Selectivity in a fragmented media environment: A novel approach

Sakshi Bhalla & Harsh Taneja

Political Communication Preconference

APSA 2024

Fragmentation and polarization

New media offers plenty of choice and audience attention gets fragmented

Concerns about polarization (based on selective exposure) often confirmed in surveys and experiments (Stroud, 2008; Garrett & Stroud, 2014; to Tyler, Grimmer & Iyengar, 2022)

Fragmentation and duplication

Research taking a macro view of audience formation has argued using passively obtained behavioral data

- Most outlets have audience overlaps with most other outlets
  - E.g., Same people visit both Fox News and NYTimes
  - Persistence of the popular
- Downplays concerns about polarization

(Webster & Ksiazek, 2012; Majó-Vázquez et al, 2019; Webster, 2014)

Fragmentation and duplication

Research taking a macro view of audience formation has argued using passively obtained behavioral data

- Most outlets have audience overlaps with most other outlets
  - E.g., Same people visit both Fox News and NYTimes
  - Persistence of the popular
- Downplays concerns about polarization in news consumption

Ultimately whether fragmentation leads to polarization is a function of how media exposure is being assessed

Past work uses Absolute Duplication

- Measure of how many individuals co-visit a pair of outlets as a share of the total internet population.
- Pairwise data scaled to create a symmetric network (matrix) of all outlets under consideration

Past work uses Absolute Duplication

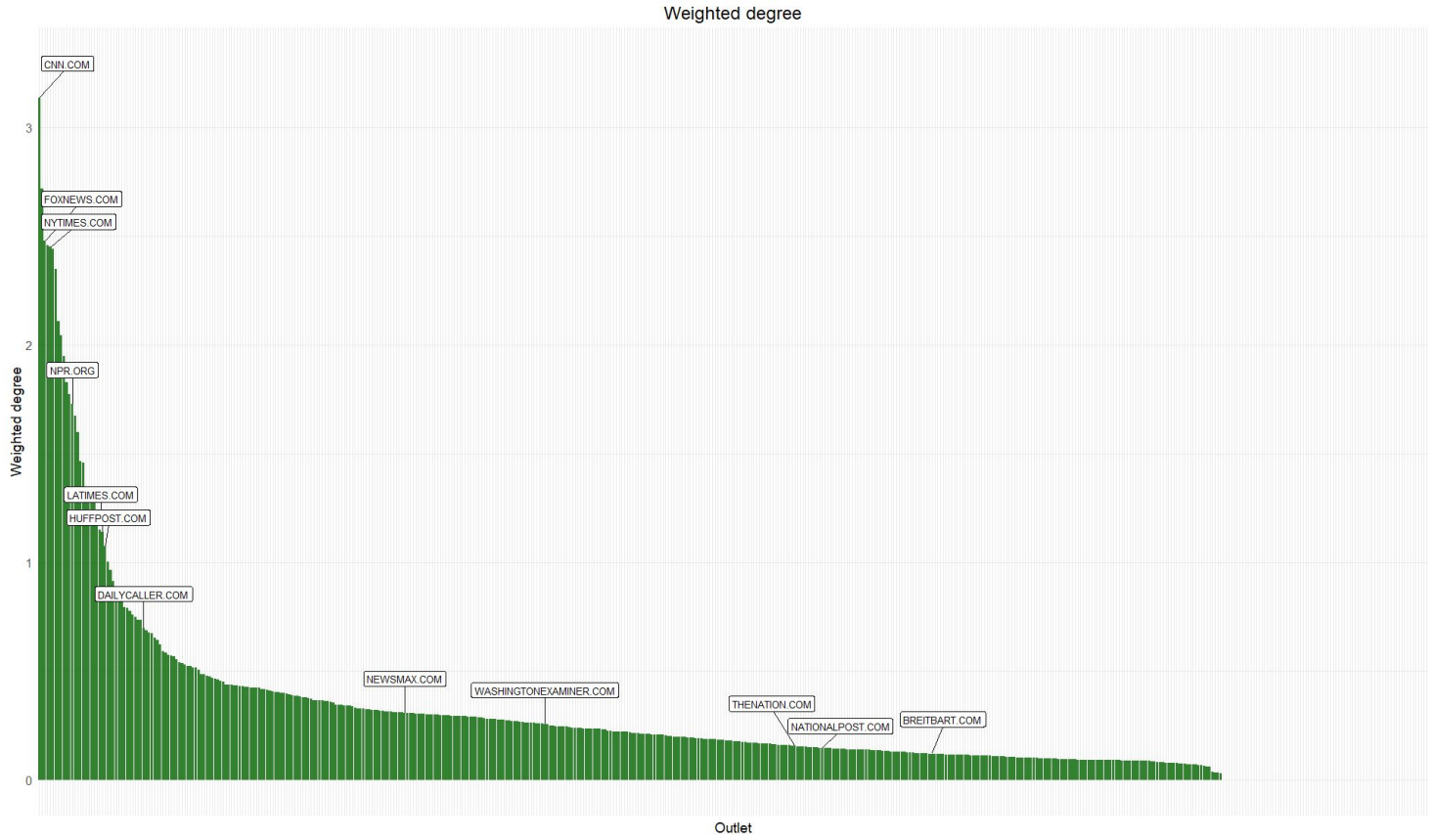
In a universe of 100 individuals
80 visit Outlet A
60 visit Outlet B
20 visit both
Absolute duplication: Visiting A and B/100 → 20/100 = 0.2

