The Impact of Consumer Multi-homing on Ad Prices: Evidence from an Online Marketplace

Yu-Hsin Liu†

† PhD Candidate, Department of Business Economics and Public Policy
Kelley School of Business, Indiana University
Contact: yuhsliu@indiana.edu

January 25, 2018
Big Picture

- Big data governance in the digitized economy
  - Economic impact
  - Non-market impact

- Online behavioral tracking (cookies, tags, pixels, etc.)
  - improves the advertising effectiveness (targeting and frequency)
  - raises the concerns on consumer privacy
Frequency Management

- Kia purchases 1k impressions and they all go to Bob...
  - Bob may be irritated
  - Kia may want to reach more other audiences

- Frequency management allows advertisers to target the right number of ad impressions exposed to a unique user – good for Bob, good for Kia
  - Frequency capping
  - Daily, weekly, monthly
Privacy Regulation

- Frequency management is implementable by storing parts of the consumer web browsing information in the ad server.
- For the counting to apply across different websites or ad networks, it requires that different ad servers to share the consumer browsing data.
- The sharing violates FTC’s self-regulatory principle on privacy protection:
  - First-party (intra-site) tracking is allowed
  - Third-party (inter-site) tracking is restricted
  - An ad network is considered as a first party
Which Website Has Higher Per-impression Price, M1, M2, or Equal?

Two-sided market
- Advertising market
- Consumer market

Three players
- A: advertisers (e.g. Kia)
- M: media/publishers (e.g. YouTube, TechCrunch)
- C: consumers (C1: single-homer; C2: multi-homer)
Hypothesis from the Economic Theory

$M_1$ has a higher per-impression price

- The advertiser can show an accurate number of ad impressions to $C_1$ via $M_1$
- The advertiser cannot do so to $C_2$, as the data in $M_1$ does not reflect the exposure frequency in $M_2$
- All else equal, $C_1$’s attention is more valuable to the advertiser
Research Question

▶ Want to know whether consumer multi-homing leads to lower online ad prices due to imprecise frequency management
  ▶ The answer reflects the level of cross-website tracking and the potential economic impact under the current US privacy regulation
  ▶ The result indicates if the multi-homing literature applies to online media

▶ The first paper to
  ① develop a continuous multi-homing level in various market definitions
  ② provide the empirical evidence of the “multi-homing effect” hypothesis in the online advertising market
  ③ employ a reduced-form technique to estimate the multi-homing effect
The trade-off between advertising effectiveness and privacy regulation (Budak et al. 2016, Goldfarb and Tucker 2011, Johnson 2013)

How consumer multi-homing may affect ad prices
  - Incremental pricing principle (Anderson et al. 2015; Anderson and Jullien 2015; Ambrus et al. 2014)
  - Greater-reach premium (Athey et al. 2016)

Empirical approach for estimating the multi-homing effect
  - Newspaper (Gentzkow et al. 2014)
  - Print magazine (Shi 2015)
Method Preview: Quasi-experiment

▶ Key identification strategy: the ad locations that are more viewable in a webpage are also more responsive to consumer multi-homing.

▶ The multi-homing effect is more negative for the “top” ads than the “bottom” ads due to the higher viewability.
Results Preview

1. Consumer multi-homing indeed hurts ad prices
   - Cross-website tracking is imperfect
   - Further merger of ad networks can improve tracking

2. Perfect tracking increases ad revenue at least by 7%
   - It means about $1 billion annual ad revenue
   - The benefit goes to small and “middle-class” websites

3. The (offline) multi-homing analyses apply to online media
   - If privacy regulation becomes stricter, the analyses become more relevant
   - Privacy regulation can affect the content provision in the consumer market
US Digital Ad Market in 2016 ($72 billion)
Publisher Data and BuySellAds Marketplace

The Atlantic
The Atlantic Monthly Online | theatlantic.com

71,300,000
Monthly Impressions

1,596,089
Followers

The Atlantic Monthly's home on the Internet, featuring current issues online alongside web-specific content on travel, literature, politics, and digital culture.

Website (CPM)

<table>
<thead>
<tr>
<th>Location</th>
<th>Impressions</th>
<th>Ad Available</th>
<th>CPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top Right 300 x 600 Top Right</td>
<td>12,808,000</td>
<td>✔️</td>
<td>$25.00</td>
</tr>
<tr>
<td>Top Right 300 x 250 Top Right</td>
<td>12,451,000</td>
<td>✔️</td>
<td>$17.85</td>
</tr>
<tr>
<td>Leaderboard 728 x 90 Top Center</td>
<td>68,000</td>
<td>✔️</td>
<td>$17.00</td>
</tr>
<tr>
<td>In Content 300 x 250 Center</td>
<td>71,297,000</td>
<td>✔️</td>
<td>$16.00</td>
</tr>
</tbody>
</table>

