README - Social Media, News Consumption, and Polarization: Evidence from a Field Experiment (openicpsr-121922)

Ro'ee Levy

November 13, 2020

1 Data

1.1 Data Availability Statements

1. **Experiment data** - The vast majority of the analysis is based on data collected in the experiment.
   - *Access and sources*: To ensure the privacy of participants and in compliance with Yale's IRB Protocol (HSC # 2000021422), access to this data is restricted. The data can be made available only for replication purposes. To access the data please see (Levy, 2020). Eligible replicators can email the author: roeelevy@mit.edu for permission.
   - *Final datasets*: The following datasets include the experiment data (the datasets are not included in this repository but they are described below):
     - `Datasets/Extension`
     - `Datasets/Likes`
     - `Datasets/Shared`
     - `Datasets/Survey`
     - `Datasets/UsersPages`
2. **Comscore data** (Comscore, 2018) - I analyze the 2007, 2008, 2017, and 2018 Comscore WRDS Web Behavior Database Panel.
   - *Access and sources*: The data was obtained through Wharton Research Data Services. The author had access to this data as a Yale researcher. While the data is not available for public use, several universities provide access to this data. It is also possible to purchase the data from Comscore. If a university subscribes to this dataset it can be downloaded by logging in to https://wrds-www.wharton.upenn.edu/ and downloading all observations for the Comscore 2007, 2008, 2017, and 2018 demographics and sessions data (transactions data is not needed)

     The WRDS Comscore demographics data should be stored in:
     - `Input/Comscore/[year]/demographics[year].csv`

     For example:
     - `Input/Comscore/18/demographics18.csv`

     The sessions files should be stored in:
     - `Input/Comscore/[year]/sessions[year].zip`

     The 2007 zip is very large and thus should be stored as four separate files:

- `Input/Comscore/07/sessions07Q1.zip`
- `Input/Comscore/07/sessions07Q2.zip`
- `Input/Comscore/07/sessions07Q3.zip`
- `Input/Comscore/07/sessions07Q4.zip`

For non-windows users the zip files should first be uncompressed to .csv files.

- *Additional Processing*:
  The following files process the Comscore data:
  - `Code/Process/External/importComscore.R`
  - `Code/Process/External/processComscore.R`
- *Final datasets*: The R files mentioned above will import the Comscore datasets, process, append the data for all years and create one dataset with all news sites visited called (the dataset is not included in this repository):
  - `Datasets/Comscore/comScoreNews.rds`

3. **US Demographics** - For reweighting some results I use survey data on the US population.
   - *Access and sources*: This data is based on the following publicly available online sources:
     (a) Pew American Trends Panel Wave 1 (Pew Research Center, 2014).
     (b) The Pew American Trends Panel Wave 23 (Pew Research Center, 2016).
     (c) The 2016 American National Election Survey (ANES) (The American National Election Studies, 2016).
     (d) The 2018 Pew Core Trends Survey (Pew Research Center, 2018).
   - *Final datasets*: The relevant variable means are stored in
     - `Datasets/externalSurveys`
4. **Media slant** - To measure the slant of media outlets, I use a dataset of the slant of 500 leading news-related domains (Bakshy, Messing and Adamic, 2015*b*).
   - *Access and sources*: The data is based on a paper by Bakshy, Messing and Adamic (2015*a*). I use the file 'top500', which should be stored in the following location:
     - `Input/Media/top500.csv`

     I also collected the Facebook Page ID of the domains. The data is stored here:
     - `Datasets/Media/raw/top500PageID.rds`
   - *Additional Processing*: The file is imported and processed in:
     - `Code/Import/importTop500.R`
     - `Code/Process/processTop500.R`
   - *Final dataset*s: The slant variable is stored in:
     - `Datasets/Media/top500Outlets.rds`
5. **Facebook posts shared by outlets** - I collected data on posts shared by outlets included in the experiment and on the URLs appearing in the posts (Levy, 2018).
   - *Access and sources*: There are no restrictions on accessing the data.
   - *Final datasets*: The relevant datasets are stored in:
     - `Datasets/News`
6. **FEC donation data** - One figure in the paper uses the share of Republican donations by zip code. I calculated these figures using FEC data.