Outlets	**NYTIMES**	**FOXNEWS**
NYTIMES	0	0.081
FOXNEWS	0.081	0

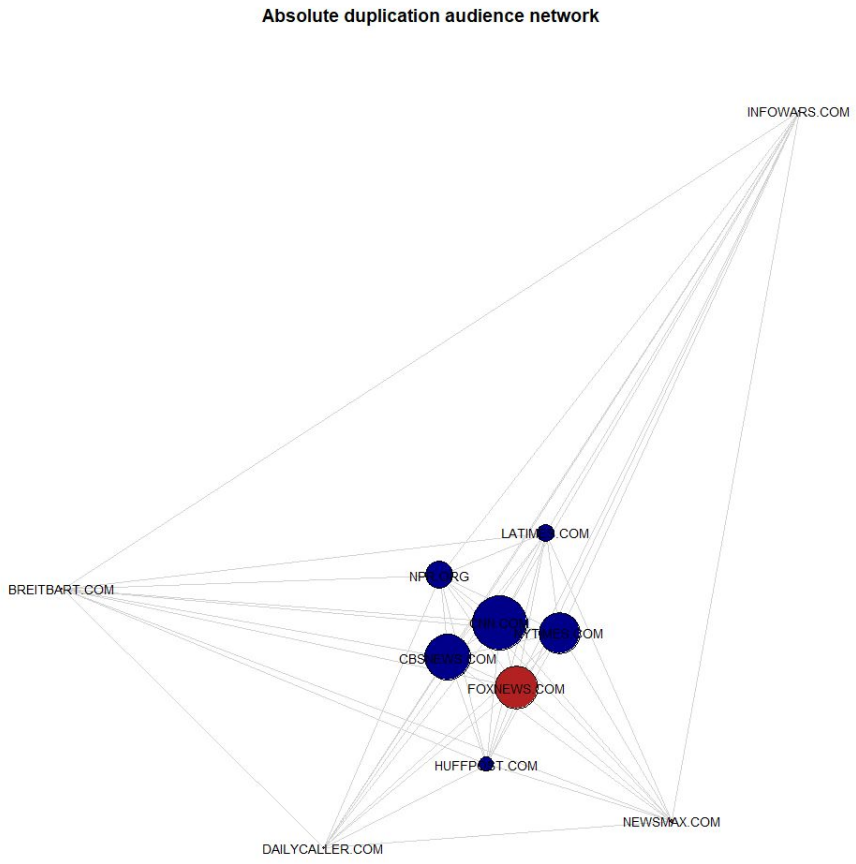
Usually results in dense matrices, as most pairs of outlets have greater than expected duplication resulting in an interpretation of no polarization in news consumption or selectivity.

