

Biased Newspapers, Naive Readers, and Political Polarization  
in the UK

A thesis submitted by

Andong LIU

in partial fulfillment of the requirement for the degree of

Master of Science

in

Economics

Tufts University

May 2021

Advisor: Eoin McGuirk, Enrico Spolaore

## **Abstract**

This research utilizes respondent-level survey data to explore the relationship between news media consumption and political polarization in the UK. Unlike the existing literature that often assumes that individuals are homogeneous in their degree of sophistication, this research introduces the concept of naivety, develops a set of naivety measures, and investigates how the heterogeneity of British readers' naiveties can potentially change to what extent readers are polarized given the same amount of newspaper consumption. The findings suggest that naive readers get less polarized to the right than sophisticated ones given the same amount of right-wing newspaper consumption.

## **Acknowledgements**

I would first like to thank my supervisors, Professor Eoin McGuirk and Professor Enrico Spolaore, whose insights greatly helped me throughout the process of writing this thesis. I would like to acknowledge my classmates who asked questions and gave suggestions that bettered this work. I would particularly like to thank Thomas Varkey for his generous financial support. In addition, I would like to thank my family for all their love and support. Finally, I would like to thank my friend, Zhaoji Tang, who provided his British school email for me to get access to UK newspaper data.

# Table of Contents

List of Figures.....	v
List of Tables.....	v
1. Introduction.....	1
2. Institutional Background .....	4
3. Data and Variable .....	6
4. Empirical Analyses .....	13
4.1. Summary statistics .....	13
4.2. Naivety and demographics .....	15
4.3. Newspaper consumption, naivety, and political polarization .....	17
5. Conclusion .....	26
6. Reference .....	28

## List of Figures

Figure 1: The Total Circulations of The Ten National Newspapers .....	6
Figure 2: How Left or Right Wing Are the Mainstream UK Newspapers by YouGov .....	9

## List of Tables

Table 1: The Ten National Newspapers' Readership and Biases .....	14
Table 2: Newspaper Readership by Readers' Political Leaning and Naivety .....	15
Table 3: Summary Statistics .....	15
Table 4: Naivety and Demographics .....	16
Table 5: OLS Regressions .....	19
Table 6: First-Stage Regressions A .....	22
Table 7: First-Stage Regressions B .....	23
Table 8: Reduced-Form Regressions .....	24
Table 9: Second-Stage Regressions .....	25

## 1. Introduction

Media play a predominant role in the democratic process. In one view, media matters because it provides information to predominately rational voters. In another view, media matters through propaganda and by exploiting cognitive mistakes that voters make (Strömberg, 2015). However, media are often biased. In the introduction of their paper, Gentzkow and Shapiro (2006) provide a vivid example: On December 2, 2003, in the Iraqi city of Samarra, American troops fought a battle, but Fox News, the New York Times, and the Al Jazeera (English version) conveyed a radically different impression of what actually happened based on the same set of underlying facts. Then should we worry about media bias?

Most empirical works on the political outcomes of media bias suggest that media bias can polarize the electorate (Bernhardt et al., 2008; Chan and Suen, 2008). Some even find direct evidence that it can influence voting (DellaVigna and Kaplan, 2007). However, when examining the effects of media on consumers, there is merely a small portion of studies that mention consumers' naivety that potentially affects to what extent they are manipulated by media. Chiang and Knight (2011) find that voters rationally account for the credibility of any endorsements of newspapers. Their results to some extent suggest that readers are sophisticated. But are all readers sophisticated? We should expect that there is at least some heterogeneity in the degree of sophistication among media consumers. Levy (2021) observed that liberal or conservative treatments did not affect political opinions even though they dramatically affected the Facebook feed of participants. Baysan (2020) also documented this kind of heterogenous effect

in other settings. This leads to our research question: do biased media have heterogeneous polarizing effects on the ideologies of naive and sophisticated readers?

Before we delve into the effects, we first need to measure this naivety. As mentioned above, the existing literature simply mentions the concept of naivety, instead of measuring it and probing its impacts. To my knowledge, the only exception is the study of social media and political polarization by Levy (2021). This research defines people's sophistication in an indirect way by asking questions like: Suppose 110 members of a local government voted on an infrastructure bill. The bill passed by a margin of 100 votes. How many members voted against the bill? But he did not find evidence supporting the existence of the heterogeneous effects. This study uses a more direct way to measure naivety at the level of an individual person. By introducing the subjective and objective bias of a media outlet, as well as the gap between the two biases, we are able to quantify this naivety through certain specifications and pinpoint characteristics that are associated with it. The subjective bias of a media outlet is the bias that consumers subjectively perceive, while the objective thereof is the bias measured and well-documented by economists. Therefore, our naivety estimate is highly dependent on the credibility of the objective bias. Gentzkow and Shapiro (2006) construct their measure of media bias by using the speaking time in local news broadcast stations given to two presidential candidates in the general election of the U.S. in 2000. If a station gives George W. Bush more time than Al Gore, then this media outlet is biased in favor of the former. Groseclose and Milyo (2005) measure the bias of a certain media outlet by examining what think tanks it cites given their ideologies.

Gentzkow and Shapiro (2010) compare phrase frequencies in newspapers with phrase frequencies in the 2005 Congressional Record and determined which party newspapers slant towards by measuring the similarities of their language to that of a congressional Democrat and Republican. Qin et al. (2018) use principal component analysis to capture how strongly the content of a newspaper reflects its political goals to measure the bias of the newspaper. These methods are basically content analysis and provide a fairly objective way to measure how slanted media outlets are. This study can thus utilize these methodologies to measure the objective bias of each outlet studied. Ultimately, by measuring the gap between an individual's subjective assessment of a newspaper's bias and the objective metric thereof, naivety measures can be generated. Despite different measures, a simple example is that a naive reader is someone who reads an objectively biased paper like the Daily Mail but who thinks that the paper is less biased than the objective metric.

