: A response to Entman. *Journal of communication*, 52(4), 870-888.
- De Vreese, C. H. (2005). News framing: Theory and typology. *Information design journal & document design*, 13(1).
- De Vreese, H., Jochen Peter, Holli A. Semetko, C. (2001). Framing politics at the launch of the Euro: A cross-national comparative study of frames in the news. *Political communication*, 18(2), 107-122.
- Deuze, M. (2005). What is journalism? Professional identity and ideology of journalists reconsidered. *Journalism*, 6(4), 442-464.
- Druckman, J. N. (2001). The implications of framing effects for citizen competence. *Political behavior*, 23(3), 225-256.
- Eisenberg, J. (1999). How individualism-collectivism moderates the effects of rewards on creativity and innovation: A comparative review of practices in Japan and the US. *Creativity and Innovation Management*, 8(4), 251-261.
- Entman, R. M. (1991). Framing US coverage of international news: Contrasts in narratives of the KAL and Iran Air incidents. *Journal of communication*, 41(4), 6-27.
- Entman, R. M. (1993). Framing: Towards clarification of a fractured paradigm. *McQuail's reader in mass communication theory*, 390-397.

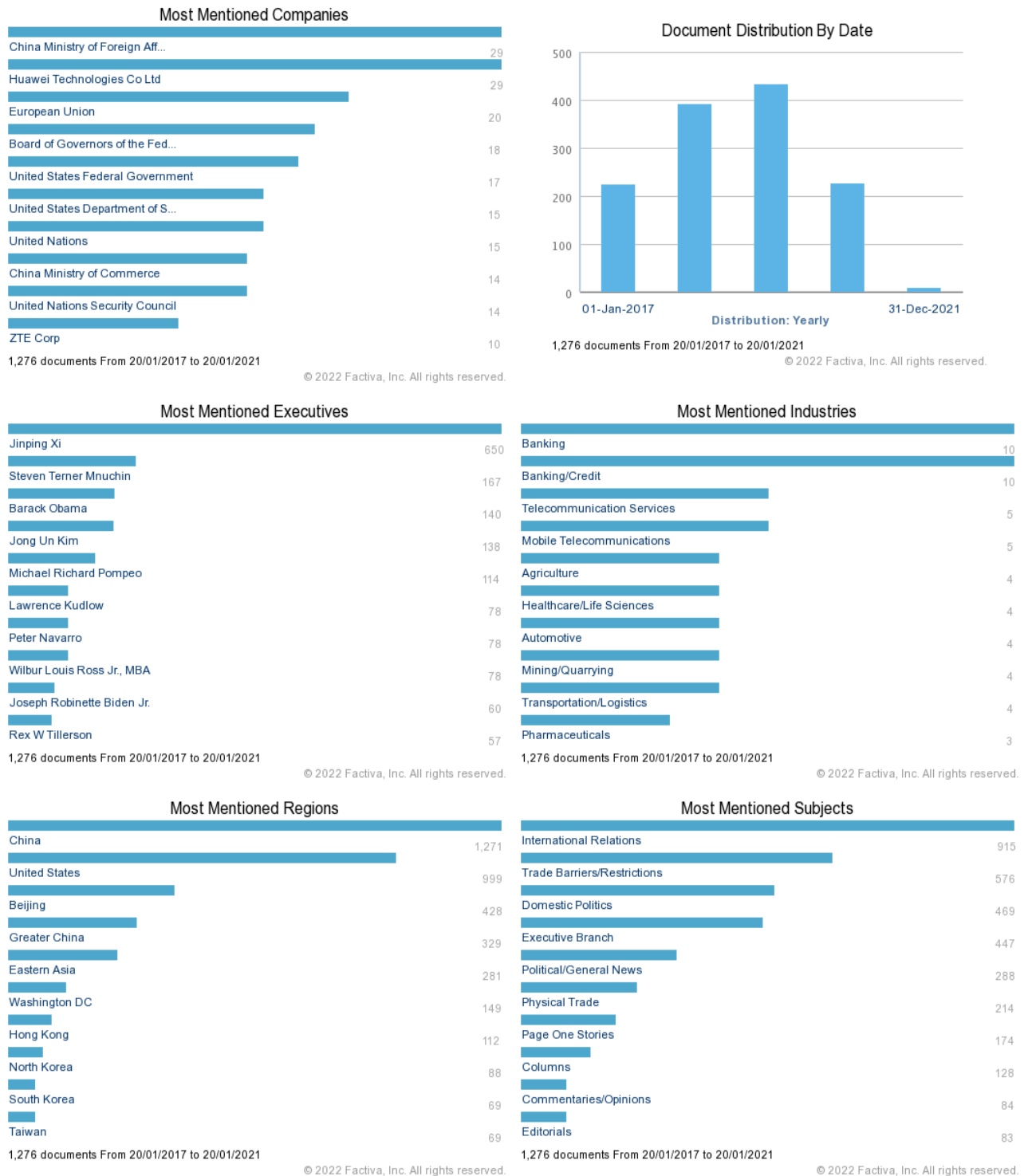
- Entman, R. M., & Rojecki, A. (1993). Freezing out the public: Elite and media framing of the US anti-nuclear movement. *Political Communication*, 10(2), 155-173.
- Entman, R. M., Matthes, J., & Pellicano, L. (2009). Nature, sources, and effects of news framing. *The handbook of journalism studies*, 175-190.
- Foundation, K. (2018). American views: Trust, media, and democracy.
- Fravel, M. T. (2011). China's strategy in the South China Sea. *Contemporary Southeast Asia*, 292-319.
- Gabrielatos, C. (2018). Keynes analysis. *Corpus approaches to discourse: A critical review*, 225-258.
- Gamson, W. A., & Modigliani, A. (1989). Media discourse and public opinion on nuclear power: A constructionist approach. *American journal of sociology*, 95(1), 1-37.
- Garyantes, D. M., & Murphy, P. J. (2010). Success or chaos? Framing and ideology in news coverage of the Iraqi national elections. *International Communication Gazette*, 72(2), 151-170.
- Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Harvard University Press.
- Golan, G. J., & Lukito, J. (2017). Newspaper editorial pages frame China similarly. *Newspaper Research Journal*, 38(2), 215-230.
- Groseclose, T., & Milyo, J. (2005). A measure of media bias. *The Quarterly Journal of Economics*, 120(4), 1191-1237.
- Guo, Y. (2020). *Freedom of the Press in China*. Amsterdam University Press.
- Hackett, R. A. (1984). Decline of a paradigm? Bias and objectivity in news media studies. *Critical Studies in Media Communication*, 1(3), 229-259.
- Hamborg, F., Donnay, K., & Gipp, B. (2019). Automated identification of media bias in news articles: an interdisciplinary literature review. *International Journal on Digital Libraries*, 20(4), 391-415.
- Han, G., & Wang, X. (2012). Understanding "Made in China" valence framing and product-country image. *Journalism & Mass Communication Quarterly*, 89(2), 225-243.
- Hitt, M. P., & Searles, K. (2018). Media coverage and public approval of the US Supreme Court. *Political Communication*, 35(4), 566-586.
- Huang, Y., & Leung, C. C. M. (2005). Western-led press coverage of Mainland China and Vietnam during the SARS crisis: Reassessing the concept of 'media representation of the other'. *Asian Journal of Communication*, 15(3), 302-318.