Degrees: Number of connections an outlet has

Extant research finds very high degrees for popular legacy media

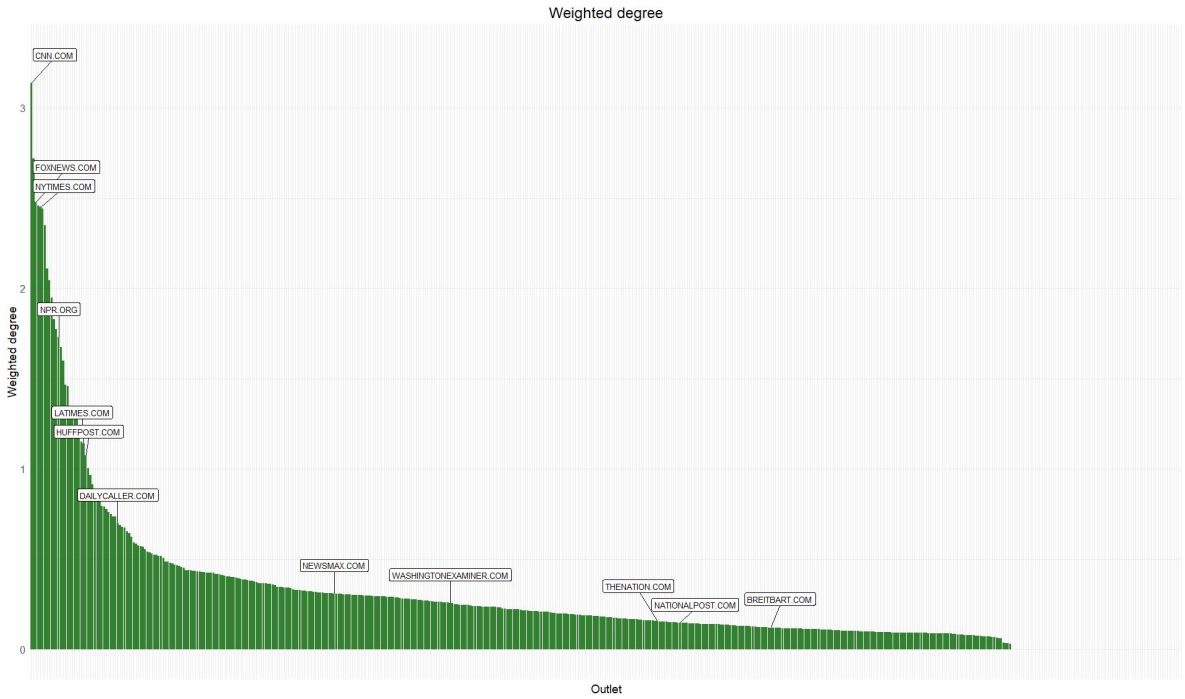


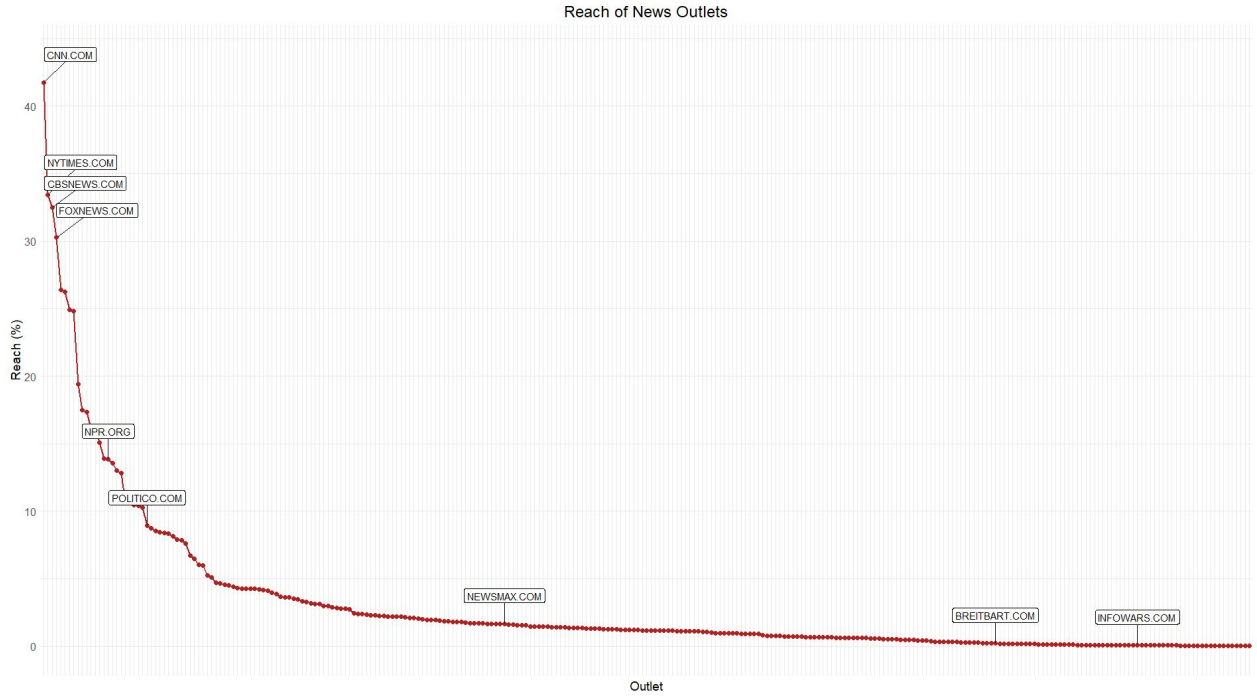
Audience formations (absolute duplication)