After the concept of naivety is well defined and measured, we try to identify what characteristics contribute to it as a set of preliminary results. Then we turn to investigate whether naive readers are more prone to be polarized by biased media outlets than those who are sophisticated. A couple of studies look at the effect of media exposure on political polarization. Bernhardt et al. (2008) propose that readers are rational and understand what coverage is missing and suggest that a biased media that suppresses the negative news about readers' preferred candidate does not benefit the preferred candidate but polarizes the electorate. Martin and Yurukoglu (2017) measure the persuasive effects of slanted news and find that cable news increases ideological



polarization among the viewing public. Levy (2020) estimates the effects of social media exposure by a field experiment based on Facebook and finds that social media algorithms may limit exposure to counter-attitudinal news and thus increase polarization. These studies invariantly conclude that media exacerbate political polarization, but they fail to consider the possibility that the heterogeneity of readers' naivety can potentially change to what extent readers are polarized given the same amount of media consumption. By filling this gap, our study might be able to shed some light on the exact mechanism of how media consumption and political polarization are related.

The main data source used is "The News Consumption in the UK" survey conducted by The Office of Communications, which regulates British communications services. This respondent-level survey provides much information regarding Briton's media consumption habits and their evaluations of many aspects of each media outlet they consume, including its accuracy, impartiality, trustworthiness, and so on. The empirical results suggest that the heterogeneous effects are present and naive readers get less polarized to the right than sophisticated ones given the same amount of right-wing newspaper consumption.

## **2. Institutional Background**

The Office of Communications (Ofcom) is the government-approved regulatory body for the communications services in the UK, including the television, radio, telecoms, and postal sectors. It sets strict broadcasting codes to ensure that news, in all forms, should be reported with due accuracy and presented with due impartiality.

Therefore, despite some complaints, British televisions and radios are generally considered to be less biased. However, the British press is not regulated by Ofcom. A major regulator of the newspaper and magazine industry is the Independent Press Standards Organization (IPSO), which reregulates and is financed by its members' publications. Most national newspapers are its members, but several newspapers refused to join the IPSO and established their own independent complaints system, for example, the Guardian and the Financial Times. It is thus not surprising that British newspapers are often criticized to be biased.

This study focuses on the ten most circulated national newspapers (excluding freesheets, for example, the Metro) due to their significance. According to the News Consumption survey conducted by Ofcom, among all print newspaper readers, around 80% of them read at least one of these ten newspapers. The Audit Bureau of Circulation that collects and releases British newspaper circulation data shows that the ten newspapers are The Sun, Daily Mail, The Times, The Daily Telegraph, Daily Express, Daily Star, Daily Mirror, i, Financial Times, The Guardian, and their Sunday counterparts (if any). Most of them have a clear political stance. The first six newspapers except Daily Star are considered to be right-wing and they all endorsed the Conservative Party in the 2019 United Kingdom general election. Daily Star has a slight right-wing stance and did not endorse any party in the 2019 general election. In contrast, the left-wing forces are relatively weak. Daily Mirror and The Guardian are believed to be fairly left-wing and endorsed the Labour Party, while i and Financial Times are generally center with a slight left-wing leaning and did not endorse any party in the

2019 general election.

British national newspapers have been crucial news sources for many centuries, but they are also experiencing a significant long-term decline in their circulations. Figure 1 shows the total circulation of the 10 national newspapers for the period from 2011 to 2020. The circulation number dropped 76.8% mainly due to the widespread of the Internet. All these newspapers thus developed their online newspaper websites or mobile applications, which also contributed to the decline of their print circulations.

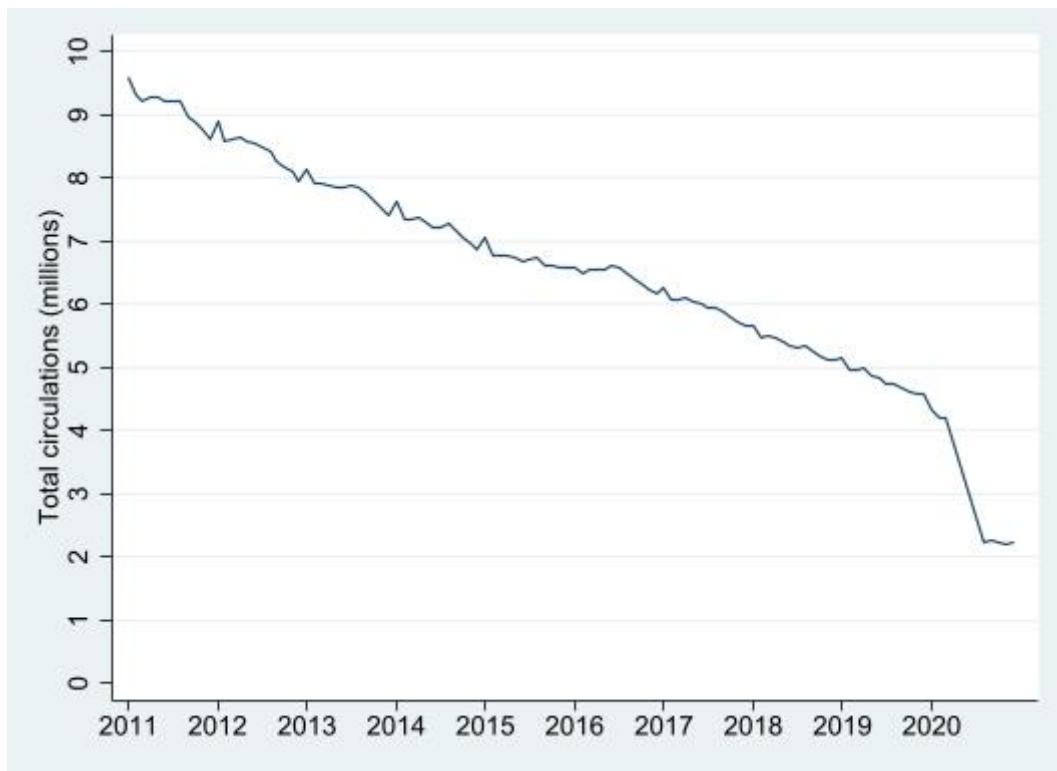


Figure 1: The Total Circulations of The Ten National Newspapers

Note: The Y-axis denotes the total circulations of the ten national newspapers. The unit is million.