- Isentyeva, A. (2020). *Corpus-based analysis of ideological bias: Migration in the British press*. Routledge.
- Johnson-Cartee, K. S. (2004). *News narratives and news framing: Constructing political reality*. Rowman & Littlefield Publishers.
- Ju, Y. (2005). The bipolar liberalism reflected in the New York Times coverage of gay marriage. In *annual meeting of international communication association*.
- Karlsson, M. (2011). The immediacy of online news, the visibility of journalistic processes and a restructuring of journalistic authority. *Journalism*, 12(3), 279-295.
- Lams, L. (2016). China: Economic magnet or rival? Framing of China in the Dutch-and French-language elite press in Belgium and the Netherlands. *International Communication Gazette*, 78(1-2), 137-156.
- Lee, CC (2002) Established pluralism: US elite media discourse about China policy. *Journalism Studies* 3: 343–357.
- Lewis, J., & Wahl-Jorgensen, K. (2005). Active citizen or couch potato? Journalism and public opinion. *Journalism: critical issues*, 98-108.
- Lichter, S. R., Rothman, S., & Lichter, L. S. (1986). *The media elite*. Adler Publishing Company.
- Lu, Q., & Atadil, H. A. (2021). Do you dare to travel to China? An examination of China's destination image amid the COVID-19. *Tourism Management Perspectives*, 40, 100881.
- Lucas, C., Nielsen, R. A., Roberts, M. E., Stewart, B. M., Storer, A., & Tingley, D. (2015). Computer-assisted text analysis for comparative politics. *Political Analysis*, 23(2), 254-277.
- Luntz, F. (2007). *Words that work: It's not what you say, it's what people hear*. Hachette UK.
- Miller, M. M., Andsager, J. L., & Riechert, B. P. (1998). Framing the candidates in presidential primaries: Issues and images in press releases and news coverage. *Journalism & mass communication quarterly*, 75(2), 312-324.
- Muller, B. (2022). Static Media Bias Chart. *Ad Fontes Media*. Retrieved April 24, 2022, from <https://adfontesmedia.com/static-mbc/>.
- Murdock, G., & Golding, P. (1977). Beyond monopoly: Mass communications in an age of conglomerates. In *Trade unions and the media* (pp. 93-117). Palgrave Macmillan, London.
- Norris, P. (1995). The restless searchlight: Network news framing of the post-Cold War world. *Political communication*, 12(4), 357-370.

