Sentiment and Inequality: A Statistical Analysis of Textual Data

Paul Rappoport[1]
Temple University and B&R Analytics
Aniruddha Banerjee
B&R Analytics

1. Introduction

1.1. Overview

Economists have recently, and enthusiastically, discovered text analysis (otherwise known as natural language processing or "NLP") and added it to the arsenal of econometric tools.[2] There is even a term and a hub, Sentometrics, which focuses on text mining, sentiment analysis, and econometrics.[3] That hub is home to all things NLP: references, data, software, and models. This new focus is not surprising given the appeal and growth of big data and data mining.

Economists have a long history of studying consumer behavior and analyzing how changes in attitudes alter the demand for products and services (Kahneman et al., 1999; Adams and Green, 1965).[4] The analysis of attitudes is prominent in econometric and social science research studies of risk, job selection, measuring attitudes towards inequality, analyzing public attitudes towards welfare policies, and measuring and tracking consumer sentiment. The data needed for tracking attitudes, confidence, and sentiment come, almost exclusively, from surveys. For example, the data used to construct both the Michigan Consumer Sentiment Index and the Conference Board's Consumer Confidence Index come from relatively small (500) monthly surveys.[5] But surveys are not without problems. There are issues of cost, sample selection, questionnaire design, and often, ambiguous measures of key concepts such as attitudes (Nayak et al., 2019; Choy, 2014).[6]

---

[1] The authors are Founding Principals at B&R Analytics. They wish to acknowledge and thank Mary Meehan, CEO of Metametrix™, and her technical team led by Adam Elliott and Sudheer Prem for their significant contributions to this work.

[2] See https://lt3.ugent.be/econlp/index.html and http://www.wikicfp.com/cfp/servlet/event.showcfp?eventid=143217&copyownerid=155568.

[3] See https://sentometrics-research.com; https://github.com/SentometricsResearch/sentometrics. The terms "text mining," "textual analysis," and "NLP" are used interchangeably in this paper.

[4] Kahneman, D., I. Ritvo, and D. Schkade, "Economic preferences or attitude expressions?: An analysis of dollar responses to public issues," Journal of Risk and Uncertainty, 19:1-3, 1999, 202-215; and Adams, F.G., and E. Green, "Explaining and predicting aggregate consumer attitudes," International Economic Review, 6()1, 1965.

[5] See http://www.sca.isr.umich.edu and https://www.conference-board.org/topics/consumer-confidence.

[6] Nayak, M., S. Durga Prasad, and K.A. Narayan," Strengths and weaknesses of online surveys," Technology, 6, 2019; and Choy, L.T., "The strengths and weaknesses of research methodology: Comparison and complimentary between qualitative and quantitative approaches." IOSR Journal of Humanities and Social Science, 19(4), 2014, 99-104.

How does text mining enhance economic analysis? Text mining offers the ability to analyze large numbers of documents without the constraints surrounding the design and implementation of surveys. Further, surveys are "reactive" in that respondents to a survey answer specific questions, whereas the data derived from text mining are unsolicited expressions of information, basic news reporting, or opinions. With increasingly more precise sentiment measurement engines available, text analysis has become a complement to traditional survey analysis.

The motivation for this paper came from the following questions: Did the Covid-19 pandemic change people's views of, and attitudes towards, inequality? Has the country become more divided (even polarized) or more insensitive? One option for answering these questions would be to conduct a survey. However, aside from cost considerations, the problem with a survey approach could be timing. Could respondents even answer what their attitudes were at different points in time? Given the need to rely on accurate recall, that would be improbable. And, is the term "inequality" itself too broad? From a problem-solving perspective, if a survey is considered impractical for answering these questions, then could text mining be a better approach? This paper reports on research that uses text mining to collect and organize text-based information to make inferences about trends that, arguably, relate to inequality.[7]

1.2. Text Mining and News

The framework for a textual analysis of inequality comes from the recent literature on using text mining to measure news sentiment. A recent paper by Shapiro et al. (2022) used text mining and sentiment analysis to create alternatives to the Michigan Consumer Sentiment Index and the Conference Board's Index of Consumer Confidence.[8] That paper showed that text mining could be used to model and track sentiment (confidence) over time without requiring surveys.

Text data are numerous, varied, and (generally) readily accessible. Text mining tools are readily available and are getting better.[9] With the tools available and the cost savings inherent in text modeling, text-based analysis of economic activity is becoming more prominent (Hoberg and Phillips, 2016; Song and Shin, 2019; NBER, 2020).[10]

---

[7] The text mining system used in this study was developed by Metametrix™ (MMX), a leader in textual analysis. See https://metametrixdata.com.

[8] Shapiro, A., M. Sudhof, and D. Wilson, "Measuring news sentiment," Journal of Econometrics, 228(2), 2022, 221-243.

[9] See https://sentometrics-research.com; and https://www.frbsf.org/economic-research/wp-content/uploads/sites/4/wp2017-01.pdf.

[10] Hoberg, G. and G. Phillips "Tested-based network industries and endogenous production differentiation," Journal of Political Economy, 124(5), 2016, 1423-1465; Song, M. and K. Shin, "Forecasting economic indicators using a consumer sentiment index: survey versus text-based data," Journal of Forecasting, 38(6), 2019, 504-518; and National Bureau of Economic Research, "Measuring the effect of firm uncertainty on economic activity: New evidence from one million documents," 2020, No. w27896.

1.3. Inequality

As noted, the objective of this paper is to use textual analysis to track attitudes over time towards inequality, as expressed in major national newspapers. Inequality is a term loosely applied to social, economic, or political situations in which unequal or divergent opportunities exist, or outcomes arise, for different groups or populations. Inequality is a broad concept that can pertain, for example, to unequal access to employment opportunities, wealth, housing, and general acceptance.[11] In some circumstances, inequality can be measured objectively. For example, economic inequality is often measured using the Gini Index, which measures the dispersion of income on a scale of 0 (everyone has the same income) to 1 (one person has all the income). Moreover, the Gini Index when used in conjunction with demographics, leads to deeper analyses of the dispersion of income, e.g., regarding the inequality associated with race, college education, and gender. The Gini index can be computed over time, thus allowing for trends in economic inequality to be tracked and analyzed. But, what about measuring inequality in other contexts, e.g., social, political, environmental, technological, etc.?

Assume that a random set of documents can be collected from multiple sources and then queried for the term "inequality." That is, only documents that touch on the topic of inequality are selected. These documents, however, may each cover only some of the many dimensions of inequality. Inequality itself cannot be meaningfully measured. However, central to all text mining processes is the measurement of sentiment. Basically, sentiment analysis involves the scoring of words, phrases, sentences, and documents on a continuum ranging from extremely negative sentiment to extremely positive sentiment. Neutral sentiment (neither positive nor negative) is represented by a sentiment score of zero on this scale. Thus, each document can have a sentiment score.[12] But, this kind of scoring must be context- and concept-specific. If the purpose is to score documents on what they say about inequality, then sentiment scoring becomes a gauge of perceptions, attitudes, beliefs, and opinions expressed about inequality in those documents.[13]

In this paper, we explore whether a measurable relationship can be detected between sentiment and inequality. For example, if a document with inequality as the concept is scored as highly negative (or highly positive) in terms of sentiment, what does that imply about inequality? We postulate that, in this example, a sentiment score is a reflection of attitudes, beliefs, and opinions about inequality, and that, over time, a change in sentiment score is likely correlated with changes in underlying attitudes towards inequality.

This may be a novel use of text mining, but it would not be sufficient to simply calculate the sentiment scores or their changes over time. When sentiment scores change or display unexpected trends over time, the real value of sentiment analysis lies in being able to relate or align those changes to specific real-world events that have a bearing on the concept in question (here,

[11] See https://dictionary.cambridge.org/us/dictionary/english/inequality.

[12] A document's sentiment score is built up from the sentiment scores of sentences within it which, in turn, are built up from sentiment scores of words or phrases within them.

[13] By making sentiment analysis context or concept-specific, sentiment can "proxy" for <u>any</u> context or concept. Inequality would be one such, but consumer confidence or demand for a product could also serve as contexts or concepts.

inequality). The rationale, of course, is that those events affect attitudes, beliefs, and opinions which, in turn, drive changes in sentiment.

To summarize, we use sentiment scores to measure the intensity of attitudes, beliefs, and opinions towards a specific concept, namely, inequality. For this, sentiment scores are derived from text obtained from several documents about inequality. We postulate that higher sentiment scores over time may indicate greater optimism towards the state of inequality, while lower sentiment scores over time may signify the opposite.

1.4. Sentiment

The measurement of sentiment using text mining is the starting point for assessing popular attitudes, beliefs, and opinions (and changes in them) regarding any concept. For Shapiro et al. (2022), the concept was news whereas, in this paper, it is inequality.

There are many approaches to measuring sentiment.[14] Sentiment analysis is applied to words, collections of words (n-grams), sentences, and paragraphs, and sentiment scores are computed ranging from extremely negative to extremely positive,[15] not unlike how scores are derived using Likert-type scales in traditional surveys. Sentiment engines require a lexicon of words and phrases, words that amplify or measure intensity, and other rules for scoring. Formal methods of sentiment analysis are covered in the review paper by Chan et al. (2021).[16]

The analysis of sentiment depends on the domain of what is being studied since specific words and phrases may be unique to that domain. Economists often look at domains such as financial reports, product reviews, SEC filings, and news. The domain for the analysis of sentiment in this paper with respect to inequality is defined by three national newspapers and the Web.

An even richer set of results from sentiment analysis can be obtained by delving deeper into possible drivers of sentiment. In this paper, we take that next step by using econometric models to establish how various measures of sentiment relate to drivers such as the passage of time, the document sources themselves, various latent factors like values and emotions, and categorical variables representing different types of inequality (namely, economic, political, and social). This further step yields a more nuanced understanding of perceptions of inequality by going beyond simply what the sentiment ratings are, but also why.