### 3. Data and Variable

This research endeavors to estimate the heterogeneous polarizing effects of newspaper consumption on readers' political ideology among naive and sophisticated readers. The main variables needed include newspaper consumption, political ideology,

and naivety measures.

To obtain these variables, two datasets are used in this paper. The first one is from the Office of Communications (Ofcom) that regulates the communications services in the UK. Their series of research “The News Consumption in the UK” is conducted on a yearly basis with consistent methods, which provides information about news consumption across television, radio, print, social media, podcasts, other internet sources, and magazines. In addition to the amount of news people consume on different platforms, this respondent-level survey also asks questions about how often people use these platforms, as well as how they rate them in terms of accuracy, trustworthiness, impartiality, and etc. This dataset also contains the respondents’ demographics, including gender, age, region, education, political ideology, and etc. Each year, around 4500 respondents across the UK are interviewed by online methods or face-to-face. As the focus of this thesis is the British press, I mainly use print newspaper data.

The second dataset is also from Ofcom. The Ofcom Broadband speeds research collects data on fixed line home broadband speeds, mobile broadband performance. It contains meticulous data on various broadband speeds by service providers, region, and type of speed measures. As newspaper consumption shapes readers’ ideology and readers choose which newspapers to read based on their ideology, the consumption variable should be endogenous and require an instrumental variable to be its proxy. The instrument, regional Internet speeds, is generated based on these data. Due to the availability of the data, the sample year is set to be the period from 2018 to 2020.

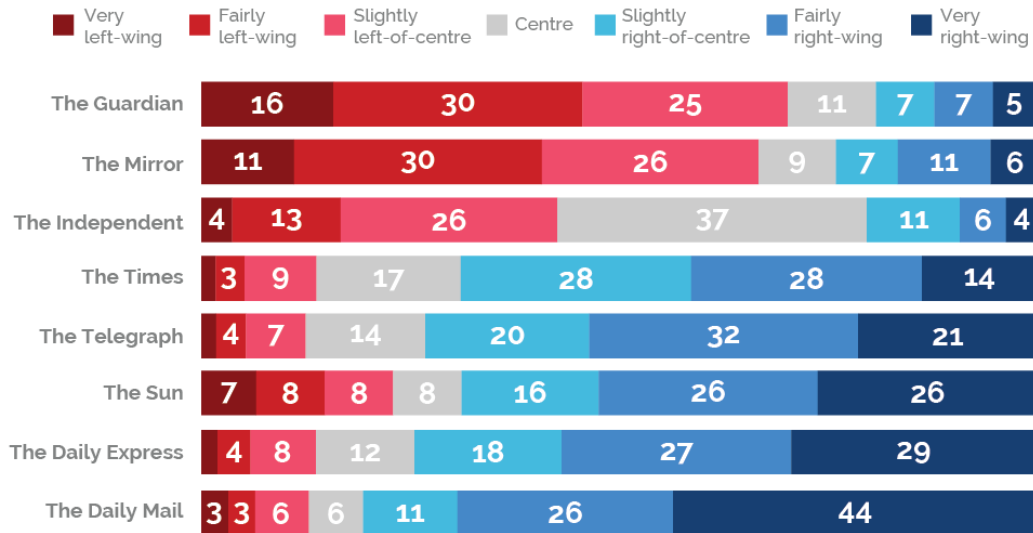
A key variable of this thesis is naivety. Certain media outlets are biased in their

reporting, but some consumers might be not fully aware of the bias. In this sense, they are naive. For the reader group of a given newspaper, I divide readers into two groups based on the gap between the subjective bias that one innately perceives and the objective bias measured and well documented by economists. Naive readers are those with a negative bias gap and sophisticated readers are those with a non-negative bias gap.

The subjective bias of a given newspaper is extracted from a question in the Ofcom survey. It asks respondents to what extent they think a newspaper brand—including its daily edition and weekly counterpart (if any)—they use is impartial and requires them to answer this question by using a scale of 1 to 10, where 1 is not at all and 10 is completely. This impartiality is converted to biasedness by reversing the order so that 1 (impartial) means 10 (biased) and vice versa. This method of conversion loses the direction of the bias (left-wing or right-wing) but this could be acceptable according to a survey conducted by YouGov that shows that the majority of newspaper readers know the direction of the mainstream UK newspapers' biases.

## How left or right wing are the mainstream UK newspapers?

Some people talk about 'left', 'right' and 'centre' to describe parties and politicians. With this in mind, where would you place each of the following? (excludes those who said "don't know" for each paper - between 39-49% of respondents)



YouGov | yougov.com

February 20-22, 2017

Figure 2: How Left or Right Wing Are the Mainstream UK Newspapers by YouGov

Note: these results exclude those who said they didn't know what the newspaper's ideology was – a rate of between 39% and 49% for each newspaper

The objective bias of a given newspaper is relatively objective as it is not evaluated by me and is based on many scientific methods developed by economists and other social scientists. The data are obtained from Media Bias/Fact Check, which is an independent online media outlet that determines the bias of media and information sources through a combination of methods. The main methodology involves using a scale of 0 to 10 to rank biases of a certain media source in four different categories: biased wording (whether it uses loaded words to convey emotion to sway its readers), factual/sourcing (whether it reports factually and backs up claims with well-sourced evidence), story choices (whether it reports news from both sides), and political

affiliation (how strongly it endorses a particular political ideology). Then these four numbers are added up and divided by 4 to get the final score of how biased a given media outlet is. Since the subjective bias ranges from 1 to 10, the objective bias is converted to be consistent with the previous scale.