- *Access and sources*: The data is based on the FEC 2015-2016 and 2017-2018 donation data (Federal Election Commission, 2018). The following datasets are used: Candidate Master, Committee Master, and Contributions by individuals, for the 2017-2018, 2015-2016 election cycles.

  The files should be downloaded, unzipped, renamed, and stored in the following location:

  - `Input/Donations/Ind/indivHeader.csv`
  - `Input/Donations/Ind/indiv16.txt`
  - `Input/Donations/Ind/indiv18.txt`
  - `Input/Donations/Cand/candMasterHeader.csv`
  - `Input/Donations/Cand/candMaster16.txt`
  - `Input/Donations/Cand/candMaster18.txt`
  - `Input/Donations/Committee/cmMasterHeader.csv`
  - `Input/Donations/Committee/cmMaster16.txt`
  - `Input/Donations/Committee/cmMaster18.txt`

  When downloading the contribution by individuals data, only the file itcont.txt is needed and should be renamed to `indiv16.txt` or `indiv18.txt`. The folder 'by_date' is not used.

  Note that the processed donation data is stored in the repository and therefore the source data is not required for any analysis.
- *Additional Processing*: The following files imports and processes the donation data:
  - `Code/Import/importDonations.R`
  - `Code/Process/processDonations.R`
- *Final datasets*: The donations data is stored in:
  - `Datasets/Donations/donationsZip.rds`

1.2 Dataset list

The following table describes each dataset. Each dataset also has a corresponding CODEBOOK.csv file which describes the variables. The codebook is stored in the same location as the dataset. The variable descriptions can also be accessed in R using the attributes of the dataset object. All the datasets are stored as .rds files with the exception of ext_facebook.fst, which is much larger and thus saved in .fst format and can be loaded using the fst package.

As explained above, the experiment datasets are described but not included in this repository since access to these datasets is restricted.

Dataset	Directory	Description	Source	Access
donationsZip.rds	Donations	Share of Republican Donations by Zip Code	FEC Donation Data	Available
slant.rds	Extension	Slant of different types of news participants engaged with, each observation is one news page/site/post liked, visited, observed, or shared	Experiment data	Restricted
ext_facebook.fst	Extension/ facebook	Posts observed by participants who installed the extension	Experiment data	Restricted

(continued)

Dataset	Directory	Description	Source	Access
fb_user.rds	Extension/ facebook	Aggregate data on posts seen in the feed in the 14 days following the intervention	Experiment data	Restricted
fb_user_day.rds	Extension/ facebook	Posts observed in the Facebook feed by day and participant	Experiment data	Restricted
fb_user_7.rds	Extension/ facebook	Aggregate data on posts seen in the feed in the 7 days following the intervention	Experiment data	Restricted
nav_user.rds	Extension/ navigate	Aggregate data on websites visited by participants	Experiment data	Restricted
nav_user_day.rds	Extension/ navigate	Websites visited by participant by day	Experiment data	Restricted
nav_user_fb_usage.rds	Extension/ navigate	Baseline visits to Facebook by participants	Experiment data	Restricted
usPop.rds	External-Surveys	Descriptive statistics of Americans	2016 ANES and Pew American Trends Panel Wave 23	Available
usPopIdeo.rds	External-Surveys	Descriptive statistics of Americans who have an ideological leaning	2016 ANES and Pew American Trends Panel Wave 23	Available
usFbPew.rds	External-Surveys	Descriptive statistics of American Facebook users	Pew American Trends Panel Wave 1 (March-April 2014) and the 2018 Pew Core Trends Survey	Available
likes_user.rds	Likes	Pages participants subscribed to two weeks after the intervention	Experiment data	Restricted
baselineLikes.rds	Likes	Pages participants liked in baseline. Each observation is a participant and each column is a popular Facebook page	Experiment data	Restricted
media.rds	Media	Media outlets offered in the experiment	Public data	Available
top500PageID.rds	Media/ raw	Facebook Page ID of leading outlets	Data collected by the author	Available
newsPosts.rds	News	Posts shared by the Facebook pages of outlets included in the experiment	Data collected by the author	Available