Our results show that sentiment scores change over time and that changes in sentiment scores appear to be aligned to specific events that have a bearing on inequality in one form or another. As expected, they also provide measurable impacts on sentiment scores of multiple hypothesized drivers.

---

[14] See https://blog.hubspot.com/service/sentiment-analysis-tools, https://getthematic.com/sentiment-analysis/.

[15] A variety of scales is possible: from the ternary (positive, neutral, negative) to more elaborate multi-point scales.

[16] See https://computationalcommunication.org/ccr/article/view/40.

1.5. Text-based Analysis of Inequality

We focus on the general term "inequality" rather than more descriptive terms such as "income or wealth inequality," "gender inequality," "racial inequality," "inequality in health outcomes," or "Covid-19 and inequality." This lack of focus is deliberate. First, our goal is to test whether text analysis can accommodate the broadest view of inequality. Secondly, we do not want to limit the analysis to a specific view of inequality.

1.6. MMX Engine

The MMX engine enables the assignment of documents to one of five categories: Society, Technology, Environmental, Economy, and Politics (STEEP). We use STEEP to classify inequality to provide a broader view of inequality.

The MMX process starts with a query. There may be multiple, embedded, or complete queries. In this instance, the subject of the query is "Inequality." We run this query to generate a large number of documents, collected and organized by source. The sources for the query are limited to the general Web and three national newspapers, namely, the Washington Post, the Wall Street Journal, and the New York Times. As is shown below, source is a key component of the process.[17]

The analysis process first generates summary data from the documents retrieved by the query from a variety of sources,. The MMX engine computes, organizes, and summarizes the outputs of the query. The summary data are contained in a spreadsheet with a number of categorical variables such as source, time, concepts, values, emotions, and STEEP categories. The MMX engine computes sentiment at both the sentence and the document level. The outcome measure is, therefore, a sentiment score.

[17] All sources have their own selective perspectives or biases. Therefore, it is imperative that sources be chosen to balance the different perspectives.

2. Composite Sentiment Scores

2.1. Measurement of Sentiment

As noted, sentiment is at the heart of the assessment. At first, the MMX analysis creates a sentiment score by sentence using a variant of a process called VADER (developed by Hutto and Gilbert, 2015).[18] The VADER process, adapted by MMX, uses a social research-oriented lexicon to measure sentiment by word, sentence, and document. MMX takes the sentence score which ranges from -4 (extreme negative sentiment) to +4 (extreme positive sentiment) and creates 9 bins (quartiles for sentiment between 0 and -4 and quartiles for the range 0 to +4. The zero value has its own bin in the middle of the range.

Next, the sentence-specific sentiment scores (on the nine-point scale) are grouped by bin to produce a frequency distribution for the scores. This is a precursor for two subsequent steps in the analysis.

1. By averaging the sentence-specific sentiment scores (over all bins) within each document, the analysis produces document-specific sentiment scores and to a composite measure called the Weighted Sentiment Score ("WSS").[19] This score and the averaging method are described in detail below.
2. By isolating extreme positive and negative scores (in particular, the bins for them), the analysis proceeds to the construction of a measure called the Extreme Sentiment Score ("ESS"), described below.

The construction of bins for the (-4, +4) range of sentiment scores achieves two objectives. First, the WSS constructed from those bins provides a realistic reading of the overall trend in sentiment. When a variety of sources is used to generate the documents (analogous to observations in balanced survey samples), a WSS is most useful when the sources themselves have no particular bias or, even if they do, an adequate mix of sources with different orientations is used to ensure some semblance of balance.

Second, the ESS can best identify and analyze changes in the extremes of the sentiment scale. The extremes of sentiment may result from responses to outliers or unusual events. In this instance, spreads between sentiment scores at the extremes can be interesting because they offer insight into

[18] Hutto, C.J. and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media, January 2015. See also https://www.analyticsvidhya.com/blog/2021/06/vader-for-sentiment-analysis/.

[19] Even though sentence-specific sentiment scores are integer-valued within the (-4, +4) scale, the document-specific average sentiment scores are unlikely to be so because of the averaging of a number of integer-valued scores. As a result, the (-4, +4) scale would no longer be represented by just the integers from -4 to 4 (with zero in the middle). Rather, a document-specific composite score like the WSS would have fractional values, either negative or positive. Then all WSS with values between -4 and -3 would fall into Bin 1, those between -3 and -2 would fall into Bin 2, and so on until those between 3 and 4 would fall into Bin 9. The midpoints of these segments, i.e., -3.5, -2.5, -1.5, -0.5, 0, 0.5, 1.5, 2.5, 3.5 will act as bin values, although Bin 5 with a "midpoint" value of 0 would be narrower than the other eight bins. This is not satisfactory because all bins should be of the same size.

events that most influence attitudes (positive or negative) towards inequality. Another way to understand the ESS is that it indicates the degree of polarization within a setting (country, market, etc.). When the ESS is positive, the extreme positive sentiment exceeds the extreme negative sentiment, signaling that those who feel strongly about an issue (like inequality) skew to a generally positive outlook. The opposite is true when the extreme negative sentiment exceeds the extreme positive sentiment. In both instances, the outlook gets stronger as the ESS (whether positive or negative) becomes larger. However, when ESS is at or near zero, extreme positive and negative sentiments offset each other, reflecting near or complete polarization between those expressing extreme sentiments.[20] We use the term "polarization" here as meaning an equal, or near-equal, division between those with the strongest beliefs or attitudes, both positive and negative.

2.1.1. Construction of the WSS

As noted earlier, the WSS is a composite sentiment score that may be computed from sentence-specific sentiment scores grouped by the nine bins. Such a composite is constructed by using the relative frequencies of the bins as weights. Because the bins span four negative levels as well as four positive levels, that composite can easily be negative for some documents and positive for others. Still, this is a perfectly valid method of construction of a weighted average and may be used to depict overall sentiment from all sources utilized for the query.

An alternative method of constructing the WSS may, however, be preferable for the statistical analysis we conduct subsequently. This method starts with mapping the (-4,+4) range for sentiment scores into the unit interval, i.e., the (0,1) range. This rescaling can certainly change computed WSS but not so the fundamental insights derived from it. Moreover, a metric defined over the unit interval has some desirable properties. For example, it shares well-known attributes of commonly used metrics like percentages, shares, or probabilities. Negative composite scores that would emerge from the (-4,+4) scale would now correspond to scores on the lower end of the (0,1) scale. In essence, the WSSs from using the (0,1) scale are more compressed (but strictly non-negative) than those from the original (-4,+4) scale. The greatest benefit comes in the form of easier econometric modeling and interpretation of relationships between WSS and various factors that plausibly "drive" it. We return to this issue with an empirical example later in the paper.

We start with the intent of mapping the nine bins for the original (-4,+4) scale into corresponding nine bins for the new (0,1) scale. That is, we divide the unit interval into nine equally-spaced segments or bins. Thus, the first such bin is the segment (0, 0.1111), the second bin is the segment (0.1111, 0.2222), and so on until the ninth bin over the segment (0.8888, 1). For purposes of constructing a weighted average, we take the midpoints in these nine bins, namely, 0.0556, 0.1667, …, 0.5, … 0.8333, 0.9444. These midpoints are proxies for the segments that constitute the nine bins.[21]

[20] A similar conclusion may be reached when WSS, the other composite sentiment measure, is at or near zero. The difference is that, while WSS is a composite of all sentiment ratings up and down the scale, ESS reflects only the frequencies of those at the extreme ends of the sentiment scale.

[21] The midpoint 0.5 of the (0,1) interval corresponds to 0, the midpoint of the (-4,+4) interval.

2.1.2. Construction of the ESS

To compute ESS, we first isolate the frequencies at the document level associated with extreme sentiment ratings, i.e., bin1 and bin 9. These correspond to sentiment ratings of -4 and 4 on the original scale, now translated into the bins that contain zero and one, respectively, on the converted scale.

The difference of the frequencies within these bins (bin 9 minus bin 1) is then normalized by dividing that difference by the sum of frequencies across all nine bins. The result, which can be either negative or positive, or even zero, is the computed value of ESS. On the converted (-1,+1) scale, it is, essentially, the percentage of all documents with sentiment at the highest end of the scale less the percentage of all documents with sentiments at the lowest end of the scale. As noted earlier, on the (-1,+1) scale, ESS reflects the greatest polarization of beliefs and attitudes when it is at, or close, to zero.

For statistical modeling of ESS (see section 4), one more transformation is useful. This transformation maps ESS values from the (-1,+1) scale to the (0,1) scale, as was done for WSS. The transformation is

Rescaled ESS = (ESS + 1) / 2

With the range of ESS now changed to the interval (0,1), the interpretation of values within that interval changes as well. All negative values (up to -1) of the original ESS now correspond to values between 0 and 0.5 for the rescaled ESS. ESS values in the range (0,0.5) thus signify that extreme negative sentiment exceeds extreme positive sentiment. Similarly, all positive values (up to 1) of the original ESS now correspond to values between 0.5 and 1 for the rescaled ESS. Moreover, values within the latter range signify that extreme positive sentiment exceeds extreme negative sentiment. When the rescaled ESS has a value of exactly 0.5, corresponding to 0 on the original ESS scale, extreme positive and negative sentiments offset, and perfect polarization occurs.[22]

2.2. Sources and STEEP Categories

The results from text mining depend on the source(s) of the text. Text selection here is somewhat analogous to sample selection in the traditional survey framework. An analysis that focuses only on social media sources such as Twitter or Facebook is likely to yield different solutions than an analysis that focuses on national sources such as the New York Times or Wall Street Journal or documents pulled from searching the Web. Source bias is a real possibility.

We employ four sources in this study for extracting documents: three major national newspapers (namely, the New York Times, the Wall Street Journal, and the Washington Post) and the Public Web. We randomly select articles/documents from these sources during the 2017–2021 study

[22] In what follows, all references to ESS mean the rescaled version of ESS within the (0.1) interval.

period, assembling a total of 12,361 documents. We do not, however, place a quota or a limit on the number of documents obtained from each source.

