

From: Shiaw-Shiuan (Sharon) Chuang

Subject: News bias on 2020 presidential election in the US and Taiwan

Date: Mar.31.2023

Background

The 2020 United States presidential election was a significant event that was both hotly contested and polarizing (Time, 2020). The election exposed the deep political divisions that exist within the country, with many people holding strongly opposing views on various issues such as immigration, healthcare, the environment, and economic policy. Moreover, evidence indicates that media bias in news coverage influenced the election (CEPR, 2021).

Similarly, Taiwan was also experiencing a polarized political climate during the 2020 presidential election, which was strongly influenced by the 2019 protest in Hong Kong. Consequently, the values of democracy and the China challenge became crucial (Hass, 2022). The election was characterized by a charged atmosphere, with allegations of vote-buying and other forms of electoral malpractice, and exposed deep-seated tensions between pro-independence and pro-China forces within the country.

In light of the highly polarized political climates in both countries, it is essential to examine the role of news bias in shaping public opinion during the 2020 presidential election.

Although both the United States and Taiwan have experienced a rise in political polarization in recent years, there are differences in the nature and extent of polarization in the two countries. For example, Taiwan's polarization is largely driven by issues related to national identity and cross-strait relations with China, while the United States' polarization is more focused on issues related to race, immigration, and economic inequality (Clark C, Tan AC, Ho K, 2019). Therefore, understanding how bipolar political status quo and media bias interacted with each other to affect the election results is crucial.

Research Questions

My research aims to provide a more comprehensive understanding of news bias during the presidential election in the polarized societies of the United States and Taiwan, using quantitative methods. Specifically, the following research questions will be explored:

1. To what extent do conservative and liberal media outlets exhibit news media bias in their coverage of the presidential election? How does this bias compare to publicly funded non-profit news media?
2. What are the most likely biased keywords in news coverage of the presidential election?
3. Which country exhibits more media bias, Taiwan or the United States?

To address these research questions, classification models will be employed to verify if the difference in news bias among selected news outlets is significant enough to be classified by the models. Multiple models will be trained to determine if varying levels of news bias can be distinguished based on the presence of specific keywords.

Through this research, the study seeks to uncover the extent to which news outlets in both countries exhibited bias in their coverage of the election and the impact that this bias had on public perceptions of the candidates and issues at stake. By analyzing news coverage from a variety of sources, including traditional news outlets, social media platforms, and online forums, the study aims to provide insights into the complex ways in which news bias operates in today's highly polarized media landscape.

Data

For this project, the ideal dataset for classifying news from different media would consist of news-based data with a labeled source. The data should include variables such as the article's title, release date (to filter the period before the election), and category (to subset to specific types of news). However, one limitation of using a pre-existing dataset is that the labeling may not be consistent across sources or may not capture all relevant information.

To address this, the project will use two separate datasets - one for the United States and one for Taiwan. For the United States, the project will use the "NELA-GT-2020" dataset from Harvard Dataverse (Horne, Benjamin; Gruppi, Mauricio, 2021). The dataset contains nearly 1.8M news articles from 519 sources collected between January 1st, 2020 and December 31st, 2020. In the 2020 dataset, there is already a subset of election-related news, so this project will use election news from CNN, Foxnews, and PBS. There are 7663 observations from CNN, 3926 observations from Foxnews, and 2280 observations from PBS. To make the classification model balanced, the project will randomly sample 2000 observations from each of these three news sources. However, one limitation of using this dataset is that it only includes news articles from three sources, which may not be representative of all news sources in the United States. Additionally, the dataset only includes news articles from the year 2020, so it may not capture the full scope of news coverage leading up to the election.

For Taiwan's news, the project will use a dataset generated through web scraping techniques. The project scraped news articles from three media sources in Taiwan: Liberty Times (pro-independence), ctinews (pro-China), and Central News Agency. The project scraped all political news that can be searched with keywords including “民進黨” (DPP, the Democratic Progressive Party, the pro-independence party), “蔡英文” (Tsai Ing-wen, Incumbent president and Candidate of the DPP), “國民黨”(KMT, the Kuomintang, the pro-China party), and “韓國瑜”(Han Kuo-yu, the Candidate of the KMT). The project subset the news to only the political category and within the period of four months before the election until a week after the election. All of the three press sources have over 3000 observations. However, a limitation of using web scraping is that it may not capture all relevant news articles, especially if they are not easily searchable or if they are behind a paywall. Additionally, there may be biases in the articles selected for scraping or in the language used in the articles.

Bibliography

1. French, D. (2020, November 4). *Polarization is only clear winner of the 2020 election*. Time. Retrieved April 2, 2023, from <https://time.com/5907318/polarization-2020-election/>
2. Boxell, L. (2021, October 13). Bias in news coverage during the 2016 US election: New evidence from images. CEPR. Retrieved April 2, 2023, from <https://cepr.org/voxeu/columns/bias-news-coverage-during-2016-us-election-new-evidence-images>
3. Hass, R. (2022, March 9). *Democracy, the China Challenge, and the 2020 elections in Taiwan*. Brookings. Retrieved April 2, 2023, from <https://www.brookings.edu/opinions/democracy-the-china-challenge-and-the-2020-elections-in-taiwan/>
4. Clark C, Tan AC, Ho K (2019). Political Polarization in Taiwan and the United States: A Research Puzzle. Taipei, Taiwan: 2019 TIGCR International Conference on "Political Polarization: Perspectives go Governance and Communication". 25/10/2019-25/10/2019.
5. Horne, Benjamin; Gruppi, Mauricio, 2021, "NELA-GT-2020", <https://doi.org/10.7910/DVN/CHMUYZ>, Harvard Dataverse, V3, UNF:6:kDCP0JkQysbiElbatOjoqg== [fileUNF]