In a reader group of a certain newspaper, different individuals can have different subject biases with respect to the newspaper, but the objective bias of that newspaper is identical across individuals in the group. It is very often that there is a gap between the subjective and objective bias of a newspaper one reads. In this thesis, the bias gap of a certain newspaper for an individual is defined as her subjective bias with respect to that newspaper minus the objective bias thereof. This gap is critical in terms of how we measure the naivety of readers.

For a given newspaper, for example, the Guardian, if a reader has a positive bias gap of it, which means that she overperceive its objective bias, then she is fully aware of the bias of the Guardian and thus a sophisticated Guardian reader. However, if her bias gap of the Guardian is negative, she underperceives its objective bias and is partially aware of its bias. In this case, she is a naive Guardian reader.

As there are 10 national newspapers are selected and studied in this research, the naivety value assigned to each individual in the dataset is computed by aggregating this individual's 10 bias gap values for the 10 newspapers. Most readers do not read all of 10 newspapers, so the actual number of the bias gaps aggregated is equal to the number of newspapers they read. Four ways of aggregation are used, generating four different specifications of measuring naivety.

The first variable *naive* is the ratio of the number of newspaper(s) that one underperceives its objective bias to the number of newspapers one read. For example, if an individual reads three newspapers—the Daily Mail, the Times, and the Sun—her bias gaps of the Times and the Sun are negative, and that of the Daily Mail is positive, then her naivety measure will be  $2/3$ . This method of measurement emphasizes the relative extent to which a reader is naive.

The second variable *naive\_d* is based on the first variable. Instead of treating each newspaper one reads equally, it is assigned with a weight calculated by the ratio of its reading frequency to the sum of that of all newspapers one reads. Back to the previous example, if reading frequencies of the Daily Mail, the Times, and the Sun are 3, 5, 7, respectively. Then the *naive\_d* for this individual will be  $4/5$ . This measure should be arguably better than the first one.

The third variable *naive\_f* only focuses on the one newspaper that an individual reads the most frequently. This variable is thus a dummy variable, which equals 1 if the bias gap of that newspaper is negative, and equals 0, vice versa. For instance, if an individual reads multiple newspapers, the Guardian is the one she reads the most frequently, and its bias gap is 2, then the individual's *naive\_f* will be 0.

The fourth variable *naive\_c* refers to a coarse version of the first variable naivety. It equals 1 as long as there is one newspaper whose bias gap is negative. In the first example, the naivety measure will be 1 instead of  $2/3$  because this individual underperceives the objective bias of the Times and the Sun. Therefore, this measure is coarse in a way that one is either absolutely sophisticated or absolutely naive and that



there is no middle ground between the two.

In the econometric analyses, the fourth measure will be mainly used and the other three will be used for comparison and checking the robustness of the main variable.

The dependent variable *pol\_lean* refers to political leaning. It is extracted from a question in the survey that asks respondents to place their political ideology on a scale of 1 to 7 where 1,2,3,4,5,6,7 denote very left-wing, fairly left-wing, slightly left-of-centre, centre, slightly right-of-centre, fairly right-wing, very right-wing, respectively. As a common nature of surveys, many questions have a “Don’t know” or “Prefer not to answer” option. In the process of coding variables, we regard these answers as missing values, so the number of observations is reduced compared to the original dataset.

The survey asks which newspaper the respondents consume and how often they read them. The newspaper consumption variable *cons* is based on these two questions. First, values are assigned to each frequency option. For example, “less often” is assigned 1, and “many times throughout a day” is assigned 10. If an individual doesn’t read one newspaper, the frequency of that newspaper is 0. Then to get *cons*, all the frequency values are added up and this sum reflects how much time one spends on reading newspapers. For the purpose of avoiding the contradicting effects of left-wing newspaper consumption and right-wing newspaper consumption on political ideology, the *cons* variable only considers the consumption of right-wing newspapers.

The instrumental variable used to proxy newspaper consumption is regional Internet speed. The rationale behind this exercise will be explained in the next section

and the estimation strategy is IV-2SLS. The Ofcom dataset provides all kinds of speed measurements by different Internet service providers (ISPs) for 12 regions in the UK and each region is further divided into urban and rural subregions. The respondent-level survey asks whether the individual lives in urban or rural area, but it provides 14 options for region. Therefore, to assign each respondent an Internet speed measure, two regions in the survey are merged into two regions in the Internet speed dataset based on their locations. The two sets of regions are not perfectly matched, but each pair is generally in the same location. The speed measure used is the average download speed in 24 hours. In any subregion, for example, rural London, there are many speed measures from different ISPs. The variable used is generated by averaging all these values.

## **4. Empirical Analyses**

### **4.1. Summary statistics**

This section presents some background summary statistics that are critical to understanding the following econometric analyses.

Table 1 presents some summary statistics of the ten newspapers investigated in this thesis. The left-wing newspapers are overpowered by the right-wing newspapers in terms of readership. The Sun and the Mail clearly have very large reader groups and are considered to be the most biased by their readers. With respect to objective biases, left-wing newspapers are generally less biased. These facts can affect the credibility of the naivety measure with the existence of central tendency. Central tendency bias is a tendency for a rater to place most of answers in the middle of a rating scale. The table shows the average subjective biases for ten newspapers are generally around 4. However, the objective biases for all right-wing newspapers exceed 4, especially the

Sun (6) and the Mail (6), which are the two most biased newspapers; left-wing newspapers are moderately biased with the objective biases being approximately 4 except the Financial Times (1) that has far fewer readers. Therefore, in general, those who read right-wing newspapers are more likely to be naive. But this issue can be easily addressed by adding newspaper fixed effects.