Table 1. Number of Documents, by Source[23]

Source	Documents (2017–2021)
New York Times ("NYT")	3,818
Wall Street Journal ("WSJ")	2,198
Washington Post ("WP")	5,537
Web ("WEB")	808

Of the five STEEP categories to which the MMX engine assigns documents, we focus on only three that are, arguably, the most likely to be linked to inequality, namely, Economy, Politics, and Society. We look at the trend in sentiment scores for three categories. The classification of a document into a STEEP category is based on the analysis of keywords in that document.

Table 2. Number and Percent of Documents, by Source and STEEP Category, 2017-2021

Source	Category	Percent distribution by STEEP	Documents (2017-2021)
NYT	Economy	19.5%	723
	Politics	33.6%	1,245
	Society	46.9%	1,740
WEB	Economy	54.8%	418
	Politics	6.7%	51
	Society	38.5%	294
WP	Economy	24.6%	1,310
	Politics	42.5%	2,260
	Society	32.9%	1,747
WSJ	Economy	40.2%	878
	Politics	27.1%	592
	Society	32.8%	717

The STEEP orientation of documents varies by source. Almost 47% of documents from the New York Times are assigned to Society, over 42% of those from the Washington Post are assigned to Politics, over 40% of those from the Wall Street Journal are assigned to Economy, and the predominant STEEP category for the Public Web is Economy. This ensures a balanced representation of the three STEEP categories across the four sources, in particular, the national newspapers.[24]

[23] Since the number of documents varies by source, a test of any size effect can be done by randomly resampling the sources to equalize the number of documents.

[24] More research is required to test the effects of this type of assignment of documents to the STEEP categories.

2.3. Dates and Time Stamps

A key part of the research plan is to track sentiment over time. This adds another requirement to the selection of sources, that every document must contain an appropriate date field (referring to the date the article was first published). For newspapers, this is not a major issue. If, on the other hand, the source is the Web (Google or Bing) then finding the first publication date can be a challenge. MMX has created a process that addresses this issue.

2.4. Trends in Sentiment Scores

In the trend analyses that follow, we use the two composite measures of sentiment defined earlier, namely, WSS and ESS. The objective is to depict trends in these two composite scores over the period of study.

2.4.1. Trends in WSS

To depict the trend in WSS, we compute the median (document-level) quarterly median WSS for 20 quarters over the study period, Q1 2017-Q4 2021. We then display the quarterly median WSS at successive disaggregations: first, the overall median WSS, then the median WSS by source, and finally the median WSS by three of the STEEP categories (namely, Economy, Politics, and Society).

Figure 1 displays the cubic trend in the overall median WSS over the study period.

Figure 1. Trend in Overall Median WSS, Q1 2017–Q4 2021

Figure 1 shows three distinct phases for the overall median WSS: rising in early 2017 to mid-2018, steady to falling slowly from mid-2018 to late 2020, and rising gradually again from late 2020 to late 2021. If overall attitudes and opinions towards inequality are reflected by WSS, then the rise in WSS during the first phase could signify a more favorable perception of the state of inequality. In the second phase, WSS suggests that perceptions about inequality may have held steady briefly but eventually declined, most significantly in the second half of 2020.[25] In the final phase, however, the WSS signals a recovery of the perception of inequality towards the level seen in 2017.[26]

To understand why these trends in the WSS emerged, it is necessary to look for significant events that preceded or coincided with them and could have caused changes in perceptions about inequality, even with a lag. We surmise that the WSS may be more of a lagging, than a coincident, indicator, in that its behavior (level, change of direction, etc.) is more likely to reflect past, rather than purely contemporaneous, events. Attitudes and perceptions, particularly at the level of a whole nation, do not change overnight. Rather, change comes gradually and over time.

The first phase was marked by the start of the Trump administration and the passage of significant tax cuts that were hailed as a boost to the US economy and projected to raise more in new revenue than the revenue forgone from the tax cuts. That optimism got tamped down in early to mid-2018 as the tax cuts became widely seen as being more favorable to the highest income classes, stock market investors, and corporations than to others. The federal debt-to-GDP ratio rose, raising the prospect of greater burdens on future generations.[27] Moreover, whatever optimism towards inequality was being created on the economic front soon began to be countered by the turmoil brought about by the Trump administration's immigration policies, large women's marches over gender inequality, nascent administration scandals (involving possible Russian interference in the US elections, marking the start of the Robert Mueller investigation), the beginning of the #MeToo movement, and the white nationalist rally in Charlottesville, Virginia. Each event, in its own way, may have strained popular perceptions of equality of different kinds and marked the beginning of the slump in the median WSS in 2018.[28]

Perceptions of inequality in the United States seem to have reached their nadir in mid-2020 following a series of political and social events or mishaps that weighed heavily on the national psyche. There was mixed news on many fronts, but the cumulative effect on attitudes towards inequality may have turned negative. Both Democrats and Republicans made electoral gains in

[25] For an alternative interpretation, note that the overall median WSS is below 0.5 (the half-way mark between 0 and 1) only in Q2 2020, signifying that, overall, negative sentiment surpassed positive sentiment in that quarter. In every other quarter, overall, positive sentiment stayed ahead of negative sentiment, but by varying margins.

[26] Although not shown here, median WSS trends by source (i.e., NYT, WP, WSJ, and WEB) follow the same cubic pattern, but with a twist. Median WSSs for NYT and WSJ at first fall, then rise, and then fall again, while those for WP and WEB exhibit the precisely opposite pattern of cyclicality. Compared to the overall median WSS, each median WSS by source displays greater amplitude (variability from a reference point). But, the mixing of the four disparate document sources causes a more attenuated and milder cycle in the overall median WWS.

[27] Amadeo, K., "How much Trump's tax cuts cost the government," November 19, 2021, in https://www.thebalance.com/cost-of-trump-tax-cuts-4586645.

[28] Corey, D., "2017 year in review," NBC News, December 26, 2017, in. https://www.nbcnews.com/news/us-news/2017-year-review-here-are-top-10-biggest-news-stories-n828881.

the 2018 mid-term elections,[29] the #MeToo movement went global and some signature prosecutions focused greater attention on misconduct towards women,[30] and President Trump engaged in a tariff war with China to a mixed reception in the U.S.[31] In 2019, following the release of the Mueller report, President Trump was impeached, possibly energizing large swaths of his — and his policies' — opponents. At the same time, environmental and political protests in the US and abroad gave voice to long-ignored causes.[32]

In the final phase through early 2021, the sustained decline in WSS may reflect the cumulative negative impact of several unsettling events that raised the collective anxiety of the country and polarized it in the process. These included the arrival of various strains of the coronavirus in an unprepared nation lacking reliable vaccines amid the hawking of unproven remedies, a bruising campaign to overcome opposition to vaccines approved for emergency use, the U.S. Congress' failure to convict an impeached president, serious racial and social unrest following the police killings of George Floyd and other black citizens, the rise of Black Lives Matter, the rapid confirmation of Justice Amy Coney Barrett to a now solidly conservative U.S. Supreme Court which seemed to bode ill for the popular and longstanding *Roe v. Wade* decision of the Court on abortion rights, the flat refusal to accept the election of President Biden by his predecessor, and the armed insurrection at the U.S. Capitol on January 6, 2021.[33] As these events unfolded through this phase, the socio-political impact of the coronavirus was magnified many times over as the nation agonized over vaccinations, mask mandates, and social distancing.[34] Popular perceptions of inequality continued to worsen in 2021 as the former President was again acquitted by the U.S. Congress after being impeached for a second time over the January 6 riots. Voter fraud and "stolen" election claims ran rampant in many parts of the nation, many states passed new laws to further restrict voting rights, and the Covid-19 death toll nationwide surpassed 800,000.

Figure 2 shows the median quarterly WSS by our four sources (three major national newspapers and the public Web) over the study period.

---

[29] Lindsay, J.M., "Ten Most Significant World Events in 2018," Council on Foreign Relations, December20, 2018, in https://www.cfr.org/blog/ten-most-significant-world-events-2018.

[30] Pomarico, N., "11 of the biggest moments in the #MeToo movement in 2018," Insider, December 19, 2018, in https://www.insider.com/me-too-movement-moments-2018-12.

[31] See fn. 29.

[32] Warren, L., "21 news stories that gripped the world in 2019," Insider, December 20, 2019, in https://www.insider.com/news-stories-that-gripped-in-the-world-in-2019.

[33] Cowan, L., "The year in review: Top news stories of 2020 month-by-month," CBS News, December 27, 2020, in https://www.cbsnews.com/news/2020-the-year-in-review-top-news-stories-month-by-month/.

[34] Carothers, T. and A. O'Donohue, "Polarization and the pandemic," Carnegie Endowment for International Peace, April 28, 2020, in https://carnegieendowment.org/2020/04/28/polarization-and-pandemic-pub-81638; and Dimock, M. and R. Wike, "America is exceptional in its political divide," Trust Magazine (Pew Research Center), March 28, 2021, in https://www.pewtrusts.org/en/trust/archive/winter-2021/america-is-exceptional-in-its-political-divide.

Figure 2. Trend in Overall Median WSS, by Source, Q1 2017-Q4 2021

Some noteworthy features of Figure 2 are the following. First, over the study period, generally higher WSS (more optimism towards inequality) is associated with documents from the Wall Street Journal and the Public Web. In particular, the former source has the highest peaks in the first half of 2018 and in late 2021 as well. The Public Web has WSS peaks in late 2017-early 2018, early 2019, and late 2020.

Second, the overall median WSS for Washington Post documents remains generally lower throughout the period, actually dropping sharply in Q2 and Q3 of 2020 and, again, in Q3 2021.

Third, of all four sources, overall median WSS for documents from the New York Times stays relatively flat throughout the study period, signifying the least fluctuation in perceptions of inequality among the four sources.