Degrees: Number of connections an outlet has

Extant research finds very high degrees for popular legacy media





But degrees correlate almost perfectly (0.91) with reach!

Our argument

Absolute duplication replicates the macroscopic structure of the audience market, where disproportionately large reach of a few large outlets (which most audiences visit) skews our interpretations.

We advocate using conditional duplication.

Our argument

In a universe of 100 individuals
80 visit Outlet A
60 visit Outlet B
20 visit both
Conditional duplication of B given A: Visiting A and B/Reach of A → 20/80 = 0.25 (25 % of A's audience also visited B)
Conditional duplication of A given B: Visiting A and B/Reach of B → 20/60 = 0.3

Outlets	**NYTIMES (,%)**	**FOXNEWS**
NYTIMES	0	29.460
FOXNEWS	26.883	0

Conditional duplication

Visits to outlet A are not independent of visits to outlet B, and thus, can be written as:

$$P\,(A|B) = \frac{P(A \cap B)}{P(B)}$$

This produces a directed network such that $P\,(A|B) \neq P(B|A)$.

Data

Audience data aggregated at the outlet level from comScore panels:

- Media Metrix (multiplatform)

A sample of 506 “news/info” websites specifically categorized as:

- General news
- Politics

Only included websites reaching at least 0.05% of the population.

Absolute vs Conditional duplication

Outlets	**Absolute duplication** with BREITBART (%)	**Conditional duplication** with BREITBART (%)
NYTIMES	0.1	55.442
CNN	0.1	64.436
FOXNEWS	0.1	80.655
NPR	0.06	34.421
HUFFPOST	0.04	27.057
NEWSMAX	0.09	21.010
DAILYCALLER	0.01	63.292

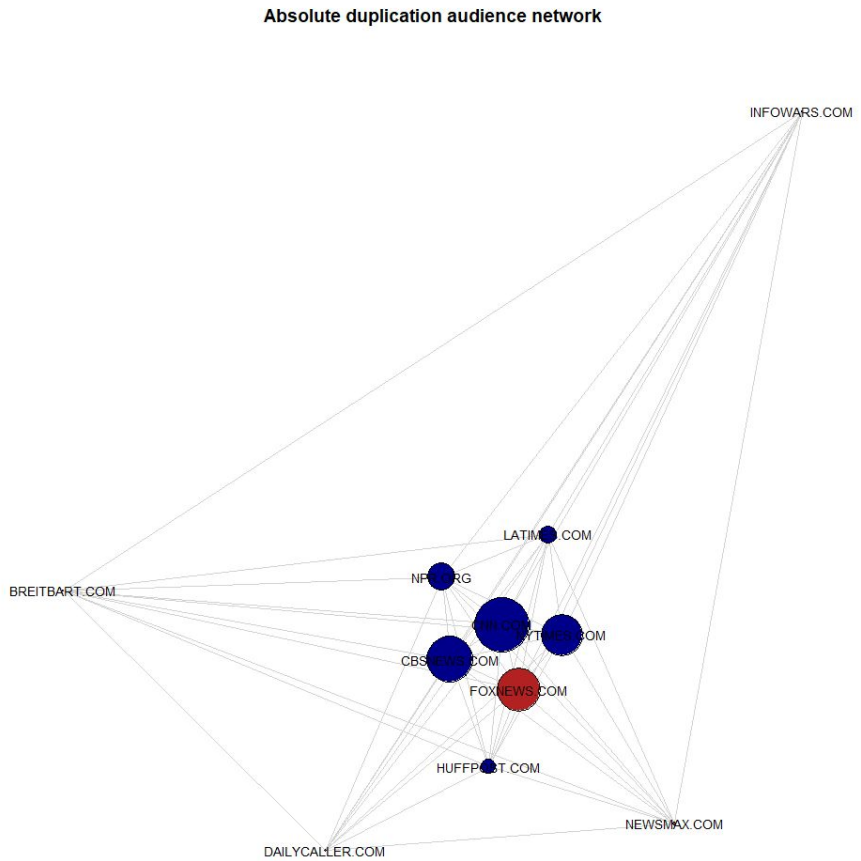
Visiting Breitbart “given” NYTimes: 0.333%
Visiting Breitbart “given” DailyCaller: 18.052%