- Norris, P. (2009). *Public Sentinel: News media and governance reform*. World Bank Publications.
- Pan, P. L., Meng, J., & Zhou, S. (2010). Morality or equality? Ideological framing in news coverage of gay marriage legitimization. *The Social Science Journal*, 47(3), 630-645.
- Powers, A., & Fico, F. (1994). Influences on use of sources at large US newspapers. *Newspaper Research Journal*, 15(4), 87-97.
- Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software*, 91, 1-40.
- Robinson, P. (2005). *The CNN effect: The myth of news, foreign policy and intervention*. Routledge.
- Ross, R. M., Rand, D. G., & Pennycook, G. (2021). Beyond "fake news": Analytic thinking and the detection of false and hyperpartisan news headlines. *Judgment & Decision Making*, 16(2).
- Scheufele, D. A. (1999). Framing as a theory of media effects. *Journal of communication*, 49(1), 103-122.
- Scheufele, B. (2004). Framing-effects approach: A theoretical and methodological critique.
- Schuck, A. R., & De Vreese, C. H. (2006). Between risk and opportunity: News framing and its effects on public support for EU enlargement. *European Journal of Communication*, 21(1), 5-32.
- Schuck, A. R., & de Vreese, C. H. (2009). Reversed mobilization in referendum campaigns: How positive news framing can mobilize the skeptics. *The International Journal of Press/Politics*, 14(1), 40-66.
- Semetko, H. A., & Valkenburg, P. M. (2000). Framing European politics: A content analysis of press and television news. *Journal of communication*, 50(2), 93-109.
- Shoemaker, P. J., & Reese, S. D. (1996). *Mediating the message* (pp. 781-795). White Plains, NY: Longman.
- Silver, L., Devlin, K., & Huang, C. (2019). US views of China turn sharply negative amid trade tensions. *Pew Research Center*, 13.
- Silver, L., Devlin, K., & Huang, C. (2020a). Republicans see China more negatively than Democrats, even as criticism rises in both parties. *Pew Research Center*.
- Silver, L., Devlin, K., & Huang, C. (2020b). People around the globe are divided in their opinions of China. *Pew Research Center*; December 5, 2019.
- Stone, G. C., & Xiao, Z. (2007). Anointing a new enemy: The rise of anti-China coverage after the USSR's demise. *International Communication Gazette*, 69(1), 91-108.

- Strömbäck, J., & Van Aelst, P. (2010). Exploring some antecedents of the media's framing of election news: A comparison of Swedish and Belgian election news. *The International Journal of Press/Politics*, 15(1), 41-59.
- Sutter, R. G. (2017). US-China relations: Perilous past, uncertain present. Rowman & Littlefield.
- Svensson, M. (2002). Debating human rights in China: A conceptual and political history. Rowman & Littlefield Pub Incorporated.
- Tefft, S. (1993). In Beijing, communism besieged. *Media Studies Journal*, 7(3), 59-64.
- Tewksbury, D., & Scheufele, D. A. (2009). News framing theory and research. In *Media effects* (pp.33-49). Routledge.
- Tuchman, G. (1978). Making news: A study in the construction of reality.
- Van Dijk, T. A. (1985). Structures of news in the press. *Discourse and communication: New approaches to the analysis of mass media discourse and communication*, 10, 69.
- Van Gorp, B. (2007). The constructionist approach to framing: Bringing culture back in. *Journal of communication*, 57(1), 60-78.
- Walter, D., & Ophir, Y. (2021). Strategy Framing in News Coverage and Electoral Success: An Analysis of Topic Model Networks Approach. *Political Communication*, 38(6), 707-730.
- Wang, M. L. (1991). Who is dominating whose ideology? New York Times reporting on China. *Asian Journal of Communication*, 2(1), 51-69.
- Wang, S.J. (1992). 'Factors Influencing Cross-National News Treatment of a Critical National Event: A Comparative Study of Six Countries' Media Coverage of the 1989 Chinese Student Demonstrations, *Gazette* 49(3): 193-214.
- Xu, X.G. & Parsons, P. (1997). The Image-Making News Flow between China and the United States, *Mass Communication Review* 24(3-4): 57-70.
- Young, L., & Soroka, S. (2012). Affective news: The automated coding of sentiment in political texts. *Political Communication*, 29(2), 205-231.