Finally, each newspaper or, for that matter, any source has its own political, economic, and social orientation. Documents taken from such sources, particularly at times of upheaval on all those fronts, are quite likely to show marked variations in sentiment (as measured by WSS). This is evidence of two plausible conclusions. First, it is clear that source matters for measuring sentiment and there is no straightforward way to find the "best" source for that purpose. Second, an overall WSS based on diverse sources of documents and information is more likely to "reflect the mood" of the nation as opposed to source-specific WSS.

The variation in WSS trends across individual sources underscores the importance of assembling a judicious mix of sources for properly understanding perceptions and attitudes through textual data mining. For any multi-faceted topic like inequality, "source bias" is likely to be high and proportional to the differences in orientation among different sources of documents. Some more light is shed on this issue by the trends observed in Figure 3.

Figure 3. Trend in Overall Median WSS, by STEEP Category, Q1 2017-Q4 2021

Figure 3 shows distinct patterns in the overall median WSS when computed by the three STEEP categories: Economy, Politics, and Society. It is not hard to see that the greatest optimism towards inequality is consistently expressed about Economy (or, possibly, economic inequality). At the other end of the spectrum, optimism about Society (or, possibly, social inequality) is lowest among the three STEEP categories, with pronounced drops into pessimism or near-pessimism from time to time (particularly, mid-2020). Optimism towards Politics (or, possibly political inequality) remains in the middle of the three STEEP categories, flirting with pessimism in Q2 2020 and Q3 2021.

An argument can be made that document sources that are more invested in one of these STEEP categories, or systematically express greater optimism or pessimism about any one of those categories, are likely to have the greatest influence on the overall median WSS.[35] That is, text mining of the type pursued in this paper can reveal interesting insights about perceptions of amorphous and multifaceted concepts like inequality, but only if the mixed outcomes encountered in this study are properly anticipated and sources and sample sizes are selected judiciously.

If attitudes and opinions towards inequality are reflected by WSS, then the rise in WSS during the first phase could signify a more favorable perception of the state of inequality. In the second phase, WSS suggests that perceptions about inequality may have held steady briefly but eventually declined, most significantly in the second half of 2020.[36] In the final phase, however, the WSS signals a recovery of the perception of inequality towards the level seen in 2017.[37]

When WSS is compared across the three STEEP categories, there are discernible differences in trends. Over the entire study period, the WSS based on documents in the Politics category has a consistently lower value than the WSS based on documents in the Economy and Society categories, respectively. Any linear or low-order polynomial trend fitted to the WSS lines for the three STEEP categories would show that for Economy consistently on top, that for Society parallel to — but below — it, and that for Politics significantly below both. Interestingly, those trend lines would appear to converge towards the very end of the study period.

2.4.2. Trends in ESS

As for the trend in ESS, we also compute the median (document-level) quarterly median ESS for 20 quarters over the study period, Q1 2017-Q4 2021. We then display the quarterly median ESS at successive disaggregations: first, the overall median ESS, then the median ESS by source and, finally, the median ESS by the three STEEP categories (namely, Economy, Politics, and Society).

Figure 4 displays the cubic trend in the overall median ESS over the study period.

[35] To keep this paper from becoming tedious and unbearably long, we don't report median WSS by both source <u>and</u> STEEP category.

[36] For an alternative interpretation, note that the overall median WSS is below 0.5 (the half-way mark between 0 and 1) only in Q2 2020, signifying that, overall, negative sentiment surpasses positive sentiment in that quarter. In every other quarter, overall, positive sentiment stays ahead of negative sentiment, but by varying margins.

[37] Although not shown here, median WSS trends by source (i.e., NYT, WP, WSJ, and WEB) follow the same cubic pattern, but with a twist. Median WSSs for NYT and WSJ at first fall, then rise, and then fall again, while those for WP and WEB exhibit the precisely opposite pattern of cyclicality. Compared to the overall median WSS, each median WSS by source displays greater amplitude (variability from a reference point). But, the mixing of the four disparate document sources causes a more attenuated and milder cycle in the overall median WWS.

Figure 4. Trend in Overall Median ESS, Q1 2017-Q4 2021

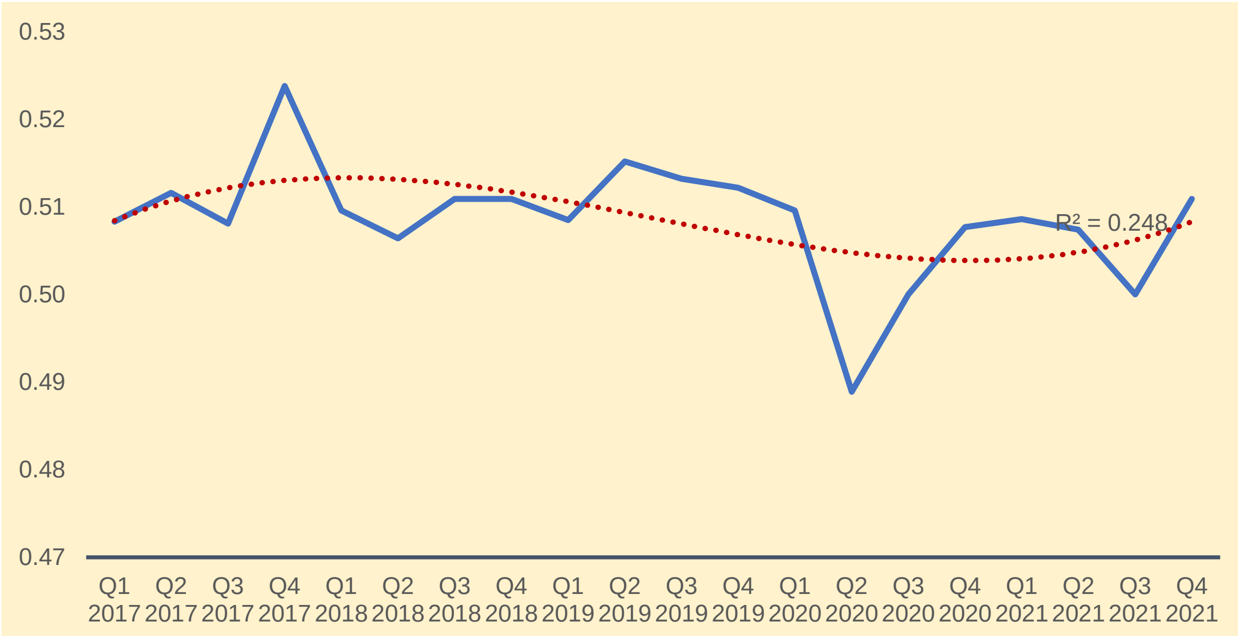


Figure 4 shows three distinct phases for the overall median ESS as well: rising in early 2017 to mid-2018, steady to falling slowly from mid-2018 to late 2020, and rising gradually again from late 2020 to late 2021. In that respect, the trend in ESS over the study period resembles the trend in WSS over that period.

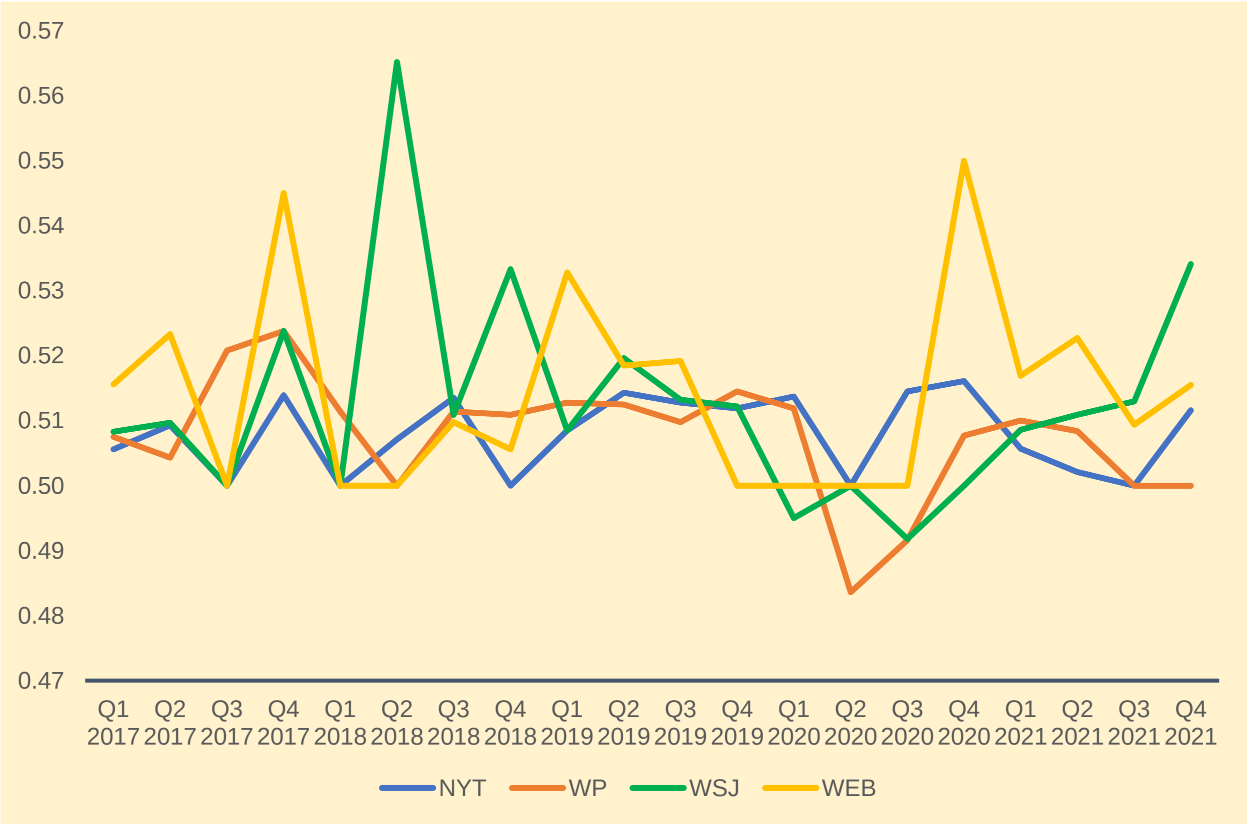
However, the trend in ESS better reveals what happened to polarization over inequality during this period. For the most part, particularly the first two-thirds of the study period, extreme positive sentiment towards inequality outweighed the extreme negative sentiment towards it. This suggests that, even among documents expressing the most extreme sentiments, a general optimism towards inequality prevailed during this sub-period. Shortly after Q1 2020, the trend changed markedly as the overall median ESS fell into “negative” territory (below 0.5), recovered slightly into “positive” territory (above 0.5) in late 2020 and early 2021, but then fell again to 0.5 in Q3 2021. This pattern of change in the overall median ESS supplements the findings from the overall median WSS in Figure 1.

