

# Auditing Partisan Audience Bias within Google Search

## Supplementary Materials

RONALD E. ROBERTSON, Northeastern University, USA

SHAN JIANG, Northeastern University, USA

KENNETH JOSEPH, Northeastern University, USA

LISA FRIEDLAND, Northeastern University, USA

DAVID LAZER, Northeastern University, USA

CHRISTO WILSON, Northeastern University, USA

### 1 BIAS SCORES

The panel dataset we utilized contained 309,564 Democrats and 209,747 Republicans who we linked to US voter registration records. In total, we collected 319,349,576 tweets, but discarded all of those not containing a URL, leaving us with approximately 113 million Tweets with dates ranging from Summer 2010 to Fall 2017.

A potential match was discarded if any the following were true: (a) there's another voter with that name and location, (b) there's another Twitter account with that name and location, (c) there's another Twitter account with that name and a blank location.

In Table 1 we include additional details about how our scores correlated with other measures of partisan bias.

### 2 EFFECT OF REMOVING WIKIPEDIA LINKS

In our dataset, links to [wikipedia.org](http://wikipedia.org) accounted for 14% of all domains appearing at the first rank. Thus, its left lean (-0.22) may drag the mean weighted bias to the left. We briefly explored the effect of removing all Wikipedia links on our primary metrics, and found that this omission had a significant but relatively unsubstantial effect on the mean average bias, moving it to the right by 0.0073 ( $t = 43.66^{***}$ ; 95% confidence interval [CI] 0.0069-0.0076), on the mean weighted bias, moving it to the right by 0.013 ( $t = 39.28^{***}$ ; 95% CI 0.0124-0.0136), and the mean ranking bias by 0.0056 ( $t = 43.87^{***}$ ; 95% CI 0.0051-0.0061). The next two most prominent domains at rank 1 were also left-leaning media domains: [nytimes.com](http://nytimes.com), which had a partisan audience bias score of -0.26 and accounted for 4.6% of all domains at the first search ranking, and [cnn.com](http://cnn.com), which had a partisan audience bias score of -0.12 and accounted for 4.5% of all domains at the first search ranking, but we did not explore how removing these might have changed our results.

### 3 BIAS BY PERSONALIZATION

As mentioned in the paper, we explored the relationship between users' characteristics and the bias of the SERPs they received using an ANOVA. Below we present the results of these analyses, which could not fit in the main text:

We fit a two-way ANOVA using participants' political party and Google account login status to predict their input bias and ranking bias. We did not find significant differences in input bias by participants' political party or login status, or the interaction term (all  $P \geq 0.38$ ). The impact of

---

Authors' addresses: Ronald E. Robertson, [rer@ccs.neu.edu](mailto:rer@ccs.neu.edu), Northeastern University, 1010-177 Network Science Institute, 360 Huntington Ave. Boston, 02115, MA, USA; Shan Jiang, [sjiang@ccs.neu.edu](mailto:sjiang@ccs.neu.edu), Northeastern University, USA; Kenneth Joseph, [k.joseph@northeastern.edu](mailto:k.joseph@northeastern.edu), Northeastern University, USA; Lisa Friedland, [lfriedland@northeastern.edu](mailto:lfriedland@northeastern.edu), Northeastern University, USA; David Lazer, [d.lazer@neu.edu](mailto:d.lazer@neu.edu), Northeastern University, USA; Christo Wilson, [cbw@ccs.neu.edu](mailto:cbw@ccs.neu.edu), Northeastern University, USA.

Table 1. Our partisan bias scores compared to similar scores developed in other recent work using Pearson’s  $r$  and Spearman’s  $\rho$ .  $N$  is the number of domains we were able to match to across scoring methods.