(continued)

Dataset	Directory	Description	Source	Access
newsEntities.rds	News	All entities appearing in posts shared by news outlets included in the experiment, Feb. 28-Apr. 25	News posts messages annotated using Spacy Natural Language Processing algorithm	Available
articlesSection.rds	News	Sections in news websites where articles appear	Data collected by the author	Available
shared_user.rds	Shared	Aggregate data on posts shared by participants	Experiment data	Restricted
shared_user_day.rds	Shared	Posts shared by participant by day	Experiment data	Restricted
sharedPosts.rds	Shared	Posts shared by participants following the intervention	Experiment data	Restricted
survey.rds	Survey	Main dataset, baseline and endline variables along with other data on the intervention and participants	Experiment data	Restricted
users_pages.rds	UsersPages	Variables at the participant by outlet level	Experiment data	Restricted

2 Code

2.1 Instructions

- Set the current working directory to the project's home directory
- Run master.R

Note: Most of the code relies on restricted-access data sets. To access these datasets see the data available statements above.

2.2 Computational requirements

2.2.1 Software Used

- R 3.5.1
  - broom (0.7.0), car (3.0-7), data.table (1.13.0), dplyr (0.8.5), ebal (0.1-6), fst (0.9.2), ggplot2 (3.3.2), ggpubr (0.2.5), grf (1.1.0), Hmisc (4.4-0), kableExtra (1.2.1), lfe (2.8-5), lmtest (0.9-37), logging (0.10-108), multiwayvcov (1.2.3), quanteda (2.0.1), sandwich (2.5-1), scales (1.1.1), stargazer (5.2.2), tidyr (1.0.2)
  - the file "`setup.R`" will install all dependencies (latest version), and should be run once prior to running other programs.

2.2.2 Memory and Runtime Requirements

The code was last run on a 4-core Intel I7 64 bit system with 16 GB RAM. Excluding package installation, the code took approximately 70 minutes to run.

2.3 Description of scripts

- The file master.R calls all the scripts.
- The scripts in `Code/Import` import external raw data.
- The scripts in `Code/Process` process the data and create datasets used in the analysis.
- The scripts in `Code/Analyze` analyze the data and create the tables and graphs.
  - All output will be stored in the output folder: `Output/`
    Note that the code also creates additional tables and figures which are not included in the paper but are presented in talks.
- Code in `Code/Utils` provides useful function when analyzing the data

The following table provides details on each R script.

File	Directory	Description	Tables and graphs created
analyzeControlSlant	Code /Analyze /Media	Provide descriptive statistics based on the control group	Figure 4, Table A.7b, Table A.9b
analyzeFB	Code /Analyze /Media	Provide summary statistics on exposure to posts observed in the feed	Figure A.6
analyzeFBContent	Code /Analyze /Media	Analyze the content of Facebook posts, and test whether specific news stories were mentioned	Table A.27
analyzeMediaBehavior	Code /Analyze /Media	Analyze effects on engagement with potential outlets: subscriptions to outlets, exposure to posts in the feed, websites visited, and posts shared	Figure 6, Figure A.2, Figure A.5, Figure A.7, Table A.10
analyzeMediaBehavior-Hetero	Code /Analyze /Media	Analyze heterogeneous effects on engagement with potential outlets: subscriptions to outlets, exposure to posts in the feed, websites visited, and posts shared	Figure A.18, Figure A.19
analyzeMediaBehaviorSlant	Code /Analyze /Media	Analyze the effects on the slant of all news participants engaged with	Figure 7, Figure A.8, Table A.11, Table A.23
analyzeMediaPersistence	Code /Analyze /Media	Analyze the effects of the intervention on media outcomes over time	Figure 8a, Figure 8b, Figure A.9a, Figure A.9b