For a while in early 2020, ESS turned sharply into negative territory. This is exactly the period in which general sentiment, depicted by the overall median WSS, also fell sharply. Thus, it was not merely the sentiments in the extreme ends of the scale that pulled in a polarizing or negative direction, but so also did sentiment up and down the scale. It can be concluded, therefore, that 2020 was a particularly volatile year for how inequality — in its many forms — was viewed in the United States. Arguably, this unusual pattern of behavior in the perception of inequality can be attributed to the cumulative effect of the contemporaneous or recent economic, political, and social upheavals of this period. It is noteworthy that the “stories” told by both WSS and ESS are mutually reinforcing. How “shocks” to the nation can produce significant changes — even if temporary —

in the national mood or viewpoint on any issue is clearly demonstrated by the joint use of these sentiment indicators.

When ESS is depicted by source, the trend in the overall median ESS looks similar to that in the overall median WSS in Figure 1.

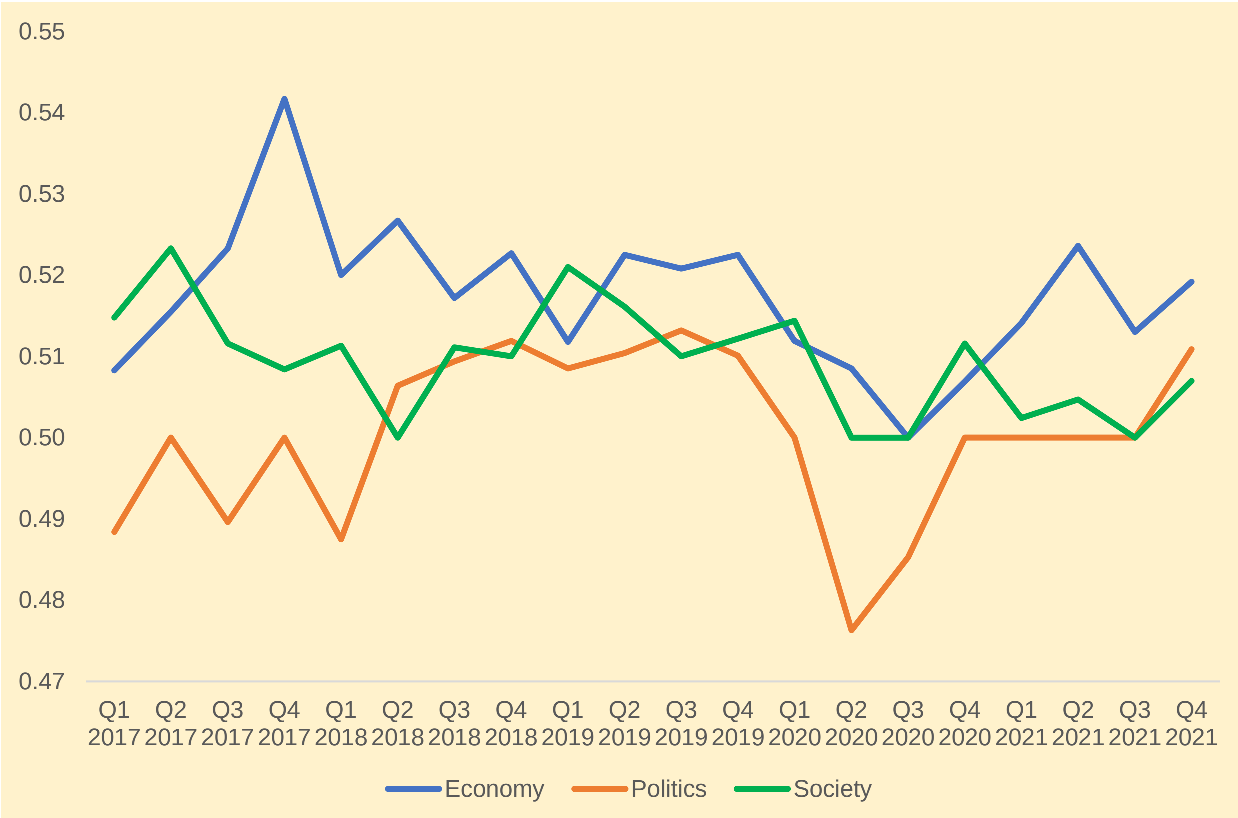
Figure 5. Trend in Overall Median ESS, by Source, Q1 2017-Q4 2021



Once again, ESS for the Wall Street Journal and the Public Web show the sharpest spikes in the positive direction at various times, the Washington Post has the sharpest drop in ESS below 0.5 in mid-2020, and the New York Times' trend is more on an even keel. Nevertheless, ESS for all four reach near-perfect or full polarization in the first half of 2018 and in Q2 2020. Individually, each source reaches such polarization several times over the study period, although not always at the same time. This finding underscores how divided the nation has been over inequality during the study period, no matter the differences of orientation among the sources.

When ESS is depicted by the three STEEP categories, once again the trend in the overall median ESS looks familiar.

Figure 6. Trend in Overall Median ESS, by STEEP Category, Q1 2017-Q4 2021



With a brief exception in Q3 2020, there is little polarization regarding Economy (or economic inequality), with even extreme sentiment staying positive almost throughout the study period.

3. Values and Emotions Analysis

MMX has the capability to determine associations between attitudes, beliefs, and opinions expressed in documents and pre-set libraries of common Values and Emotions. For example, the Values library has 123 Values like Access, Diversity, Empowerment, and Wealth, among others. Similarly, the Emotions library has 24 emotions, such as Admiration, Compassion, Hate, Trust, etc. When a query (on any topic) is run on MMX, the standard output includes the top 10 Values and the top 10 Emotions associated with the documents queried. These Values and Emotions provide significant insight into what drives the attitudes and perceptions expressed in the queried documents. These are the non-quantitative counterparts of sentiment scores in the analysis of documents. It is natural to ask not merely which Values and Emotions are most closely associated with the attitudes, beliefs, and opinions expressed in queried documents, but also how they shape (or are shaped by) the sentiments measured in them.

3.1. Values

MMX identifies key Values associated with the set of documents included in a query. For this, MMX uses a machine learning approach with logistic regression as a probabilistic classifier. After training the logistic regression model for every Value using a proprietary training data set, MMX uses the trained models to probabilistically classify documents by the different Values in its training model.

The 123 Values in MMX's Values library represent a set of behavior and opinion qualifiers/modifiers that shape attitudes and perceptions in the long run. They provide insight into actions and decisions. In this section, we explore, in particular, the relationship between Values and sentiment in the context of inequality. Table 3 lists the top 10 Values identified by a query about inequality in this study.

Table 3. Top Ten Values from Query about Inequality

Value	Description	Weight[38]
Access	The opportunity to obtain, use or benefit from something, such as information, services, information, membership, education, or elected officials	15.34%
Assistance	Helping a person or an entity with a job, task, or goal; contributing to the fulfillment of a need	13.03%
Balance	Establishing and maintaining equilibrium and harmony between demanding forces of life; equal and appropriate proportions of work, personal and home life	6.50%
Diversity	The mixture of people who are of different races or who have different cultural backgrounds in a group or organization	6.61%
Empowerment	To promote or provide a sense of power in others to accomplish goals, control their life, or claim their rights	17.95%

[38] Weight is the relative frequency of each Value in the Top 10 among all Top 10 Values detected over the period 2017-2021. The top 10 Values account for over 60% of all Values.

Equality	Having the same rights, social status, or opportunities as others	10.98%
Growth	The process of personal development; economic change in a company or business	8.74%
Value	The importance, worth, or usefulness of something and its relationship to price or cost	5.21%
Wealth	Abundance of valuable material possessions or resources	5.76%
Youth	The vigor, freshness, or immaturity associated with someone young	9.87%

How do Values, and changes in them, affect sentiment? While this raises the question of what drives changes in Values in the first place, it may be reasonable to postulate that they drive sentiment and not vice versa. This section tracks sentiment by Value and compares the resulting trends by key Values.

Table 4. WSS, by Top Ten Value and Time

Value	2017	2018	2019	2020	2021
Access	0.548	0.553	0.542	0.545	0.540
Assistance	0.551	0.525	0.540	0.504	0.538
Balance	0.562	0.623	0.543	0.535	0.546
Diversity	0.590	0.572	0.548	0.527	0.543
Empowerment	0.516	0.524	0.552	0.524	0.497
Equality	0.563	0.535	0.561	0.573	0.544
Growth	0.589	0.595	0.556	0.537	0.560
Value	0.504	0.561	0.554	0.511	0.570
Wealth	0.658	0.686	0.638	0.696	0.627
Youth	0.530	0.455	0.526	0.455	0.519
Overall WSS	**0.544**	**0.539**	**0.543**	**0.519**	**0.542**

The WSS scores in Table 4 represent WSS by Value. The Value **Wealth** stands out in that the WSS score for each year is consistently higher than the overall average WSS. To a lesser extent, the same is true for the value **Growth**. In contrast, the WSS associated with the Value **Youth** is consistently lower over time than the overall average WSS. A more rigorous analysis of this purported association is pursued in Section 4.