Table 1: The Ten National Newspapers' Readership and Biases

Newspaper	Readership	Subjective biases	Objective biases
The Sun	1145	4.57	6.00
The Mail	1533	4.56	6.00
The Star	234	4.42	3.00
The Express	430	4.33	6.00
The Mirror	653	4.43	5.00
The Guardian	546	3.91	4.00
The i	212	3.78	4.00
The Times	745	4.00	3.00
The Telegraph	572	4.17	6.00
The Financial Times	215	3.33	1.00

Note: 1. The number of readers is the number of readers who have at least one bias gap because some readers did not report their subjective bias(es) for newspaper(s) they read; 2. The subjective bias is an average number of all individual subjective biases.

Table 2 presents the distribution of right-wing and non-right-wing newspaper consumption among naive and sophisticated readers. The right-wing newspapers clearly dominate the press as 88.4% of readers at least consume one right-wing newspaper while this percentage for left-wing newspapers is 31.1%. Naive readers outnumber sophisticated readers, accounting for 63.8% of the sample. Moreover, among right-wing newspaper readers, naive ones are almost twice as many as sophisticated ones, while the sets of readers are evenly distributed in the other group.

Table 2: Newspaper Readership by Readers' Political Leaning and Naivety

	Right-wing readership	Left-wing readership	Full sample
Naive	2134	740	2342
Sophisticated	1110	402	1326
Total	3244	1142	3668

Note: 1. Right-wing readership is the number of readers who read at least one right-wing newspaper among the 10 newspapers studied, and similarly, left-wing readership is the opposite; 2. The naive measure used here is naive\_c as it is binary.

Table 3 presents the summary statistics for all variables used in the econometric analyses, including their mean, median, standard deviation, maximum, and minimum.

Table 3: Summary Statistics

Variable	mean	median	sd	min	max
Political leaning	4.195	4.000	1.394	1.000	7.000
Right-wing newspaper consumption	7.552	7.000	5.348	0.000	54.000
Left-wing newspaper consumption	2.141	0.000	3.656	0.000	35.000
Naive	0.555	0.667	0.455	0.000	1.000
Naive_d	0.557	0.696	0.460	0.000	1.000
Naive_f	0.579	1.000	0.494	0.000	1.000
Naive_c	0.638	1.000	0.481	0.000	1.000
Regional Internet speed	80.408	84.074	20.241	14.930	111.684
Male	0.542	1.000	0.498	0.000	1.000
Age	53.736	55.000	17.938	0.000	94.000
White	0.861	1.000	0.346	0.000	1.000
Household income	34.887	27.500	22.938	2.250	100.000
Education	0.311	0.000	0.463	0.000	1.000
Online interview	0.622	1.000	0.485	0.000	1.000
Interested in political news	0.883	1.000	0.322	0.000	1.000

Note: t-statistics in parentheses. The unit for regional Internet speed is Mbit/s. Household income is in 1,000 pounds.

#### 4.2. Naivety and demographics

The naivety measure in this thesis is a novelty. It can be useful to examine whether certain demographic characteristics can be associated with people's naivety. The econometric model used is shown below:

$$naive_i = \alpha_0 + \alpha_1 male_i + \alpha_2 age_i + \alpha_3 white_i + \alpha_4 education_i + \alpha_5 household\_income_i + \alpha_6 int\_polnews_i + \varepsilon_i$$

In the model above, male equals 1 if the individual is male and equals 0 otherwise, white equals 1 if the individual is white and equals 0 otherwise, and int\_polnews equals 1 if the individual is interested in political news and equals 0 otherwise. If the individual has a college degree or equivalent, education equals 1, and it equals 0 otherwise. The household\_income is the total income of the household to which the individual belongs.

Table 4: Naivety and Demographics

	naive	naive_d	naive_f	naive_c
male	-0.013 (0.78)	-0.013 (0.81)	0.008 (0.42)	-0.003 (0.15)
age	0.001 (1.74)	0.001 (1.47)	0.000 (0.58)	-0.001 (1.09)
white	-0.051 (2.15)	-0.051 (2.11)	-0.075 (2.72)	-0.095 (3.73)
education	-0.105 (5.55)	-0.114 (5.92)	-0.114 (5.42)	-0.069 (3.38)
household income	-0.001 (2.08)	-0.001 (2.15)	-0.001 (1.68)	-0.001 (1.41)
interested in political news	0.061 (2.09)	0.061 (2.09)	0.058 (1.91)	0.093 (3.09)
Subregion fixed effects	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES
<i>N</i>	3,086	3,086	3,086	3,086

Note: t-statistics in parentheses.

The results of a simple linear regression shown in Table 4 demonstrate that the determinants of naivety are consistent across the four measures. In terms of the first three inherent demographics, gender is uncorrelated with naivety, while age is slightly correlated but the coefficients are very small. Being white is negatively correlated with naivety. What is more interesting is the last three variables. Education and household

income are negatively correlated with naivety, indicating that a more educated person will less likely be naive and that those who live in a poor family are more likely to be naive than those from a wealthy family. Whether an individual is interested in news or not is positively correlated with naivety, which conveys a pessimistic message that those who are interested in political news and thus very likely to engage in politics tend to be naiver than others. Therefore, it might be important to examine whether newspaper consumption polarizes naive readers more than sophisticated readers.