(continued)

File	Directory	Description	Tables and graphs created
analyzeNewsAnnotation	Code /Analyze /Media	Plot names that were mentioned most often by the news outlets	Figure 2
analyzePostContent	Code /Analyze /Media	Analyze text in posts participants were exposed to in their feed, posts they clicked, and posts they shared	Figure A.14, Figure A.15, Figure A.16, Figure A.17, Table A.22a, Table A.22b, Table A.22c
decomposeExposure	Code /Analyze /Media	Decompose effect on exposure to posts from the pro- and counter-attitudinal outlets offered in the experiment	Figure 10, Figure A.21, Table A.28
analyzeKnowledge	Code /Analyze /Surveys	Analyze effect on endline survey knowledge outcomes	Table A.26
analyzeOtherOutcomes	Code /Analyze /Surveys	Analyze additional outcomes from the endline survey	Figure A.4
analyzeOutcomesUserPage	Code /Analyze /Surveys	Analyze effects at the outlet by participant level including whether an outlet was seen, is trusted, and its perceived slant	Figure A.3
analyzePersuasion	Code /Analyze /Surveys	Analyze the effects on the political opinions outcomes	Figure A.10, Table A.12a
analyzePolarization	Code /Analyze /Surveys	Analyze the effects on the affective polarization outcomes	Figure A.12, Table A.12b, Table A.13, Table A.17, Table A.19
analyzePrimaryOutcomes	Code /Analyze /Surveys	Analyze effect on the political opinions and polarization indices	Figure 9, Figure A.11, Figure A.20, Table A.14a, Table A.14b, Table A.15a, Table A.15b, Table A.18a, Table A.18b, Table A.24a, Table A.24b
analyzeSurveyPurpose	Code /Analyze /Surveys	Analyze the perceived objective of the baseline survey	Table A.20a, Table A.20b, Table A.22a, Table A.22b, Table A.22c
analyzeCompliance	Code /Analyze	Analyze compliance with the intervention	Table 3, Table A.1
analyzeHeteroForest	Code /Analyze	Analyze heterogenous effects using random forests	Table A.25
analyzePolarization-WithMedia	Code /Analyze	Use extension and endline survey data to estimate the effect of exposure to news on affective polarization	Table A.16a, Table A.16b
balanceTables	Code /Analyze	Create balance tables and additional descriptive statistics of participants in the study	Table 2, Table A.2, Table A.3, Table A.4, Table A.5, Table A.6
comscoreDemonstrate	Code /Analyze	Plot news consumption descriptive statistics based on Comscore data	Figure 5a, Figure 5b

(continued)

File	Directory	Description	Tables and graphs created
comscoreSegregation	Code /Analyze	Create tables providing descriptive statistics on news consumption based on Comscore data	Figure 3, Table A.7a, Table A.8a, Table A.8b, Table A.9a
importComscore	Code /Import	Import 2007, 2008, 2017, 2018 Comscore demographic and browsing data	
importDonations	Code /Import	Import FEC donation data for 2016, 2018 election cycles	
importTop500	Code /Import	Import outlet dataset	
processComscore	Code /Process	Process Comscore data and save one dataset for all news sites across years	
processDonations	Code /Process	Process donation data to determine ideology by zipcode	
processTop500	Code /Process	Process dataset of leading news outlets	
causalForestFunctions	Code /Utils	Functions to run causal forest analysis	
commonFunctions	Code /Utils	Two common functions: define a new factor variable and set the label attributes	
commonVars	Code /Utils	Common vars and basic functions used in many files, mostly for variable names and creation of plots	
indexFunctions	Code /Utils	Useful functions when creating an outcome index	
log	Code /Utils	Functions to generate logs	
robustReg	Code /Utils	Functions to find robust standard errors, print regression tables with robust standard errors, find differences between coefficients	
segregationFunctions	Code /Utils	Functions used to create news consumption descriptive stat.: isolation, segregation, congruence, share of counter-attitudinal news	
textAnalysis	Code /Utils	Userful functions to process text throughout the project	
utilURL	Code /Utils	Useful functions when working with URLs	
master	Code	Master file to create all tables and graphs	
setup	Code	Setup R environment	