Table 5. ESS, by Top Ten Value and Time

Value	2017	2018	2019	2020	2021
Access	0.520	0.500	0.514	0.500	0.509
Assistance	0.515	0.500	0.513	0.492	0.513
Balance	0.519	0.538	0.500	0.500	0.513
Diversity	0.526	0.513	0.510	0.487	0.513
Empowerment	0.506	0.510	0.512	0.500	0.500
Equality	0.515	0.509	0.526	0.533	0.506

Growth	0.529	0.520	0.512	0.508	0.519
Value	0.508	0.509	0.511	0.500	0.524
Wealth	0.556	0.500	0.559	0.570	0.560
Youth	0.500	0.478	0.500	0.469	0.500
Overall ESS	**0.510**	**0.510**	**0.512**	**0.500**	**0.510**

3.2. Emotions

As with Values, the MMX engine also identifies key Emotions associated with the set of documents included in a query. To identify the Emotions, MMX uses a proprietary rule-based algorithm. The document to be analyzed is split into sentences, which are then individually analyzed and scored based on the strength of Emotion in each sentence using a proprietary lexicon created by MMX. Once the sentence-level Emotions are identified, the entire document's Emotion is scored using a probability score based on the number of sentences and the strength of Emotions in these sentences.

The 24 emotions in MMX's Emotions library represent short run responses to events or activities. In this section, we explore, in particular, the relationship between Emotions and sentiment in the context of inequality.

Table 6. Top Ten Emotions from Query about Inequality

Emotion	Description	Weight[39]
Admiration	A feeling of respect and appreciation or an object of esteem	15.34%
Anger	Intense emotional state of strong feelings of displeasure usually in response to some threat or provocation	13.03%
Compassion	Sympathetic consciousness of others distress, together with a desire to alleviate it; empathy and concern for others	6.50%
Disappointment	A defeated or deflated feeling when expectations or hopes are not fulfilled	6.61%
Guilt	The feeling of having done wrong; having committed an offense or violated a moral standard and deserving of blame	17.95%
Hate	An intense dislike, animosity or resentment directed at individuals, objects, ideas, or situations	10.98%
Pride	Satisfaction and contentment in one's or another's actions or choices	8.74%
Sadness	Afflicted with grief or unhappiness	5.21%
Trust	Reliance on the character, ability, strength, or truth of someone or something	5.76%
Worried	Troubled or concerned, or showing concern or anxiety about actual or potential problems	9.87%

[39] Weight is the relative frequency of each Emotion in the Top 10 among all Top 10 Emotions detected over the period 2017-2021. The top 10 Emotions account for over 76% of all Emotions.

How do Emotions, and changes in them, affect sentiment? As with Values earlier, we postulate that Emotions drive sentiment and not vice versa. This section tracks sentiment by Emotion and compares the resulting trends by key Emotions.

Table 7. WSS, by Top Ten Emotion and Time

Emotion	2017	2018	2019	2020	2021
Admiration	0.567	0.583	0.570	0.553	0.561
Anger	0.485	0.481	0.506	0.464	0.497
Compassion	0.576	0.563	0.560	0.545	0.579
Disappointment	0.505	0.523	0.553	0.543	0.519
Guilt	0.458	0.500	0.457	0.447	0.484
Hate	0.477	0.475	0.492	0.458	0.456
Pride	0.601	0.593	0.625	0.606	0.572
Sadness	0.554	0.491	0.512	0.504	0.510
Trust	0.535	0.557	0.552	0.551	0.552
Worried	0.513	0.553	0.532	0.522	0.497
Overall WSS	**0.544**	**0.539**	**0.543**	**0.519**	**0.542**

We note that, in Table 7, **Anger**, **Guilt** and **Hate** (all "negative" Emotions) have WSS scores that are consistently below the overall average WSS in every year. **Disappointment**, **Sadness**, and **Worried** — arguably also negative Emotions — all have WSS scores that are below the overall average WSS in most, if not all, of the years. In contrast, **Admiration**, **Compassion,** and **Pride** (all "positive" Emotions) have consistently larger WSS scores than the overall WSS average score over the five years. Also, so does **Trust** (another positive Emotion), except in 2017.

Table 8. ESS, by Top Ten Emotion and Time

Emotion	2017	2018	2019	2020	2021
Admiration	0.512	0.529	0.530	0.525	0.523
Anger	0.490	0.488	0.508	0.473	0.500
Compassion	0.530	0.513	0.516	0.510	0.529
Disappointment	0.500	0.511	0.512	0.500	0.500
Guilt	0.490	0.500	0.500	0.487	0.500
Hate	0.481	0.468	0.477	0.469	0.469
Pride	0.543	0.532	0.559	0.558	0.531
Sadness	0.500	0.486	0.500	0.500	0.500
Trust	0.507	0.521	0.515	0.518	0.513
Worried	0.508	0.500	0.500	0.500	0.513
Overall ESS	**0.510**	**0.510**	**0.512**	**0.500**	**0.510**

In Table 8, the four positive Emotions (**Admiration**, **Compassion**, **Pride**, and **Trust**) have ESS values that not only exceed 0.5 (the perfect polarization point) but remain consistently above the overall ESS average for all five years (with one narrow exception for **Trust** in 2017). These Emotions, on net, evoke stronger extreme positive sentiment than extreme negative sentiment. In contrast, the six negative Emotions are mostly below the perfect polarization point in most years, with the ESS scores of **Anger** and **Hate** also below the overall ESS average in all five years. For these, extreme negative sentiment tends to outweigh extreme positive sentiment, significantly so in some years.

These results, along with the Value results, point to the potential role that Values and Emotions play in influencing sentiment. In the next section, we undertake a more in-depth analysis of that role using statistical modeling. Instead of looking at a single Value or Emotion, one at a time, statistical modeling enables us to understand how all Top Ten Values and Emotions, taken together and along with Time (or Year), Document Source, and STEEP Categories, impact WSS and ESS.

4. Statistical Analysis of Sentiment Towards Inequality

Previous sections of this paper established the building blocks for a full-blown statistical analysis of how sentiment scores are determined. Section 2 tracked trends in the evolution of sentiment over time, using the overall median WSS and ESS for the purpose, and doing so separately for documents by source or in various STEEP categories. Section 3 introduced the idea that Values and Emotions may have a role in determining those trends. This section uses statistical modeling to more rigorously confirm the ability of Values and Emotions to drive sentiment scores.

4.1. Statistical Models

Discrete choice models are frequently used to analyze survey data. This class of statistical models is ideal for dependent variables that are discrete or categorical, whether binary or multi-category. Most commonly used are logistic (or "logit") regression and "probit" regression models.[40]

Binary Logit Regression is designed for discrete dependent variables that take either the value 0 or the value 1. Because the dependent variable cannot take any value in between or outside that range, it cannot be considered a continuous variable and, hence, Ordinary Least Squares (or a variant of it for a continuous dependent variable) cannot be used to estimate the parameters of the model.[41] But, what if the dependent variable is a proportion or fraction of some kind, which can be continuous in the range from 0 to 1 but never falls outside that range. In some instances, it can even take on the boundary values themselves, namely, 0 and 1. The best way to view such a variable is as a <u>bounded</u> continuous variable that falls within a narrow and fixed range. How can the parameters of such a model be estimated?

A special method, most often called Fractional Logistic Regression (or "Fractional Logit"), is well-suited to models with continuous but bounded dependent variables. It has a similar structure to the Binary Logit Regression model and can be estimated by quasi-maximum likelihood methods.[42] That is the modeling methodology we apply in the present context.

With Logit-type regression models, we have the option to report results either in the form of coefficients (which are estimates of the parameters attached to independent variables) or their odds

---

[40] Discrete choice analysis is now a standard part of micro-econometrics, which typically deals with data on individuals (whether consumers, firms, voters, or other micro-units). There is a large literature on the specification, estimation, testing, and interpretation of discrete choice models. One standard reference is Greene, W.H., <u>Econometric Analysis</u>, 8th Edition, Pearson Education Limited, 2018. See especially Part IV.

[41] Some researchers use the Linear Probability Model (LPM) for the binary outcome and estimate it using least squares methods, regardless of several violations of the assumptions underlying least squares estimation. This gives rise to several well-known problems, such as heteroskedasticity, non-normality, possible non-linearity, and misleading tests of statistical significance for estimated model parameters. Most seriously, the LPM also leads to model-based predictions that lie outside the [0,1] range, which are essentially meaningless.

[42] Papke, L.E., and J.M. Wooldridge, "Econometric Methods for Fractional Response Variables with an Application to 401(K) Plan Participation Rates," <u>Journal of Applied Econometrics</u>, 11, 1996, 619-632. Also, Wooldridge, J. M., <u>Econometric Analysis of Cross Section and Panel Data</u>, 2nd ed., Cambridge, MA: MIT Press, 210.

ratios (OR). There is a close correspondence between a coefficient estimate and its OR,[43] and statistical inference using any one of the two is equivalent to doing so using the other.

A positive coefficient signifies that the impact of an independent variable (or driver) is to boost the value (or probability) of the dependent variable, while a negative coefficient signifies just the opposite, i.e., a reduction. When a coefficient estimate is not statistically significant, it signifies no impact at all, i.e., neither positive nor negative. When a coefficient estimate is positive, the corresponding OR is a number greater than one. When the coefficient estimate is negative, the corresponding OR is a number between zero and one. A coefficient estimate of zero (in effect, when it is not statistically significant), corresponds to an OR that is exactly equal to one.

The range of an OR is bounded at the bottom by zero, but it has no upper bound. As long as it is zero or higher, but not one or higher, the independent variable or driver in question has a negative impact, which grows in magnitude as the OR gets closer to zero. If the OR is greater than one, then the independent variable has a positive impact, which grows as the OR itself gets larger. Thus, an OR of one is the threshold or separating value between positive and negative impacts.[44]

4.2. Statistical Model Specification for WSS and ESS

For the overall median of both WSS and ESS, we specify a statistical relationship in which that overall median is the dependent variable (or, outcome) and independent variables (or, drivers) fall into five sets. The purpose of such a model is to estimate how — and how much — all the drivers simultaneously impact the outcome, here the overall median WSS or ESS.