Bias Score	$N$	$r$	95%CI	$\rho$	$p - value(\rho)$
Facebook (Bakshy et al., 2015)	447	0.959	(0.958, 0.960)	0.945	$1.09 \cdot 10^{-218}$
Budak et al., 2015	16	0.930	(0.928, 0.932)	0.732	0.0013
AllSides Patented	200	0.814	(0.809, 0.818)	0.734	$4.11 \cdot 10^{-35}$
Pew Trust Survey 2014	35	0.784	(0.778, 0.789)	0.550	0.0006
AllSides Community	200	0.728	(0.721, 0.735)	0.639	$2.43 \cdot 10^{-24}$
MTurk domain raters	400	0.570	(0.560, 0.580)	0.497	$2.17 \cdot 10^{-26}$

participants login status on input and output bias was suggestive but not significant (input bias:  $F(1, 15265) = 2.88, P = 0.09$ ; output bias:  $F(1, 15265) = 2.88, P = 0.10$ ). Differences by political party and the interaction were not significant either (all  $P \geq 0.38$ ). Similarly, we conducted the same tests while aggregating SERPs by root query and did not find any significant differences after applying a Bonferroni correction. These results suggest that there is, at most, a weak association between a user’s login status or political party affiliation and the partisan bias of the SERPs they receive, casting doubt on concerns of filter bubbles in Google Search.

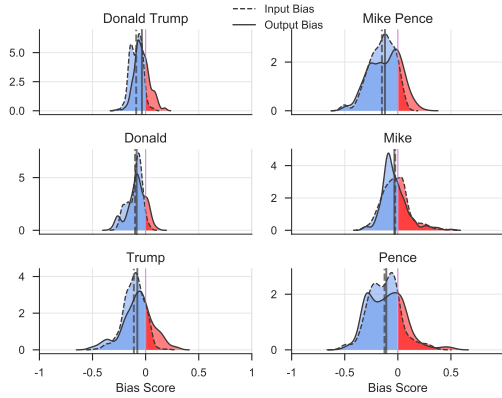
#### 4 ROOT QUERIES

Below are the root queries we used and a list of their unique corresponding suggestions (Table 1), as well as the aggregated crowdsourced ratings we collected for them and their corresponding suggestions (Table 2).

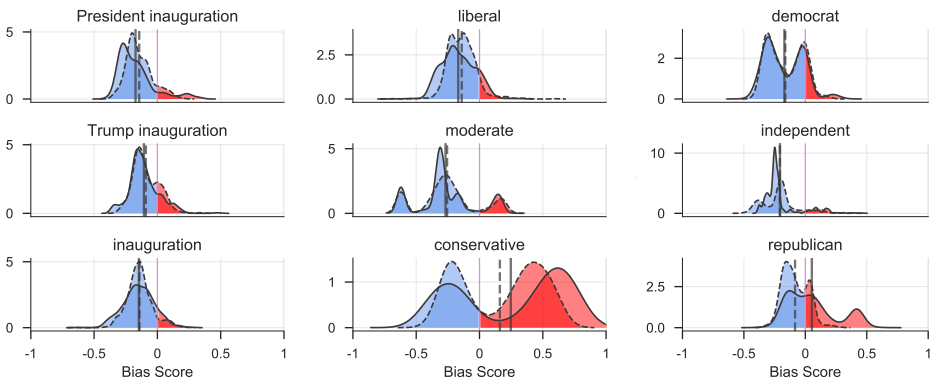
#### 5 INPUT AND OUTPUT BIAS DISTRIBUTIONS BY ROOT QUERY

We compared the input and output bias distributions to obtain a clearer picture of how bias varied by query, but these figures were too large to fit in the main text (Figure 1).

Received April 2018; revised June 2018; accepted September 2018



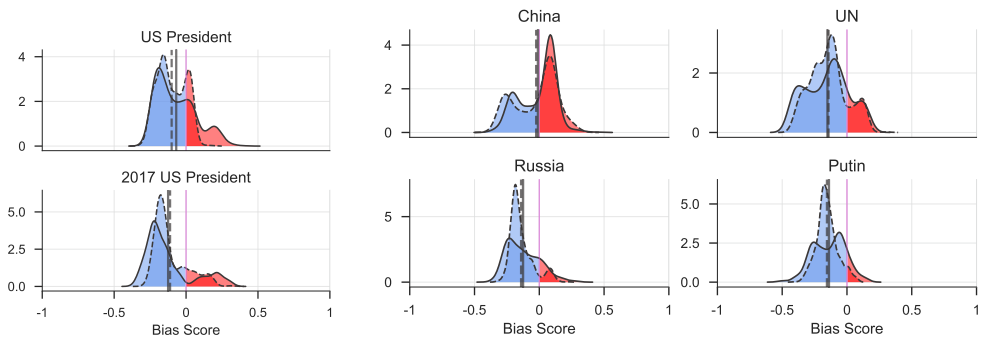
(a) Political actor root queries.