2.4 Descriptions of figures and tables

The following table provides details on each figure and table. Figures and tables starting with A. appear in the appendix.

Figure / Table	Source Code	Output Dir	Output File
Figure 1	Not generated by code		
Figure 2	Analyze/Media/ analyzeNewsAnnotation.R	News	Figure2_NamesPrimaryOutlets
Figure 3	Analyze/comscoreSegregation.R	Descriptive/Comscore	Figure3_segregationMeasureBySource
Figure 4	Analyze/Media/ analyzeControlSlant.R	Descriptive/Extension	Figure4_isolationMeasureByMedium
Figure 5a	Analyze/comscoreDemonstrate.R	Descriptive/Comscore	Figure5a_comScoreSegregation
Figure 5b	Analyze/comscoreDemonstrate.R	Descriptive/Comscore	Figure5b_comScoreIdeologyDonations
Figure 6	Analyze/Media/ analyzeMediaBehavior.R	Combined	Figure6_ITT_FullDataMatch
Figure 7	Analyze/Media/ analyzeMediaBehaviorSlant.R	Combined	Figure7_slantScaleFullData
Figure 8a	Analyze/Media/ analyzeMediaPersistence.R	Persistence	Figure8a_PersistentEffectSlantScaleExt
Figure 8b	Analyze/Media/ analyzeMediaPersistence.R	Persistence	Figure8b_PersistentEffectSlantScaleShared
Figure 9	Analyze/Surveys/ analyzePrimaryOutcomes.R	Combined	Figure9_PrimaryOutcomes
Figure 10	Analyze/Media/ decomposeExposure.R	Exposure	Figure10_decomposition
Figure A.1	Not generated by code		
Figure A.2	Analyze/Media/ analyzeMediaBehavior.R	Exposure	ExposureFriendsAdsPages
Figure A.3	Analyze/Surveys/ analyzeOutcomesUserPage.R	AdditionalSurvey-Outcomes	outcomesUserOutletShort
Figure A.4	Analyze/Surveys/ analyzeOtherOutcomes.R	Combined	additionalOutcomes
Figure A.5	Analyze/Media/ analyzeMediaBehavior.R	Browsing	NavFromFB
Figure A.6	Analyze/Media/ analyzeFB.R	Browsing	shareVisitsByRank
Figure A.7	Analyze/Media/ analyzeMediaBehavior.R	Shared	ITTSharedMatch
Figure A.8	Analyze/Media/ analyzeMediaBehaviorSlant.R	Combined	slantScaleExcludingFullData
Figure A.9a	Analyze/Media/ analyzeMediaPersistence.R	Persistence	PersistentEffectMonthSlantScaleExt
Figure A.9b	Analyze/Media/ analyzeMediaPersistence.R	Persistence	PersistentEffectMonthSlantScaleShared
Figure A.10	Analyze/Surveys/ analyzePersuasion.R	PoliticalOpinions	indPoliticalOpinionsOrder
Figure A.11	Analyze/Surveys/ analyzePrimaryOutcomes.R	PoliticalOpinions	opinons4Treat
Figure A.12	Analyze/Surveys/ analyzePolarization.R	Polarization	polIndOrder
Figure A.13	Not generated by code		
Figure A.14	Analyze/Media/ analyzePostContent.R	Combined	sharedPoliticalPosts
Figure A.15	Analyze/Media/ analyzePostContent.R	Exposure	outletAndSection