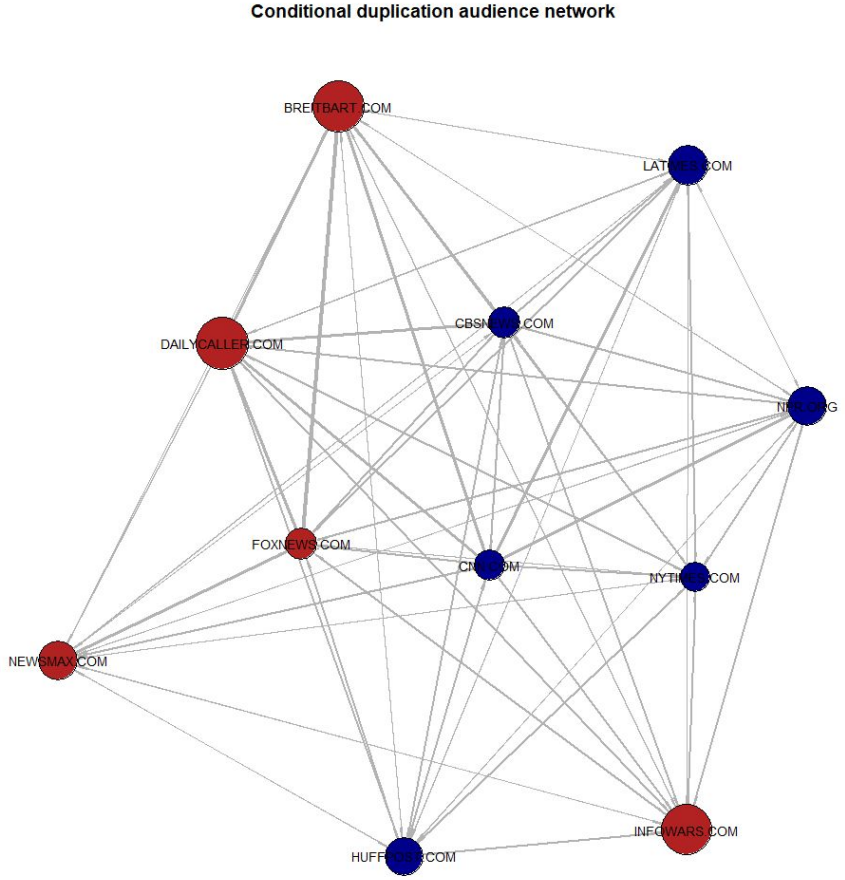
Conditional duplication

Outlets	**Absolute duplication** with BREITBART (%)	**Conditional duplication** with BREITBART (%)	**Index shared visitors** by reach of Outlet_i and BREITBART
NYTIMES	0.1	55.442	1.827
CNN	0.1	64.436	1.697
FOXNEWS	0.1	80.655	2.443
NPR	0.06	34.421	4.497
HUFFPOST	0.04	27.057	3.945
NEWSMAX	0.09	21.010	14.519
DAILYCALLER	0.01	63.292	99.203

- Network more sensitive to instances where less popular outlets are disproportionately visited given visits to another outlet.
- High or low “absolute duplication” is simply a function of the reach of the outlet.

Audience formations





- Network more sensitive to instances where less popular outlets are disproportionately visited given visits to another outlet.
- High or low “absolute duplication” is simply a function of the reach of the outlet.