### **4.3. Newspaper consumption, naivety, and political polarization**

This section studies how readers' naivety changes the extent to which newspaper consumption can polarize them. The political ideologies of naive readers can be either more malleable or more rigid than sophisticated ones, which might produce heterogeneous polarizing effects of newspaper consumption on their ideologies. To capture this difference an interaction term between consumption and naivety is needed. Therefore, the econometric model takes the form

$$leaning_i = \alpha_0 + \alpha_1 cons_i + \alpha_2 naive_i + \alpha_3 cons_i \times naive_i + \alpha_4 X_i + \varepsilon_i$$

where  $leaning_i$ ,  $cons_i$ ,  $naive_i$  are individual  $i$ 's political leaning, right-wing newspaper consumption, and degree of naivety, respectively;  $X_i$  is a set of control variables of individual  $i$ , including  $i$ 's left-wing newspaper, gender, age, race, education, as well as an important indicator  $online_i$  that equals 1 if the respondent chose to take the survey online and equals 0 if she chose the face-to-face interviews. According to Ofcom News Consumption Survey: Technical Report for Adults 2020, it was noted that online participants were more likely to have a larger range of interests and a greater number of technology devices in their households. These differences between two

groups of respondents can bias the regressions so it is necessary to include it in the regression.

The key variable in this model is the interaction term. There are three hypotheses to be tested in terms of the sign of its coefficient.

- (1)  $\alpha_3 > 0$ . The rationale behind it is that naive readers are prone to be polarized by newspapers as they might easily believe what these media outlets say without critical thinking. Thus, their ideology is malleable.
- (2)  $\alpha_3 < 0$ . The theory underlying this hypothesis is that naive readers are more stubborn than sophisticated readers. Therefore, their ideology is rigid.
- (3)  $\alpha_3 = 0$ . This hypothesis claims that the effect of consuming right-wing newspapers on political ideology does not depend on people's naivety.

Table 5 shows the results of OLS regressions.

Table 5: OLS Regressions

Political leaning	naive	naive_d	naive_f	naive_c
Right-wing newspaper consumption	0.041 (2.76)	0.044 (2.95)	0.053 (3.46)	0.063 (4.32)
Naivety	0.264 (2.76)	0.268 (2.84)	0.284 (3.32)	0.347 (4.06)
Right-wing newspaper consumption×Naivety	-0.013 (1.08)	-0.016 (1.35)	-0.024 (2.38)	-0.035 (3.22)
Left-wing newspaper consumption	-0.068 (4.35)	-0.068 (4.34)	-0.067 (4.28)	-0.066 (4.17)
Male	0.058 (1.22)	0.057 (1.20)	0.055 (1.14)	0.053 (1.10)
Age	0.005 (3.07)	0.005 (3.07)	0.005 (3.06)	0.005 (2.95)
White	0.152 (2.00)	0.149 (1.97)	0.147 (1.94)	0.149 (1.96)
Education	-0.132 (2.31)	-0.132 (2.30)	-0.132 (2.31)	-0.132 (2.32)
Online interview	0.231 (4.16)	0.232 (4.17)	0.228 (4.10)	0.224 (4.03)
Newspaper fixed effects	YES	YES	YES	YES
Subregion-year fixed effects	YES	YES	YES	YES
<i>N</i>	3,095	3,095	3,095	3,095

Note: t-statistics in parentheses. The dependent variable is leaning and all those naive\_\* measures are referring to the variable used for naive on the right-hand side. The subregion-year fixed effects here are done by using an interaction term between subregion and year factor variables along with their main effects.

The results of the four regressions shown above are similar. It is notable that the coefficient of naive is very big in magnitude, indicating a strong positive correlation between naivety and right-wing leaning. This shows that naive newspaper readers are more right-wing than sophisticated readers. The coefficient of the variable of interest supports the second hypothesis, indicating that naive readers get less polarized to the right than sophisticated ones given the same amount of right-wing newspaper consumption. With their rigidity of ideology, naive readers who consume right-wing newspapers will turn to be more right-wing slightly, but their ideology is fairly right-



wing compared to sophisticated readers in the first place. Their ideology will only shift to the left very slowly even if they consume fewer right-wing newspapers.

Given that right-wing newspapers rule the British press and that about 64% of readers are naive (shown in section 4.1), it could be the case that people inevitably tend to be more right-wing and the majority of them get fixated to the right due to their naivety. If naive people are more likely to engage in politics, which is to some extent implied by the finding in section 4.2 that they are more likely to be interested in political news, then rightists could dominate democracy.

With respect to control variables, gender has no significant impact while age exerts a negligible impact on people's ideology. Whites and less educated people are more right-wing than others.

All arguments made above are based on correlations instead of causalities. Clearly, right-wing newspaper consumption is endogenous as it shapes readers' ideology and readers choose this consumption based on their ideology. Ideally, if we assume that naivety is an exogenous trait conditional on the control variables and there is an exogenous shock to the consumption, then it is possible to establish a causal relationship between right-wing newspaper consumption, naivety, and political ideology.

The plausibly exogenous shock used in this thesis is the Internet speed in the region where the individual lives. The regional Internet speed, denoted by  $speed$ , is negatively correlated with print newspaper consumption conditional on basic demographics. The advent of the Internet spawned websites, applications, and social

media for news, which greatly hit the press and many newspapers thus developed their websites or mobile applications. In any case, the increasing Internet speed shifts people from print newspapers to either online newspapers or other platforms for news. Conditional on some basic demographics, as well as other regional factors—for example, Liverpool boycotts the Sun due to the Hillsborough disaster—that can potentially affect people’s ideology, regional Internet speed should plausibly be uncorrelated with the error term.

The first stage depicts how newspaper consumption varies with regional Internet speed and how the interaction term between regional speed and naivety influences the interaction term between newspaper consumption and naivety. As two instruments are used, the two corresponding econometric models are shown below:

$$cons_i = \beta_0 + \beta_1 speed_i + \beta_2 naive_i + \beta_3 speed_i \times naive_i + \beta_4 X_i + \varepsilon_i$$

$$cons_i \times naive_i = \gamma_0 + \gamma_1 speed_i + \gamma_2 naive_i + \gamma_3 speed_i \times naive_i + \gamma_4 X_i + \varepsilon_i$$

Table 6 and Table 7 presents the first-stage estimates for these two models, respectively.