(b) Inauguration root queries.

(c) Political view root queries.

(d) Political party root queries.



(e) President root queries.

(f) Foreign entity root queries.

Fig. 1. Overall distributions of SERP bias metrics in the standard windows, as well as by root query. The distribution beneath the dashed line is the input bias (SERP bias due to Google’s filtering algorithm or underlying corpus), and the distribution beneath the solid line is the output bias (SERP bias after Google’s ranking algorithm was applied). For each root, the vertical dashed line is the mean input bias and the vertical solid line is the mean output bias.

Table 2. The root queries we used and their corresponding unique suggestions.

Root Query	Unique Google Search Suggestions
Donald Trump	donald trump age, donald trump cabinet, donald trump jr, donald trump memes, donald trump news, donald trump twitter, donald trump wall, donald trump wife
Donald	donald glover, donald penn, donald sterling, donald trump, donald trump news
Trump	trump approval rating, trump australia, trump cabinet, trump children, trump executive orders, trump family, trump h1b, trump hotel dc, trump impeachment, trump inauguration time, trump las vegas, trump news, trump polls, trump tower, trump twitter, trump wall, trump wife
Mike Pence	mike pence age, mike pence bio, mike pence black history, mike pence black history month, mike pence black history month tweet, mike pence children, mike pence dance party, mike pence daughters, mike pence family, mike pence height, mike pence house, mike pence indiana, mike pence net worth, mike pence treason, mike pence twitter, mike pence west point, mike pence wife, mike pence wiki, mike pence young
Mike	mike adame, mike and eds, mike and molly, mike bass ford, mike bishop, mike causey, mike coffman, mike colter, mike comrie, mike conley, mike connors, mike duman, mike epps, mike evans, mike fisher, mike francesca, mike gesicki, mike greenwell, mike gundy, mike iltch, mike johnson, mike leach, mike locksley, mike macintyre, mike madigan, mike matheny, mike mccarthy, mike mikurak, mike milken, mike morgan, mike myers, mike napoli, mike nugent, mike pence, mike pompeo, mike posner, mike riley, mike rowe, mike shaw, mike shaw kia, mike sherm, mike smith, mike the tiger, mike torino, mike toler, mike tomlin, mike trout, mike turner, mike tyson, mike vick, mike wallace, mike weber, mike workman, mike zimmer
Pence	pence age, pence and flynn, pence and trump, pence auction, pence black history, pence black history month, pence brothers, pence calls wife mother, pence children, pence coin, pence conversion, pence dance party, pence daughters, pence definition, pence family, pence gallery, pence gay dance party, pence giants, pence hall, pence indiana, pence kia, pence ks, pence law library, pence lincoln tweet, pence march for life, pence meaning, pence money, pence msu, pence news, pence nissan, pence orthodontics, pence park, pence plumbing, pence rally, pence schedule, pence septic, pence springs wv, pence subaru, pence to dollars, pence to usd, pence twitter, pence vs kaine, pence west point, pence wife, pence winery
democrat	democrat and chronicle, democrat and chronicle archives, democrat and chronicle jobs, democrat and chronicle pets, democrat and chronicle sports, democrat articles, democrat bugs, democrat color, democrat definition, democrat donkey, democrat missourian, democrat news, democrat or republican, democrat symbol, democrat voter fraud, democrat who voted for sessions, democratic party
independent	independent adoption center, independent agencies, independent alligator, independent american party, independent assortment, independent assortment of chromosomes is a result of, independent auto sales, independent bank, independent bank ionia mi, independent bank memphis, independent bar, independent brewing, independent burger, independent clause, independent contractor, independent core strategies, independent day school, independent definition, independent events, independent flooring, independent glass, independent health, independent herald, independent imaging, independent insurance agents of iowa, independent insurance brokers, independent journal review, independent living, independent living madison wi, independent living resource center, independent living resources, independent mail, independent online, independent opportunities, independent party, independent presbyterian church, independent records, independent review, independent santa barbara, independent school, independent sf, independent trucks, independent variable, independent variable definition, independent weekly
republican	republican american, republican city ne, republican color, republican eagle, republican elephant, republican herald, republican journal, republican logo, republican motherhood, republican national committee, republican party, republican party of florida, republican party of iowa, republican party of kentucky, republican party of new mexico, republican party of wisconsin, republican party platform, republican platform, republican presidents, republican senators, republican symbol, republican voter guide
liberal	liberal arts, liberal arts advising, liberal arts and science academy, liberal arts and sciences advising uf, liberal arts and sciences uiuc, liberal arts auburn, liberal arts building tamu, liberal arts college, liberal arts colleges, liberal arts majors, liberal arts math, liberal arts ole miss, liberal arts psu, liberal arts tamu, liberal arts utep, liberal cup, liberal cup menu, liberal definition, liberal developmentalism, liberal education, liberal education electives uiuc, liberal feminism, liberal finance, liberal forum, liberal high school, liberal kansas, liberal ks, liberal ks weather, liberal meaning, liberal media, liberal meltdown, liberal memes, liberal mo, liberal party, liberal redneck, liberal redneck manifesto, liberal snowflakes, liberal studies, liberal studies cal poly, liberal studies cornell, liberal studies fresno state, liberal studies wcu, liberal tears, liberal tears, liberal tears gun oil, liberal tears meme, liberal tears mug