(continued)

Figure / Table	Source Code	Output Dir	Output File
Figure A.16	Analyze/Media/ analyzePostContent.R	Browsing	outletAndSection
Figure A.17	Analyze/Media/ analyzePostContent.R	Shared	outletAndSection
Figure A.18	Analyze/Media/ analyzeMediaBehaviorHetero.R	Combined	proHetero_Sep
Figure A.19	Analyze/Media/ analyzeMediaBehaviorHetero.R	Combined	counterHetero_Sep
Figure A.20	Analyze/Surveys/ analyzePrimaryOutcomes.R	Combined	MainOutcomesHetero_Sep
Figure A.21	Analyze/Media/ decomposeExposure.R	Exposure	filterBubbleRobust
Table 1	Not generated by code		
Table 2	Analyze/balanceTables.R	Balance	Table2_balanceBaseline
Table 3	Analyze/analyzeCompliance.R	Likes	Table3_Compliance
Table A.1	Analyze/analyzeCompliance.R	Outlets	OutletsOffered
Table A.2	Analyze/balanceTables.R	Balance	descriptiveSubsample
Table A.3	Analyze/balanceTables.R	Balance	balanceBaselineMatch
Table A.4	Analyze/balanceTables.R	Balance	balanceFollow
Table A.5	Analyze/balanceTables.R	Balance	balanceFollowMatch
Table A.6	Analyze/balanceTables.R	Balance	complianceDescriptive
Table A.7a	Analyze/comscoreSegregation.R	Descriptive/Comscore	segregationComscoreUser
Table A.7b	Analyze/Media/ analyzeControlSlant.R	Descriptive/Control	segregationControlUser
Table A.8a	Analyze/comscoreSegregation.R	Descriptive/Comscore	segregationComscoreFBUser
Table A.8b	Analyze/comscoreSegregation.R	Descriptive/Comscore	segregationComscoreYears
Table A.9a	Analyze/comscoreSegregation.R	Descriptive/Comscore	segregationComscore
Table A.9b	Analyze/Media/ analyzeControlSlant.R	Descriptive/Control	segregationControl
Table A.10	Analyze/Media/ analyzeMediaBehavior.R	Combined	allResultsMediaMatchPoisson
Table A.11	Analyze/Media/ analyzeMediaBehaviorSlant.R	Combined	slantBySample
Table A.12a	Analyze/Surveys/ analyzePersuasion.R	PoliticalOpinions	politicalOpinionControls
Table A.12b	Analyze/Surveys/ analyzePolarization.R	Polarization	affectiveControls
Table A.13	Analyze/Surveys/ analyzePolarization.R	Polarization	exMeasureAffective
Table A.14a	Analyze/Surveys/ analyzePrimaryOutcomes.R	PoliticalOpinions	RobustToPrimaryOutletsOpinions
Table A.14b	Analyze/Surveys/ analyzePrimaryOutcomes.R	Polarization	RobustToPrimaryOutletsAffective
Table A.15a	Analyze/Surveys/ analyzePrimaryOutcomes.R	PoliticalOpinions	RobustToPrimarySampleOpinions
Table A.15b	Analyze/Surveys/ analyzePrimaryOutcomes.R	Polarization	RobustToPrimarySampleAffective
Table A.16a	Analyze/analyzePolarization-WithMedia.R	Combined	elasticityAffective
Table A.16b	Analyze/analyzePolarization-WithMedia.R	Combined	elasticityAffectiveControlGroup
Table A.17	Analyze/Surveys/ analyzePolarization.R	Polarization	affectiveOwnOther
Table A.18a	Analyze/Surveys/ analyzePrimaryOutcomes.R	PoliticalOpinions	persuasionAnderson