## Appendix A



**Figure 3.**

Factiva (2022), *Dow Jones Factiva*, retrieved from <https://global.factiva.com/> Copyright © 2022 Factiva, Inc.

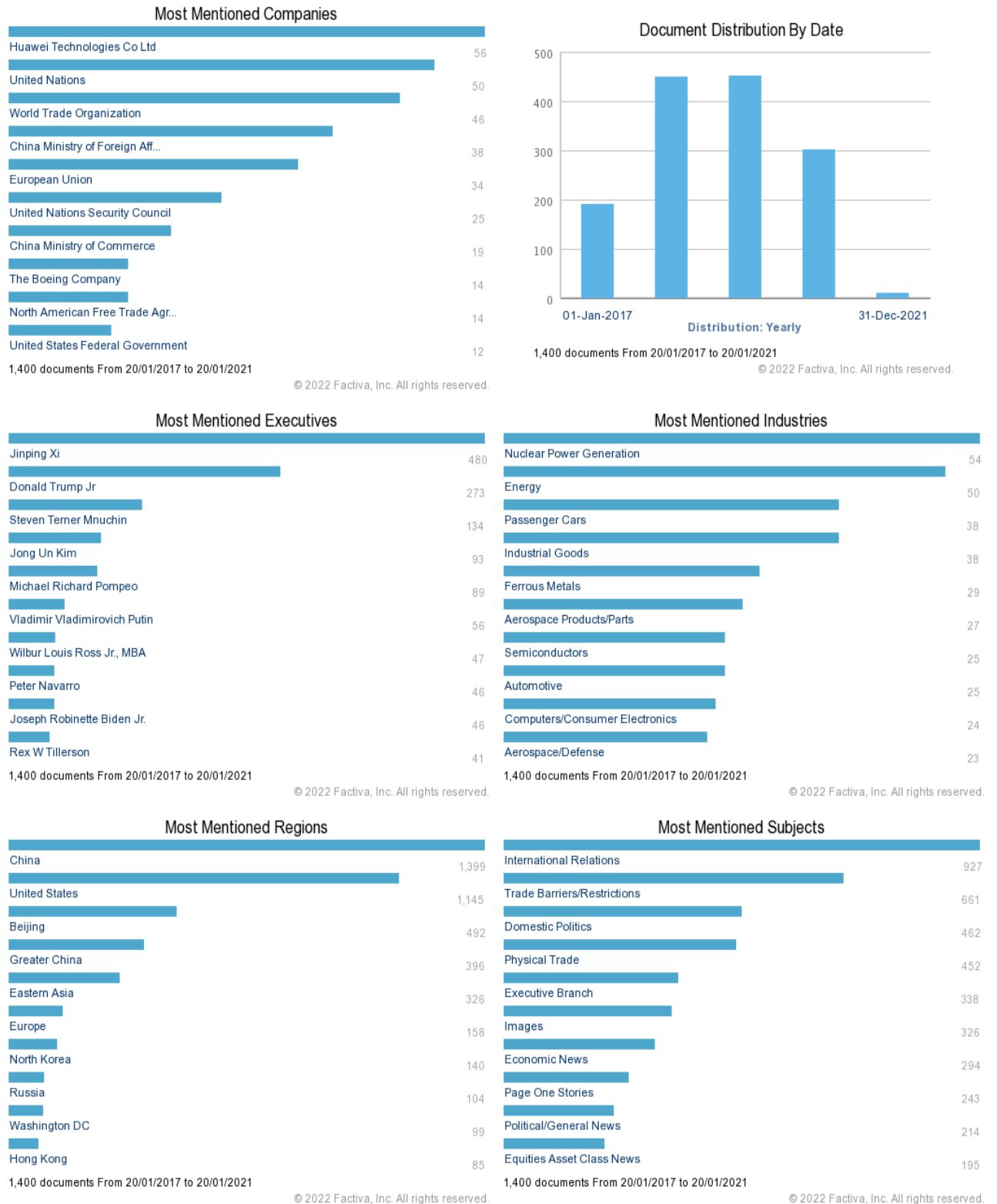


Figure 4.

Factiva (2022), *Dow Jones Factiva*, retrieved from <https://global.factiva.com/> Copyright © 2022 Factiva, Inc.