moderate	moderate conservative, moderate constitution 1795, moderate definition, moderate depression, moderate enlightenment, moderate exercise, moderate hikes in the smokies, moderate in a sentence, moderate income housing, moderate incongruity, moderate intellectual disability, moderate intensity statin, moderate means, moderate muslim, moderate nat type, moderate news sources, moderate persistent asthma, moderate political views, moderate realism, moderate republican, moderate republican senators, moderate republicans, moderate risk of isolation training, moderate sedation, moderate sedation certification, moderate sleep apnea, moderate sliding scale, moderate stress, moderate stroke impairment, moderate synonym, moderate to severe plaque psoriasis, moderate use of caffeine by athletes
conservative	conservative definition, conservative news, conservative newspapers, conservative party, conservative review, conservative talk radio, conservative treehouse, conservative tribune
Trump inauguration	trump inauguration 2017, trump inauguration artists, trump inauguration bad lip reading, trump inauguration ball, trump inauguration cartoon, trump inauguration crowd photos, trump inauguration dance, trump inauguration date, trump inauguration entertainment, trump inauguration gigapixel, trump inauguration lip sync, trump inauguration performance, trump inauguration performers, trump inauguration picture, trump inauguration protests, trump inauguration speech, trump inauguration speech youtube, trump inauguration tickets, trump inauguration time, trump inauguration voice over, trump inauguration youtube
inauguration	inauguration 2017, inauguration bad lip reading, inauguration ceremony, inauguration countdown, inauguration crowd, inauguration date, inauguration day, inauguration day 2017 time, inauguration day federal holiday, inauguration day lip reading, inauguration day protest seattle, inauguration definition, inauguration lip reading, inauguration live, inauguration live streaming, inauguration meaning, inauguration performers, inauguration protest, inauguration schedule, inauguration schedule 2017, inauguration schedule of events, inauguration tickets, inauguration trump, inauguration voice over
President inauguration	president inauguration 2016, president inauguration 2017, president inauguration 2017 date, president inauguration 2017 live, president inauguration date, president inauguration day, president inauguration day 2017, president inauguration live, president inauguration schedule, president inauguration speech, president inauguration stanford, president inauguration tickets, president inauguration time
US President	us president 2016, us president 2017, us president election, us president fdr, us president impeached, us president list, us president pension, us president polls, us president popular vote, us president salary, us president timeline, us president trivia
2017 US President	2017 us president candidates, 2017 us president election, 2017 us president inauguration, 2017 us president prediction
China	china 1, china 10, china 88, china a go go, china airlines, china anne mcclain, china bear, china bell, china blossom, china bowl, china buffet, china cafe, china chef, china chilcano, china city, china cottage, china doll, china dragon, china dynasty, china eastern airlines, china eatery, china express, china flag, china flavor, china food, china fun, china garden, china garden terre haute, china go, china gourmet, china hing hull, china house, china house kenosha, china house storm lake, china hut, china inn, china inn blacksburg, china king, china king goleta, china king menu, china kitchen, china lights, china moon, china news, china ocean, china one, china one dubois pa, china palace, china pavilion, china peak, china peak weather, china poblano, china population, china post, china post tracking, china southern airlines, china spring, china spring high school, china spring isd, china star, china star menu, china sun, china tango, china taste, china villa, china wok, china wok menu, china wok owasso, china wok terre haute
UN	an ambassador, un bank, un beso, un bien, un careers, un charter, un chien andalou, un contributions to world peace, un cruise, un declaration of human rights, un deux trois, un di, un dia a la vez, un dia a la vez letra, un dia de estos, un dia la vez, un dia sin inmigrantes, un dia sin mexicanos, un foundation, un head kofi, un in english, un in spanish, un internship, un internships, un israel, un jobs, un journal, un lab test, un member states, un members, un nuevo dia, un poco, un poquito, un posto al sole, un prefix, un refugee agency, un reloj de udstedes, un resolution israel, un rinconcito en el cielo, un security council, un small arms treaty, un webmail, un women, un unlocked games, unf
Russia	russia hough, russia map, russia news, russia nuclear war, russia ohio, russia population, russia sanctions, russia spy ship, russia today, russia ukraine, russia vs usa, russia war]
Putin	putin and trump, putin bay, putin height, putin memes, putin net worth, putin news, putin on the ritz, putin poison, putin prostitute remark, putin russia, putin trump, putin ukraine