Table 6: First-Stage Regressions A

Right-wing newspaper consumption	naive	naive_d	naive_f	naive_c
Regional Internet speed	-0.018 (1.43)	-0.018 (1.47)	-0.019 (1.51)	-0.017 (1.41)
Naivety	0.589 (2.13)	0.511 (1.87)	0.412 (1.47)	0.388 (1.37)
Regional Internet speed ×Naivety	-0.000 (0.04)	0.001 (0.27)	0.002 (0.61)	-0.000 (0.09)
Left-wing newspaper consumption	0.183 (4.02)	0.183 (4.03)	0.182 (4.00)	0.188 (4.08)
Male	0.317 (4.16)	0.318 (4.17)	0.312 (4.09)	0.315 (4.11)
Age	0.007 (2.79)	0.007 (2.83)	0.007 (2.84)	0.007 (2.82)
White	-0.046 (0.14)	-0.047 (0.33)	-0.044 (0.31)	-0.051 (0.36)
Education	0.001 (0.01)	0.003 (0.04)	0.008 (0.08)	-0.008 (0.09)
Online interview	-0.033 (0.43)	-0.026 (0.35)	-0.018 (0.23)	-0.036 (0.48)
Newspaper fixed effects	YES	YES	YES	YES
Subregion-year fixed effects	YES	YES	YES	YES
F-test statistics	1.07	1.07	1.20	1.01
<i>N</i>	3,095	3,095	3,095	3,095

Note: t-statistics in parentheses. The dependent variable is right-wing newspaper consumption and all those naive\_\* measures are referring to the variable used for naive on the right-hand side. F-test statistics are also reported and in general, it should be over 10 when the instrument is valid.

Table 7: First-Stage Regressions B

Right-wing newspaper consumption×Naivety	naive	naive_d	naive_f	naive_c
Regional Internet speed	-0.036 (2.45)	-0.037 (2.48)	-0.053 (3.05)	-0.042 (2.59)
Naivety	5.990 (16.89)	5.953 (16.93)	6.183 (16.03)	6.080 (15.53)
Regional Internet speed ×Naivety	0.012 (2.70)	0.013 (2.82)	0.018 (3.70)	0.012 (1.78)
Left-wing newspaper consumption	0.152 (2.28)	0.131 (1.95)	0.147 (1.86)	0.189 (2.63)
Male	0.190 (2.02)	0.170 (1.80)	0.193 (1.71)	0.127 (1.25)
Age	-0.001 (0.35)	-0.001 (0.34)	0.000 (0.00)	-0.002 (0.66)
White	-0.052 (0.29)	-0.132 (0.71)	-0.120 (0.53)	-0.083 (0.42)
Education	0.183 (1.51)	0.196 (1.60)	0.206 (1.43)	0.154 (1.14)
Online interview	-0.097 (1.05)	-0.111 (1.17)	-0.191 (1.74)	-0.237 (2.35)
Newspaper fixed effects	YES	YES	YES	YES
Subregion-year fixed effects	YES	YES	YES	YES
F-statistics	5.97	6.35	9.79	4.17
<i>N</i>	3,095	3,095	3,095	3,095

Note: t-statistics in parentheses. The dependent variable is the interaction term between right-wing newspaper consumption and naivety and all those naive\_\* measures are referring to the variable used for naive on the right-hand side. F-test statistics are also reported and in general, it should be over 10 when the instrument is valid.

The results above indicate that regional Internet speed is negatively correlated with newspaper consumption. However, this correlation is weak, and the two F tests of excluded instruments demonstrate that the two instruments used might be weak instruments as a valid instrument generally has a F-statistic exceeding 10. Nevertheless, we can proceed with this flaw and turn to the second stage. To begin with, Table 8 presents the results for the reduced form.

Table 8: Reduced-Form Regressions

Political leaning	naive	naive_d	naive_f	naive_c
Regional Internet speed	0.014 (1.73)	0.013 (1.69)	0.013 (1.65)	0.014 (1.76)
Naivety	0.392 (2.05)	0.335 (1.79)	0.220 (1.24)	0.344 (1.91)
Regional Internet speed ×Naivety	-0.002 (1.07)	-0.002 (0.86)	-0.001 (0.53)	-0.003 (1.20)
Left-wing newspaper consumption	-0.061 (3.87)	-0.061 (3.85)	-0.061 (3.80)	-0.059 (3.73)
Male	0.071 (1.49)	0.071 (1.49)	0.069 (1.45)	0.070 (1.47)
Age	0.005 (3.31)	0.005 (3.32)	0.005 (3.32)	0.005 (3.33)
White	0.148 (1.94)	0.147 (1.93)	0.145 (1.90)	0.145 (1.90)
Education	-0.134 (2.34)	-0.134 (2.35)	-0.136 (2.37)	-0.138 (2.41)
Online interview	0.232 (4.16)	0.233 (4.19)	0.234 (4.19)	0.229 (4.11)
Newspaper fixed effects	YES	YES	YES	YES
Subregion-year fixed effects	YES	YES	YES	YES
<i>N</i>	3,095	3,095	3,095	3,095

Note: t-statistics in parentheses.

Although the interaction term is not significant, the coefficient of speed is informative. The positive correlation between regional Internet speed and right-wing leaning indicates that increasing Internet speed drives people to the right of the political spectrum. Faster Internet speed could mean that people will be more likely to consume social media, websites, and mobile applications for news. Then, it might be the case that like the press, online platforms are also dominated by right-wing forces. This can have broad implications in terms of understanding the media landscape in the UK.

Next, we present the second-stage regression results in Table 9.