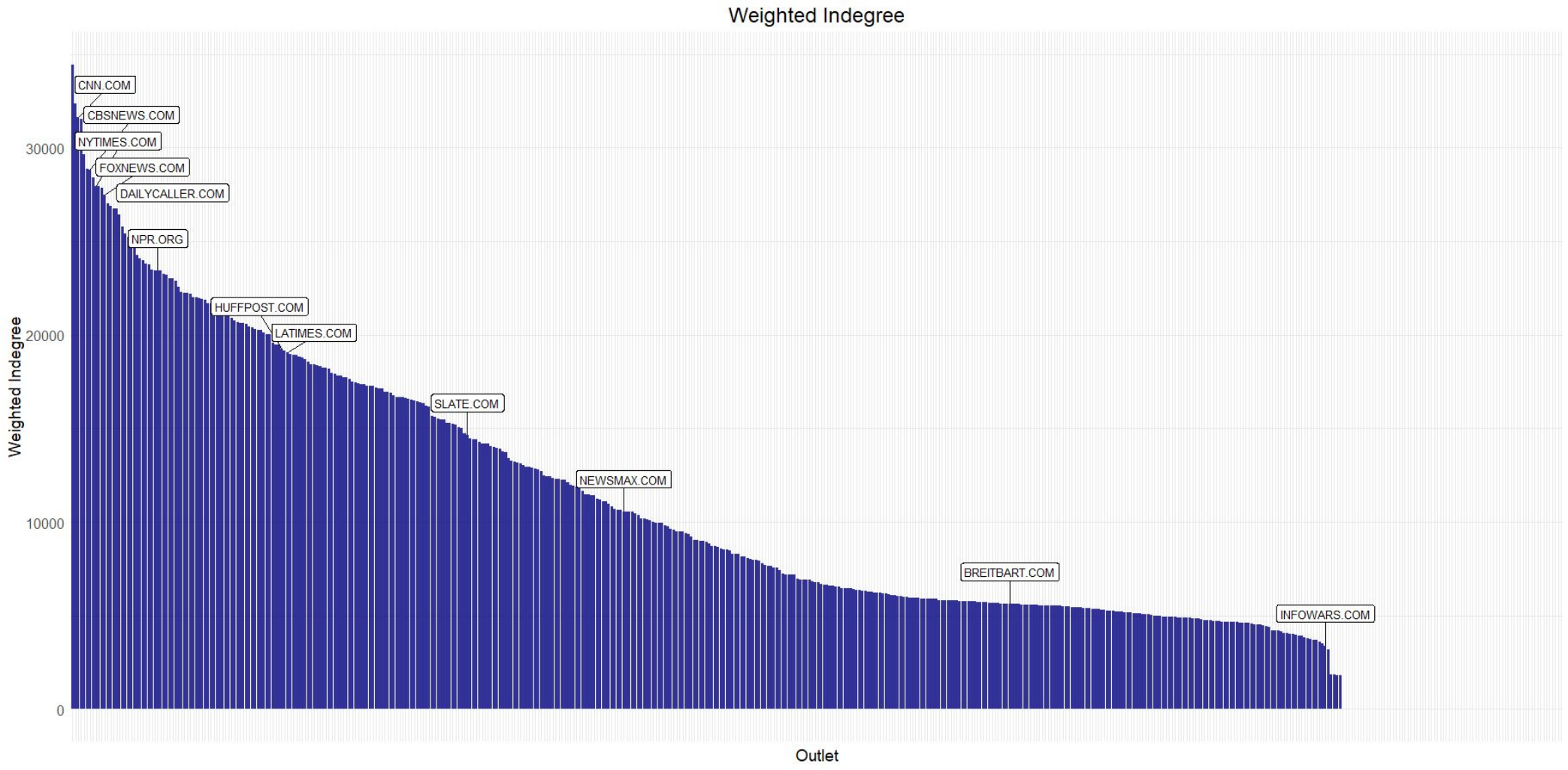
Conditional duplication: Indegrees and Outdegrees

Indegree: The number of edges that come "into" a vertex in a directed graph (Visiting outlet A, having visited any other outlet)