(continued)

Figure / Table	Source Code	Output Dir	Output File
Table A.18b	Analyze/Surveys/ analyzePrimaryOutcomes.R	Polarization	polarizationAnderson
Table A.19	Analyze/Surveys/ analyzePolarization.R	Polarization	polarizationFamilies
Table A.20a	Analyze/Surveys/ analyzeSurveyPurpose.R	AdditionalOutcomes	surveyObjective3gram
Table A.20b	Analyze/Surveys/ analyzeSurveyPurpose.R	AdditionalOutcomes	surveyObjective2gram
Table A.22a	Analyze/Surveys/ analyzeSurveyPurpose.R	AdditionalOutcomes	surveyObjectiveProControl
Table A.22b	Analyze/Surveys/ analyzeSurveyPurpose.R	AdditionalOutcomes	surveyObjectiveCounterControl
Table A.22c	Analyze/Surveys/ analyzeSurveyPurpose.R	AdditionalOutcomes	surveyObjectiveProCounter
Table A.22a	Analyze/Media/ analyzePostContent.R	Exposure	commonWords
Table A.22b	Analyze/Media/ analyzePostContent.R	Browsing	commonWords
Table A.22c	Analyze/Media/ analyzePostContent.R	Shared	commonWords
Table A.23	Analyze/Media/ analyzeMediaBehaviorSlant.R	Combined	effectOnSlantReweighted
Table A.24a	Analyze/Surveys/ analyzePrimaryOutcomes.R	PoliticalOpinions	primaryOpinionsWeighted
Table A.24b	Analyze/Surveys/ analyzePrimaryOutcomes.R	Polarization	affectiveWeighted
Table A.25	Analyze/analyzeHeteroForest.R	Combined	causalForestResults
Table A.26	Analyze/Surveys/ analyzeKnowledge.R	Information	allPrimaryInformation
Table A.27	Analyze/Media/ analyzeFBContent.R	Information	informationExtension
Table A.28	Analyze/Media/ decomposeExposure.R	Exposure	explainingExpGapWithUsage

References

Bakshy, Eytan, Solomon Messing, and Lada A Adamic. 2015*a*. "Exposure to Ideologically Diverse News and Opinion on Facebook." *Science*, 348(6239): 1130–1132.

Bakshy, Eytan, Solomon Messing, and Lada Adamic. 2015*b*. *Replication Data for: Exposure to Ideologically Diverse News and Opinion on Facebook.* Harvard Dataverse, https://doi.org/10.7910/DVN/AAI7VA.

Comscore. 2018. *Web Behavior Database Panel 2007, 2008, 2017, 2018.* Wharton Research Data Services, University of Pennsylvania.

Federal Election Commission. 2018. *Campaign Finance Data 2015-2018.* https://www.fec.gov/data/browse-data/?tab=bulk-data (last accessed Nov 5, 2020).

Levy, Ro'ee. 2018. "Posts Shared by Outlets on Facebook."

Levy, Ro'ee. 2020. *Replication Data for: Social Media, News Consumption and Polarization: Evidence from a Field Experiment.* Havard Dataverse https://doi.org/10.7910/DVN/V4ZFIT.

Pew Research Center. 2014. *American Trends Panel Wave 1.* https://www.journalism.org/dataset/american-trends-panel-wave-1/.

Pew Research Center. 2016. *American Trends Panel Wave 23.* https://www.pewresearch.org/politics/dataset/american-trends-panel-wave-23/.

Pew Research Center. 2018. *Jan. 3-10, 2018 - Core Trends Survey.* https://www.pewresearch.org/internet/dataset/jan-3-10-2018-core-trends-survey/.

The American National Election Studies. 2016. *The ANES 2016 Time Series Study.* Stanford University and the University of Michigan https://electionstudies.org/data-center/2016-time-series-study/.