Discussion of "Local News Online: Aggregators, Geo-Targeting and the Market for Local News" by George, L. and C. Hogendorn

Levy, Daniel

Bar-Ilan University, Emory University, and RCEA

10 October 2013

Discussion of

Local News Online: Aggregators, Geo-Targeting and the Market for Local News

Lisa George and Christiaan Hogendorn

Daniel Levy
Bar-Ilan University, Emory University, and RCEA

October 10, 2013

News Aggregators – Thieves?

- Robert Murdoch (News Corp):

  "… parasites, content kleptomaniacs, vampires, ..., thieves who steal all our copyright." A. Huffington (FTC, 2009)

- Robert Thomson (Managing Ed., *WSJ*):

  "… tech tapeworms in the intestines of the internet."

- Bill Keller (Exec. Ed., *NYT*):

  "… Somali piracy."

Two Observations

- A survey (2009) found that
  - 44% of Google News readers only scan the headlines
  - They don't follow the links
  - They never visit the original news source site

- Simon Dumenco (Columnist, *Advertising Age*, 2012)
  - His blogpost at AdAge.com went viral
  - Reported by Techmeme, Huffington Post, and others.
  - Google Analytics: #visits to the original post
  - Generated by Techmeme – 746, Huffington Post – 57

News Aggregators – Value Producers?

- Arianna Huffington (FTC, 2009):

  "We link to the *WSJ* daily, … [yet] we have never heard from them. You know why? Because we drive a lot of traffic to them, and they like it."

  "… because so many other sites understand it, we get hundreds of requests from news outlets every day to link back to them. It's not a zero sum game, it's … the link economy."

Two Views – Two Possibilities?

- News aggregators – substitute the content creators
  - Content creators incur the costs
  - Aggregators free-ride
  - Grab the potential readers
  - Adv. revenue loss and less incentive to gather news
- News aggregators – complement the content creators
  - Curate
  - Direct the readers
  - Match between readers and stories
  - Create more traffic at originating news sites

News Aggregators – News Corp Too?

- Arianna Huffington (FTC, 2009):

  "Let's be honest, many of those complaining the loudest are working both sides of the street… The ***WSJ*** has a tech section that's nothing more than a parasite -- I mean aggregator -- of outside content. **Foxnews.com** is a politics bloodsucker that blood sucks -- sorry, I mean aggregates -- and links to storage from… ***NYT***, ***WP***, **MSNBC,** and others…News Corp owns **IGN**, which has a variety of web properties, including the **Rotten Tomatoes Movie Review** aggregator site, which is entirely made up of movie reviews pulled together from other places."

News Aggregators – More Complicated

- Arianna Huffington (FTC, 2009):

  “Of course, let me just remind Rupert Murdoch … that you can shut down the indexing of your content by Google right now, this very minute …simply by actually clicking “disallow” in your robots.txt file. …It's actually much faster than whining. But … as soon as you do that, and this is why you haven't done it, you will start denying your content to other sites that aggregate and link back to your original source, and you stand to lose a large part of your traffic overnight.”

Role of News Aggregators

- Substitute or complement?
- In theory they may be either one
- Theory cannot help in settling the question
- Empirical question
- That's the main motivation of the empirical papers that study this issue

Substitute or Complement?

- Effect of information technology on information markets
- Important question
- Regulatory implications
- Implications for market structure
- Optimal internet-era business model/strategy

George and Hogendorn (2013)

- Emphasize the possibility that (instead of directly substituting or complementing the original news creators), the aggregators may shift consumption patterns across media outlets
- Aggregators may reduce search costs uniformly
- Aggregators may reduce the cost of consuming some types of news relative to others, alter relative demands

Existing Literature

- Role of aggregators – relatively new question
- Not too many papers
- Directly relevant theoretical papers – not too many
- Directly relevant empirical papers – perhaps even fewer

Chiou and Tucker (2011)

- How ICT technology affect consumer info gathering
- Nat. experiment: dispute GN-AP, in Jan10 GN removed AP
- Compare GN readers' visits before and after
- Control - Yahoo! News readers, who had cont. access to AP
- Aggregator users go on to visit content websites
- After AP links removal, fewer GN readers visited other news sites

Athey and Mobius (2012)

- Aggregators' impact on the Q and composition of news read
- Case study: Nov. 2009, France GN introduced "localization"
- People can enter zip code to enable the feature
- Control group – people who did not enable the feature
- Use of GN leads to a greater local news consumption
- The increase in the local news consumption is temporary

George and Hogendorn (2013)