Table 3. The root queries we used, the number of unique queries we retrieved through Google’s autocomplete, and the results of our crowdsourced query classification and partisan bias rating aggregated by root query.

Topic Area	Root Query	Unique Queries	Classification			Bias Rating	
			Political	Ambiguous	Not Political	Mean	SD
Political Actors	Donald Trump	9	100.0%	0.0%	0.0%	0.08	0.17
	Donald	6	33.3%	33.3%	33.3%	0.00	0.00
	Trump	18	83.3%	16.7%	0.0%	-0.03	0.15
	Mike Pence	20	95.0%	0.0%	5.0%	-0.03	0.11
	Mike	55	7.3%	21.8%	70.9%	0.00	0.00
	Pence	46	58.7%	13.0%	28.3%	-0.01	0.12
Inauguration	Trump inauguration	22	100.0%	0.0%	0.0%	-0.01	0.15
	inauguration	25	92.0%	4.0%	4.0%	-0.02	0.11
	President inauguration	14	100.0%	0.0%	0.0%	0.04	0.08
Political Party	democrat	18	94.4%	0.0%	5.6%	0.01	0.19
	independent	46	4.3%	26.1%	69.6%	0.00	0.00
	republican	23	91.3%	8.7%	0.0%	0.09	0.15
Political Ideology	liberal	47	38.3%	17.0%	44.7%	0.07	0.23
	moderate	33	24.2%	15.2%	60.6%	0.01	0.04
	conservative	9	100.0%	0.0%	0.0%	0.15	0.17
US President	US President	13	100.0%	0.0%	0.0%	-0.03	0.10
	2017 US President	5	100.0%	0.0%	0.0%	0.05	0.11
Foreign Entities	China	70	1.4%	8.6%	90.0%	0.00	0.00
	UN	46	17.4%	19.6%	63.0%	0.00	0.02
	Russia	13	69.2%	23.1%	7.7%	0.00	0.00
	Putin	13	92.3%	0.0%	7.7%	-0.02	0.07