The first set captures the impact of time. Setting the year 2021 as the "reference level," we study the impact of time, i.e., effects of the years 2017, 2018, 2019, and 2020 on the overall median WSS or ESS.[45] That is, the impact of any of those four years is estimated relative to the year 2021. Note that time or year is itself just a proxy for possible drivers that change over time and impact WSS or ESS, but are not directly observed and cannot be included in the model.

The second set comprises document sources. In our model, document source has four levels: New York Times, Washington Post, Wall Street Journal, and the General Web. All other sources (in the form of URL) represent the reference level.

The third set comprises the Top Ten Values listed in Tables 4 and 5, and the reference level consists of a composite of all other Values that are not in the top ten. With a common denominator like

[43] OR is simply the exponentiated value of a coefficient, i.e., OR = exp(coefficient).

[44] One reason for preferring to report ORs rather than their corresponding coefficient estimates is that ORs are invariant to the units in which the independent variables are measured. Moreover, ratios of ORs for different levels of an independent variable can be used to measure relative impacts of those levels. This is explained in fn. 49 below with an example from Table 9.

[45] If an independent variable in a logit or probit regression model is also categorical, it is necessary to set any one level of that variable as the reference level against which every other level of that variable is measured.

"All Other Values," it is also possible to make head-to-head comparisons among the Top Ten Values that represent the levels of the driver Values.

Similarly, the fourth set comprises the Top Ten Emotions listed in Tables 7 and 8, and the reference level consists of a composite of all other Emotions that are not in the top ten. In this, "All Other Emotions" is the common denominator which allows head-to-head comparisons among the Top Ten Emotions themselves.

The final set comprises the three STEEP categories considered in this study namely, Economy, Politics, and Society, while the reference level is a composite of the two remaining STEEP categories, namely, Environment and Technology.

All five sets of drivers and a "constant" representing all other impacts are included simultaneously in the statistical model. That model is then estimated using the quasi-maximum likelihood method (e.g., the `fracreg` program in Stata).[46]

4.3. Statistical Model for Overall Median WSS

The fractional logistic regression model estimated for the overall median WSS is reported in Table 8. All ORs shown in red are statistically significantly different from the threshold value of one at the 5% level (i.e., probability value of 0.0500). The rest are not statistically significantly different.

Table 9. Fractional Logistic Regression Model for WSS and its Drivers

Dependent Variable:	WSS				
Reference Variable	**Independent Variable**	**Odds Ratio**	**Std Error**	**Z-statistic**	**Prob Value**
Year 2021	Year 2017	1.0313	0.0095	3.36	0.0010
	Year 2018	1.0375	0.0090	4.24	0.0000
	Year 2019	1.0326	0.0082	4.03	0.0000
	Year 2020	0.9696	0.0072	-4.13	0.0000
Other Source (URL)	NYT	0.9469	0.0127	-4.06	0.0000
	WP	0.9386	0.0125	-4.74	0.0000
	WSJ	1.0111	0.0142	0.78	0.4330
	WEB	0.9671	0.0169	-1.91	0.0560
All Other Values	Access	1.0358	0.0097	3.78	0.0000
	Assistance	1.0285	0.0090	3.20	0.0010
	Balance	1.0421	0.0142	3.03	0.0020
	Diversity	1.0203	0.0133	1.54	0.1220
	Empowerment	1.0384	0.0130	3.02	0.0030
	Equality	0.9608	0.0085	-4.54	0.0000
	Growth	1.0909	0.0127	7.46	0.0000

[46] Stata® is a statistical software package from StataCorp LLC, 4905 Lakeway Drive, College Station, TX 77845-4512. Version 17 is now available.

	Value	1.0377	0.0145	2.64	0.0080
	Wealth	1.4998	0.0256	23.74	0.0000
	Youth	0.8993	0.0103	-9.25	0.0000
All Other Emotions	Admiration	1.1170	0.0110	11.20	0.0000
	Anger	0.8192	0.0085	-19.28	0.0000
	Compassion	1.1175	0.0102	12.16	0.0000
	Disappointment	0.9221	0.0099	-7.54	0.0000
	Guilt	0.7554	0.0116	-18.20	0.0000
	Hate	0.7600	0.0119	-17.47	0.0000
	Pride	1.1748	0.0163	11.61	0.0000
	Sadness	0.8422	0.0123	-11.79	0.0000
	Trust	1.0723	0.0101	7.38	0.0000
	Worried	0.9045	0.0132	-6.90	0.0000
All Other STEEP	Economy	1.1359	0.0245	5.90	0.0000
	Politics	1.0146	0.0210	0.70	0.4840
	Society	1.0929	0.0224	4.33	0.0000
	Constant	1.1030	0.0287	3.77	0.0000

Interpretation of the estimates in Table 9 is as follows:

Time: Within the first set of independent variables, Years 2017-2020 all have ORs that are statistically significant.[47] For three of those years (2017-2019), the ORs exceed one, signifying that the median WSS was boosted relative to 2021. That is, there was optimism (i.e., a net positive outlook) about inequality, in general, in those years compared to 2021.[48] Moreover, 2018 registered the highest OR among those years, suggesting that in that year, the inequality outlook was the most optimistic among all five years for which documents were studied.[49] In contrast, the OR for 2020, while statistically significant, was below one, signifying a more pessimistic outlook towards inequality I that year compared to 2021 and, therefore, every other year in the study.

[47] When testing the statistical significance of a coefficient estimate, it is commonplace to do so against a null hypothesis value of zero. That is, failure to find statistical significance amounts to the independent variable (driver) in question having a zero value, i.e., no impact on the dependent variable (outcome). When testing for the statistical significance of an OR, the corresponding null hypothesis value is one (which is when no relationship or impact is found). The use of ORs instead of coefficient estimates in this paper means that all statistical tests of ORs are made against the null hypothesis value of one.

[48] In this section, we use the terms "optimism" and "pessimism" to mean a net positive sentiment/outlook (more favorable perception) and a net negative sentiment/outlook (less favorable perception), respectively.

[49] As noted in fn. 44, relative impacts of different levels of an independent variable can be compared by taking the ratio of their ORs. For example, in Table 9 (which reports ORs for Years 2017-2020 relative to Year 2021), the ORs for 2017 and 2018 are, respectively, 1.0313 and 1.0375. Two conclusions may be drawn from these. First, the odds of 2017 boosting the median WSS were 3.13% higher than for 2021 (and 3.75% higher for 2018 than for 2021). Second, in a head-to-head comparison using the ratio of ORs for 2017 and 2018, the latter year had a 3.75%/3.13% = 1.2 times (or 20%) higher odds than the former year of boosting the median WSS.

Document Source: The second set of independent variables comprise the three major newspapers and the general Web, evaluated relative to all other online document sources. Only two of those sources (**New York Times** and **Washington Post**) have statistically significant ORs below one, signifying that views expressed about inequality by them are, in general, more pessimistic than those expressed by other sources. **Washington Post** documents are the most pessimistic of all sources.

Values: All Top Ten Values (shown in alphabetical order in Table 9) have statistically significant ORs. Of these Values, eight have ORs greater than one and two (**Equality** and **Youth**) have ORs below one. **Wealth** has an OR that far exceeds any other — 37% higher than for **Growth**, the Value with the next-highest OR. So, relative to all other Values, **Wealth** is the most significant impetus for expressing optimism about inequality. In contrast, **Equality** and **Youth** are associated with pessimistic sentiment about the state of inequality. This finding likely reflects a perceived failure to achieve desired standards of equality or the aspirations of youth regarding equality.

Emotions: All Top Ten Emotions (shown in alphabetical order in Table 8) have statistically significant ORs. Of these Emotions, four have ORs greater than one and six have ORs below one. Arguably, the four Emotions with ORs greater than one, namely, **Admiration**, **Compassion**, **Pride**, and **Trust** are all "positive" Emotions that associate the documents analyzed for sentiment with optimism about inequality. Similarly, the six Emotions with ORs below one, namely, **Anger**, **Disappointment**, **Guilt**, **Hate**, **Sadness**, and **Worried** may be thought of as "negative" Emotions that associate those documents with pessimism about inequality. **Pride** has the most favorable outlook, while **Guilt** has the least.

STEEP Categories: Only **Economy** and **Politics** have statistically significant ORs, one above one (**Economy**) and the other below (**Politics**). The third STEEP Category, namely, **Society** appears to be on par with all other STEEP Categories (**Environment** and **Technology**) in how favorably (or otherwise) the state of inequality is viewed. **Economy** has a net positive outlook but only slightly so because its OR exceeds one only marginally. The economic "lift" that may have been experienced in the earlier years of the study period were attenuated somewhat in the later years as Covid-19 had several adverse effects on the economy. In contrast, **Politics** has a net negative outlook, probably because of the combined effect of the fraught state of Presidential politics and political upheavals in the second half of the study period.

4.4. Statistical Model for Overall Median ESS

The fractional logistic regression model estimated for the overall median ESS is reported in Table 10. Only the ORs shown in red are statistically significantly different from the threshold value of one at the 5% level (i.e., probability value of 0.0500).