- Effect of adding geo-targeted local news links to the GN on the readers visit to local news sites
- Case study: July 2010, GN started automatic “localization”
- No need to enter a zip code, the system identifies IP
- The study analyses the patterns of local news consumption
- Test 1 – GN users vs to Yahoo! News users
- Test 2 – GN users before vs GN users after redesign

Data

- Rich and interesting
- Observation at a h-hold/machine/browsing level
- Many details to take care of
- Many decisions to make and defend, not easy
- The authors succeed in giving attention to lots of details, and in overcoming many challenges
- Very thorough analysis

Data - Households

- April 1, 2010 – Sept 30, 2010
- Panel of 24,859 household news visits
- No. of news visits and number of local news visits
- No. of news and local news visits referred by GN and YN
- MSA, race, income

Data – Site Visits

- Browsing history of ~ 50,000 h-holds who allowed tracking
- Machine-level data
- Possible difficulties:
  - More than one user
  - Mobile devices, office computers are excluded
  - Only the h-holds that agreed (sample not representative)
  - Problems of interpretation/generalization
  - These difficulties are not unique to this study. Studies that use these types of data, face similar difficulties

Data – News Outlets and Local Visits

- 3,184 (top-level) domains
- For each domain – the MSA with the highest number of visits is defined as the home MSA
- This way, each domain is assigned a home MSA/market
- "Local visit" if visitor's and the domain's home MSAs match
- Non-local domains: with < 10% of visits from home MSA (to preserve the NYT-11% and the NYP-10% as local for NYers)
- These criteria identify national media/domains quite well

Data – Intermediation

- Raw session data contain info on referrals from GN and YN
- Difficulty separating referrals from GN and Google
- Develop a measure based on whether or not the domain's name was listed on the referral site on the day of referral
- Identify "intense" GN users – the share of all news visits referred by GN in the first six months of 2010
  - The measure captures only referrals with clicks
  - Cannot separate the users who never visit from those who visit but never click on the link

Data – Measuring Referrals

- 1-3 snapshots are captured daily. If the GN pages are updated more often, then GN referrals will be miscounted in the visit data
- User customization before redesign is not captured, which could lead to under-counting referrals.
- Additional issues introduced by the way scraping is done

Estimation

- Fixed-effect Model
- Four different measures of *Y* (local news consumption)
- Two measures of *X* (treatment specification)
  - One uses YN readers as a control (control ~ treatment)
  - Second employs a measure of intensity of GN use
- Three specifications
  - Semi-log
  - Linear probability
  - Binomial fixed effect

Findings

- GN redesign increased local news consumption and also shifted attention from non-local to local news sources
- The effect was the largest for the most intense users
- Increases in local news consumption seem to arise from more frequent visits to familiar sites rather than to new sites
- The effect, however, is small in magnitude, and not long-lasting

Thoughts/Suggestions/Improvements 1

- Is there a way to conduct the analysis at the zip-code level rather than MSA-level? Wouldn't that be better because many demographic and economic data are available at the zip-code level.

- That way, the households living outside MSAs can be kept.

- Using state-dummies could capture variation that exists across states in the patterns of news consumption. For example, some states consume more national news (e.g., NY, MA, CA, etc.) while in other states there is greater interest in local news.

Thoughts/Suggestions/Improvements 2

- Many choices/decisions had to be made.
- I very much appreciate the detailed discussion of these decisions, and the rational offered.
- Most of them seemed straightforward and made sense.
- Few were somewhat arbitrary (the authors had no other option) and the authors noted those as well.

Thoughts/Suggestions/Improvements 3

- In light of these, the authors conduct various robustness checks, but more robustness/sensitivity analysis can help make an even stronger case, to convince a skeptic.
- For example: run a model in which the control group consists of only those YN users who don't use GN. The sample will be smaller, and the estimate less precise, but if its sign and magnitude are similar, that will be good.
- Another example: run a model where the referral threshold is 15% and 25% instead of 20% to see robustness.
- There are numerous examples like these.

Thoughts/Suggestions/Improvements 4

- Most of these studies, including the current one, cover relatively short period of time
- The reason is that over longer periods, many other things change with confounding effects
- Nevertheless, extending the sample periods covered can be useful and informative

Thoughts/Suggestions/Improvements 5

- I like the paper. I learned a lot.
- They convinced me. The results reported here are in line with the results reported by other studies.
- The suggestions I have made might lead to marginal improvements but are unlikely to reverse the main results.
- I got so convinced in the findings reported that I have even decided to conduct an experimental analysis of my own.
- Let me take you to my web-page at BIU before-and-after:

Daniel Levy's Webpage at BIU