Table 9: Second-Stage Regressions

Political leaning	naive	naive_d	naive_f	naive_c
Right-wing newspaper consumption	-0.372 (0.53)	-0.513 (0.60)	-0.792 (0.67)	-0.149 (0.16)
Naivety	-0.202 (0.85)	-0.116 (0.42)	0.029 (0.10)	-0.273 (0.85)
Right-wing newspaper consumption × Naivety	1.819 (1.25)	1.288 (0.80)	0.362 (0.22)	2.057 (1.10)
Left-wing newspaper consumption	0.037 (0.31)	0.048 (0.34)	0.080 (0.42)	0.020 (0.15)
Male	0.227 (1.06)	0.253 (1.00)	0.310 (0.92)	0.151 (0.56)
Age	0.008 (1.39)	0.009 (1.32)	0.011 (1.26)	0.006 (0.77)
White	0.121 (1.03)	0.108 (0.87)	0.114 (0.82)	0.115 (1.05)
Education	-0.097 (1.09)	-0.110 (1.14)	-0.136 (1.29)	-0.097 (1.08)
Online interview	0.200 (2.73)	0.207 (2.70)	0.225 (2.49)	0.159 (1.85)
Newspaper fixed effects	YES	YES	YES	YES
Subregion-year fixed effects	YES	YES	YES	YES
<i>N</i>	3,095	3,095	3,095	3,095

Note: t-statistics in parentheses.

Since the instruments are weak, the second-stage results are not significant. A better instrument could still be regional Internet speed but with more variations in time and regions, for example, postcode-level Internet speeds for the last 10 years. The current Internet speed variable only covers 24 regions and 3 years, so these variations might be not strong enough to produce sufficient variations in newspaper consumption. However, a challenge of using Internet speed as instrument is that there might be other channels that can potentially confound the results. In the analyses, the consumption of other media platforms, including television, radio, social media, etc., is not controlled, which means that internet speed can affect political leaning through channels other than

newspaper consumption, violating the assumption of a valid instrument. But can it be resolved by controlling these channels? It depends on the exact mechanism of how internet speed affects newspaper consumption. If not controlling other channels, we are allowing substitution effect that people shift from newspapers to other channels due to higher Internet speed. If controlling other channels, people are only allowed to decrease newspaper consumption while holding other consumption constant, which should be less likely the case. Therefore, Internet speed might not be a good choice in the first place. Overall, the IV method is not a successful attempt, but it can still shed some light on future research.

## **5. Conclusion**

This thesis utilizes respondent-level survey data to explore the relationship between news media consumption and political polarization in the UK. Unlike the existing literature that often assumes that individuals are homogeneous in their degree of sophistication, this research introduces the concept of naivety and provides four ways of measuring it. By incorporating reader's naivety into the analyses, it probes which characteristics of readers are associated with their naivety and investigates whether British right-wing newspapers have heterogeneous polarizing effects on the ideologies of the naive and sophisticated readers. As newspaper consumption and readers' ideology might have a reverse causality issue, instrumental variable, regional Internet speed, is used in the 2SLS regression. However, as the instrument is generally not satisfying, the second-stage estimates are not significant.

The media landscape in the UK has a strong right-wing leaning: most national

newspapers are right-wing and the majority of readers read right-wing newspapers. In terms of naivety, it is negatively associated with education, indicating that a more educated person should be less naive, and positively associated with the interest in political news. The latter might imply that those who are interested in political news and thus engaging in politics are more likely to be naive. This conjecture reflects the importance of examining whether the political ideologies of naive readers, who constitute 64% of the reader group of the 10 national newspapers, are relatively rigid or malleable under the influence of right-wing newspaper consumption. The results show that naive readers get less polarized to the right than sophisticated ones given the same amount of right-wing newspaper consumption. This finding could imply that the situation in the UK where rightists and right-wing newspapers greatly outnumber the left-wing forces could continue because naive right-wing readers tend to stick to their ideology.

Future research could go in many ways. Most obviously, the instrument employed is too weak to give significant results. Ideally, harnessing more data over time—for example, Internet availability at the postcode level—should be able to provide satisfying first-stage results and thus improve the efficiency of the IV estimation. It might be useful to examine if the results would be the same for television viewers as TV channels are the dominant media outlets for news in the UK. Social media could be another candidate and the analyses could be applied to other countries.



## 6. Reference

- [1] Baysan, C., 2021. Persistent Polarizing Effects of Persuasion: Experimental Evidence from Turkey. Unpublished.
- [2] Bernhardt, D., Krasa, S., Polborn, M., 2008. Political polarization and the electoral effects of media bias. *Journal of Public Economics*. 98: 1092–104.
- [3] Chan, J., Suen, W., 2008. A spatial theory of news consumption and electoral competition. *The Review of Economic Studies*. 75: 699–728.
- [4] Chiang, C., Knight, B., 2011. Media bias and influence: evidence from newspaper endorsements. *Review of Economic Studies*. 78: 795–820.
- [5] DellaVigna, S., Kaplan, E. 2007., The Fox News effect: media bias and voting. *Quarterly Journal of Economics*. 122: 1187–234.
- [6] Gentzkow, M., and Shapiro, J.M., 2006. Media Bias and Reputation. *Journal of Political Economy* 114 (2): 280–316.
- [7] Gentzkow, M., Shapiro, J.M., 2010. What drives media slant? Evidence from US daily newspapers. *Econometrica* 78 (1): 35–71.
- [8] Groseclose, T., and Jeffrey M., 2005. A Measure of Media Bias. *Quarterly Journal of Economics* 120 (4): 1191–237.
- [9] Levy, R., 2021. Social Media, News Consumption, and Polarization: Evidence from a Field Experiment. *American Economic Review* 111 (3): 831-70.
- [10] Martin, G.J., and Yurukoglu, A. 2017. Bias in Cable News: Persuasion and Polarization. *American Economic Review*. 107 (9): 2565–99.
- [11] Qin, B., Strömberg, D., Wu, Y., 2018. Media bias in China. *American Economic Review*. 108 (9): 2442–2476.
- [12] Strömberg, D., 2015. Media and Politics. *Annual Review of Economics* 7 (1): 173-205.