Table 10. Fractional Logistic Regression Model for ESS and its Drivers

Dependent Variable:	**ESS**				
Reference Variable	**Independent Variable**	**Odds Ratio**	**Std Error**	**Z-statistic**	**Prob Value**
Year 2021	Year 2017	1.0157	0.0054	2.90	0.0040
	Year 2018	1.0027	0.0049	0.55	0.5790
	Year 2019	1.0251	0.0048	5.34	0.0000
	Year 2020	0.9774	0.0041	-5.41	0.0000
Other Source (URL)	NYT	0.9616	0.0073	-5.13	0.0000
	WP	0.9539	0.0072	-6.23	0.0000
	WSJ	0.9828	0.0077	-2.23	0.0260
	WEB	0.9557	0.0098	-4.40	0.0000
Other Values	Access	1.0142	0.0050	2.89	0.0040
	Assistance	0.9986	0.0050	-0.28	0.7790
	Balance	1.0155	0.0074	2.13	0.0340
	Diversity	0.9789	0.0080	-2.60	0.0090
	Empowerment	1.0147	0.0093	1.59	0.1120
	Equality	0.9792	0.0047	-4.41	0.0000
	Growth	1.0260	0.0061	4.33	0.0000
	Value	1.0302	0.0083	3.69	0.0000
	Wealth	1.2061	0.0145	15.60	0.0000
	Youth	0.9329	0.0059	-10.96	0.0000
Other Emotions	Admiration	1.0472	0.0060	8.11	0.0000
	Anger	0.9103	0.0053	-16.28	0.0000
	Compassion	1.0568	0.0056	10.49	0.0000
	Disappointment	0.9578	0.0058	-7.14	0.0000
	Guilt	0.8892	0.0073	-14.34	0.0000
	Hate	0.8487	0.0099	-14.00	0.0000
	Pride	1.1088	0.0119	9.62	0.0000
	Sadness	0.9293	0.0068	-9.96	0.0000
	Trust	1.0251	0.0051	4.95	0.0000
	Worried	0.9813	0.0061	-3.03	0.0020
Other STEEP	Economy	1.0271	0.0106	2.60	0.0090
	Politics	0.9784	0.0099	-2.16	0.0310
	Society	1.0093	0.0100	0.94	0.3500
	Constant	1.0795	0.0140	5.90	0.0000

Interpretation of the estimates in Table 10 is as follows:

Unlike WSS, the role of ESS is not so much to reveal the average sentiment as to determine the directionality of the overall sentiment (with perfect polarization as the pivot point). With that in mind, the ORs in Table 9 should be understood as follows. When, relative to the reference level, any given level of an independent variable has an OR greater than one, there is a skewing of net extreme sentiment in a positive direction, i.e., the effect of that level of the independent variable is to create an extreme positive sentiment that outweighs the extreme negative sentiment. If, on the other hand, the OR is less than one, then the opposite is true: extreme negative sentiment outweighs extreme positive sentiment. Only when the OR is one, do the two types of extreme sentiment offset, i.e., perfect polarization occurs.

Time: Once again, ORs are interpreted relative to the year 2021. Table 10 shows that ORs are statistically significant in the years 2017, 2019, and 2020, but not 2018. Moreover, ORs are greater than one in 2017 and 2019, but less than one in 2020. That means that, relative to 2021, net extreme sentiment skewed slightly positive in 2017 and 2019, but skewed negative in 2020. These findings reflect the pattern in the overall median ESS shown in Figure 4.

Document Source: ORs for all four document sources considered in this study, relative to all other sources, are statistically significant <u>and</u> below one. Comparing the four sources themselves, **Washington Post** has the lowest OR, while **Wall Street Journal's** OR, while still below one, is the highest. This is also the pattern observed in Figure 5. Relative to all other sources, extreme negative sentiment outweighs extreme positive sentiment for all four sources, most for **Washington Post** and least for **Wall Street Journal**.

Values: Two of the Top Ten Values, namely, **Assistance** and **Empowerment**, have ORs that are not statistically significant and are, in effect, equal to one. Relative to all other Values, their net ESS are neither higher nor lower. However, of the other eight Top Ten Values that are statistically significant, five (**Access**, **Balance**, **Growth**, **Value**, and **Wealth**) are greater than one and three (**Diversity**, **Equality**, and **Youth**) are below one.

The former five Values, particularly **Wealth**, are most likely to <u>increase</u> the overall median ESS (as extreme positive sentiment outweighs extreme negative sentiment). Stated another way, those five Values make it likely for documents that receive extreme sentiment ratings to view inequality optimistically. In contrast, the latter three Values, particularly **Youth**, are most likely to <u>reduce</u> the overall median ESS (as extreme negative sentiment outweighs extreme positive sentiment). These three Values make it likely for documents that receive extreme sentiment ratings to view inequality pessimistically.

Emotions: While ORs for all Top Ten Emotions are statistically significant, four are greater than one and the other six are below one. In fact, the pattern is exactly the same as that found for the ORs in the WSS statistical model. All supposedly positive Emotions (**Admiration**, **Compassion**, **Pride**, and **Trust**) have ORs that exceed one, most **so** for **Pride**. These four Emotions are likely to <u>increase</u> the overall median ESS (as extreme positive sentiment outweighs extreme negative sentiment), thus making it likely for documents that receive extreme sentiment ratings to view inequality optimistically. In contrast, all supposedly negative Emotions (**Anger**, **Disappointment**, **Guilt**, **Hate**, **Sadness**, and **Worried**) have ORs that are below one, most so for **Hate**. These six Emotions are most likely to <u>reduce</u> the overall median ESS (as extreme negative sentiment outweighs extreme positive sentiment). That is, those six Emotions pull documents that receive extreme sentiment ratings towards viewing inequality pessimistically.

STEEP Categories: Only **Economy** and **Politics** have statistically significant ORs, the former above one and the latter below. The third STEEP Category, namely, **Society** appears to be on par with all other STEEP Categories (**Environment** and **Technology**) in how extreme sentiment ratings shape views about the state of inequality. For **Economy**, extreme positive sentiment outweighs slightly the extreme negative sentiment, implying that documents about the economy likely view inequality optimistically, although only marginally so. In contrast, for **Politics**, extreme negative sentiment outweighs slightly the extreme positive sentiment, implying that documents about the country's political situation likely view inequality pessimistically, although (again) only marginally so.

5. Conclusions

Textual analysis is, and will continue to be, a part of the toolkit available to a researcher. In a broad sense, our research to date shows the promise of text mining. The research suggests that information about attitudes and perceptions can be extracted from documents that are chosen to investigate any concept, e.g., the term “inequality.” If the process works for such a broad concept, then it should work for analyses employing more specific or precisely-worded queries. Indeed, an important part of text mining is accurately framing the query or concept itself. In most circumstances, precisely-worded queries are likely to yield better or more meaningful insights. The research objectives should dictate the type and complexity of the queries. A more complex query, such as for “income inequality,” would be more direct and better suited for an analysis of sentiment regarding a specific type of inequality. However, even the simple query about inequality pursued in this study is sufficiently interesting and successful at underscoring the value of text mining.

We postulated that there is a link between the query (about inequality in this study) and the measurement of sentiment. This is a foundational issue for measuring sentiment through text mining. But, it is even more important to be able to interpret or explain sentiment trends using a heuristic review of key contemporaneous events that have a bearing on attitudes (implicit or explicit) towards, or perceptions of, a concept like inequality. This is analogous to using event history analysis[50] to explain trends in real or historical data.

Whereas some type of sentiment analysis is part of most text mining projects, the paper proposes alternative approaches to measuring and scoring sentiment. As its principal contribution, this paper demonstrates how WSS (to capture the average level of sentiment) and ESS (to capture the directionality of sentiment) can determine whether the outlook or attitude expressed in textual data is net positive (optimistic) or net negative (pessimistic). More importantly, while both WSS and ESS can be applied to the assessment of individual drivers of outcomes, the paper shows how a simultaneous assessment of <u>all</u> hypothesized drivers may be conducted using a statistical modeling approach. Statements about both level and directionality of the impact of drivers on sentiment can be made using the approach to answer questions like the following:

1. What is (are) the most important driver(s) of perception or attitude, whether in a positive or negative direction?
2. How does such a perception or attitude change over time or with different sources of textual data?
3. Compared to structured surveys, can sentiment analysis of textual data (balanced for different viewpoints) better, and less expensively, reveal the impacts of unobservable or complex drivers (like Values and Emotions) on perceptions and attitudes?
4. Can the same “query → sentiment-based analytics → insights/inferences” framework be used repeatedly to understand widely differing contexts or concepts, thus avoiding the need to create dedicated and separate surveys for each context or concept?

---

[50] See https://methods.sagepub.com/Book/handbook-of-data-analysis.

5. How should sources of textual data be marshaled when using the Sentometrics approach, whether for general concepts like inequality or for more specific concepts like consumer confidence or product acceptance?

Although Sentometrics (especially of the kind demonstrated in this paper) advances the cause of understanding the world around us, it is not just another arrow in the quiver. Its biggest virtue is to recognize the power of opinion, attitudes, beliefs and the socio-psychological makeup of actors — whether economic, political, social, etc. — to shape human actions and choices and, thus, to determine the course of any environment in which human beings interact and conduct transactions. Traditional sources of data — whether historical or survey-based — can be cumbersome, expensive, and often only lead to insights long after the data were realized. Textual data are akin to "revealed preferences," in that they rely on attitudes and perceptions already expressed by thousands of documents or individuals, and can be harnessed at relatively low cost or investment of time by the use of a well-designed MMX-type query mechanism. As a complement to this mechanism, the Sentometrics format of this paper creates a parallel channel for understanding human behavior by recognizing the link between experience, sentiment, and choice. This channel is grounded in the use of contemporaneous real-world events to explain trends in such behavior and, simultaneously, the use of rigorous quantitative or statistical tools to yield objective insights.

This paper presents a novel approach to addressing a complex concept (here, inequality) using a methodology that extracts sentiment (and its possible drivers like Values, Emotions, STEEP categories, etc.) from textual data. But, we believe, this is just the beginning. This alternative approach can be expected to open up new avenues of research without having to rely on conventional quantitative surveys. A new direction to explore would be the building of dynamic Sentometric models that track two important changes over time (at, say, monthly, quarterly, or annual intervals): (1) in each period's Top Ten Values and Emotions and (2) in the direction and size of the sentiment scores associated with the Values and Emotions that are their drivers. The purpose here would be to identify/measure how changes in the leading Values and Emotions change sentiment scores and, eventually, behaviors or perceptions. These developments can potentially advance an analysis such as that pursued in this paper from a descriptive to a more predictive exercise. In turn, that will open doors for forecasting future trends, causality testing, simulation exercises, and integration with more conventional models of consumer and market behavior.