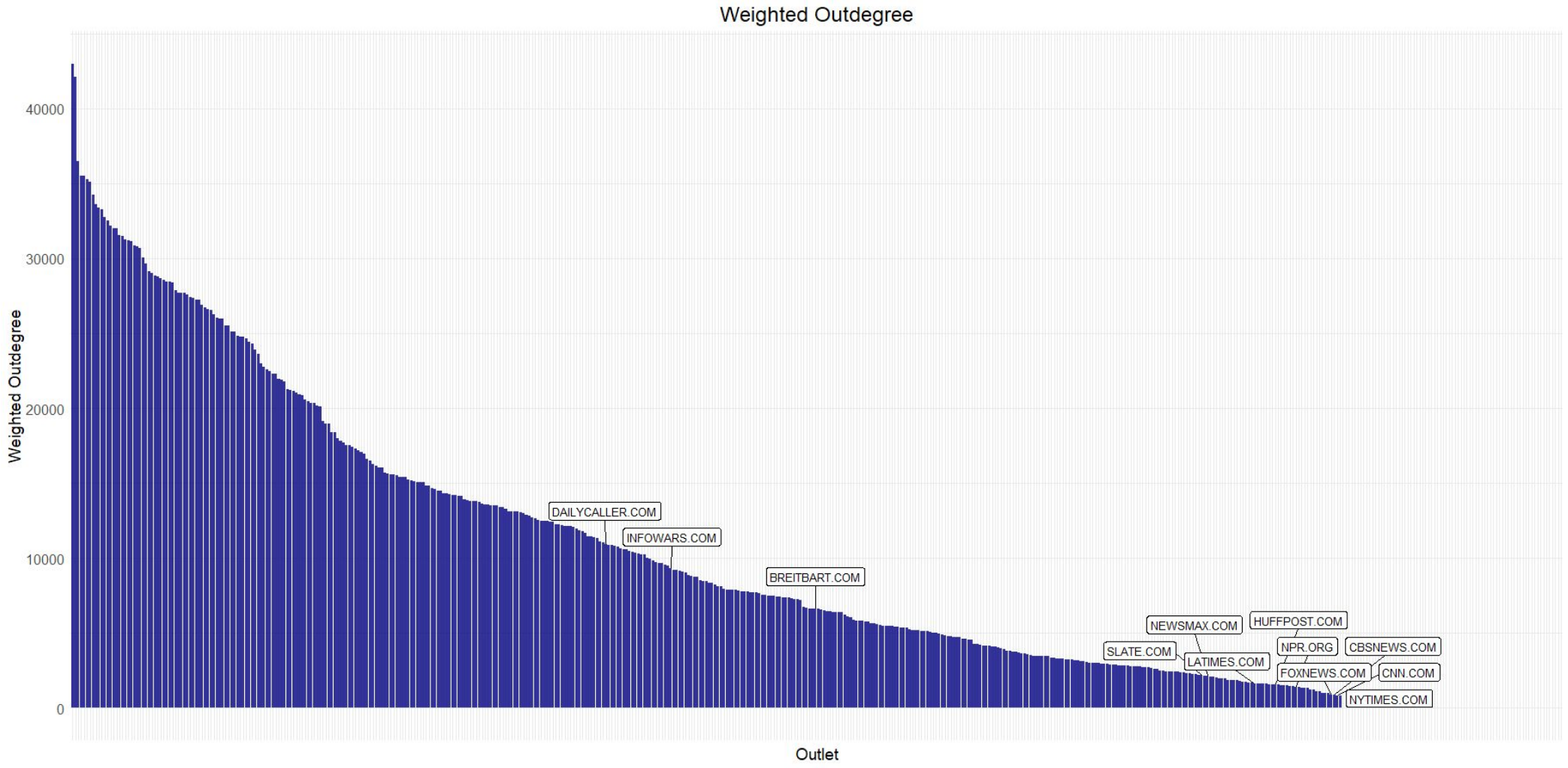
Outdegree: The number of edges that go "out" of a vertex in a directed graph. (Visiting any other outlet, having visited outlet A)

Measure relative *flow* of attention and relative influence in the network.

(A different) View from 30,000 feet



(A different) View from 30,000 feet



(A different) View from 30,000 feet

- Relative audience flows: how much visiting one outlet influences the likelihood of visiting the other
- Selective exposure & niche behaviours: Small outlets may have low reach but extremely high conditional duplication with other big/small outlets revealing *tendencies* towards certain types of content.
- Independence from size: Reveals meaningful patterns of behaviour which may have been obscured earlier.