Channel: Government & Politics & Business & Finance
Member Since October 2014
Supplementary Source: Web Archive in 2016
Multi-homing Measures and comScore 2016

- Web browsing tracking for 80K+ US households with demographics
- I count the number of unique websites (multi-homing level) a user visits within a “narrow” genre, a “broad” genre, and across the entire web space
  - Genre definition depends on web categorization (rely on SimilarWeb)
  - The multi-homing counting is in a daily basis
- I take a weighted average of user-level measures to publisher-level measures and match with ad prices
Supplementary Source: SimilarWeb

- SimilarWeb reports a categorization for the top 50 sites within each refined category – match it with the categorization in BuySellAds.
Publishers in My Sample: the “Middle Class”

![Graph 1: Total Pages Viewed in 2016 (comScore)]

- Min: 1
- Med: 3
- Avg: 162
- Max: 6272860
- Obs: 550746

![Graph 2: Total Pages Viewed in 2016 (BuySellAds)]

- Min: 1
- Med: 10
- Avg: 259
- Max: 13001
- Obs: 534
# Multi-homing Level of Sample Publishers

<table>
<thead>
<tr>
<th>Category</th>
<th>Publishers Categorized</th>
<th>Sample</th>
<th>$H^{\text{narrow}}$ Mean</th>
<th>S.D.</th>
<th>$H^{\text{broad}}$ Mean</th>
<th>S.D.</th>
<th>$H^{\text{universal}}$ Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Business &amp; Politics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Business &amp; Finance</td>
<td>435</td>
<td>1.88</td>
<td>1.02</td>
<td>1.95</td>
<td>1.11</td>
<td>29.14</td>
<td>23.78</td>
</tr>
<tr>
<td></td>
<td>Government &amp; Politics</td>
<td>219</td>
<td>1.14</td>
<td>0.18</td>
<td>1.61</td>
<td>0.39</td>
<td>24.40</td>
<td>8.92</td>
</tr>
<tr>
<td><strong>Entertainment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Entertainment</td>
<td>338</td>
<td>1.33</td>
<td>0.62</td>
<td>1.61</td>
<td>1.17</td>
<td>30.27</td>
<td>35.30</td>
</tr>
<tr>
<td></td>
<td>Gaming</td>
<td>823</td>
<td>1.92</td>
<td>0.75</td>
<td>2.11</td>
<td>0.72</td>
<td>15.61</td>
<td>8.25</td>
</tr>
<tr>
<td><strong>Fashion</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Beauty &amp; Fashion</td>
<td>260</td>
<td>1.20</td>
<td>0.39</td>
<td>1.26</td>
<td>0.57</td>
<td>32.46</td>
<td>48.30</td>
</tr>
<tr>
<td></td>
<td>Weddings</td>
<td>131</td>
<td>1.34</td>
<td>0.87</td>
<td>1.67</td>
<td>0.93</td>
<td>34.11</td>
<td>11.97</td>
</tr>
<tr>
<td><strong>Health</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Food &amp; Drink</td>
<td>798</td>
<td>1.84</td>
<td>0.81</td>
<td>1.88</td>
<td>0.80</td>
<td>25.61</td>
<td>13.33</td>
</tr>
<tr>
<td></td>
<td>Health &amp; Fitness</td>
<td>260</td>
<td>1.15</td>
<td>0.20</td>
<td>1.34</td>
<td>0.50</td>
<td>29.89</td>
<td>22.22</td>
</tr>
<tr>
<td><strong>Life</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Parenting &amp; Education</td>
<td>357</td>
<td>1.32</td>
<td>0.37</td>
<td>1.36</td>
<td>0.40</td>
<td>21.65</td>
<td>9.97</td>
</tr>
<tr>
<td></td>
<td>Home &amp; Architecture</td>
<td>340</td>
<td>1.48</td>
<td>0.64</td>
<td>1.66</td>
<td>0.66</td>
<td>36.42</td>
<td>29.51</td>
</tr>
<tr>
<td></td>
<td>Pets</td>
<td>874</td>
<td>1.55</td>
<td>0.67</td>
<td>1.78</td>
<td>0.63</td>
<td>20.93</td>
<td>5.22</td>
</tr>
<tr>
<td><strong>Recreation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Automotive</td>
<td>1057</td>
<td>2.01</td>
<td>1.11</td>
<td>2.15</td>
<td>1.29</td>
<td>16.14</td>
<td>7.14</td>
</tr>
<tr>
<td></td>
<td>Sports</td>
<td>2249</td>
<td>3.30</td>
<td>3.68</td>
<td>3.64</td>
<td>3.40</td>
<td>36.87</td>
<td>57.91</td>
</tr>
<tr>
<td></td>
<td>Travel</td>
<td>880</td>
<td>1.71</td>
<td>0.95</td>
<td>1.75</td>
<td>0.98</td>
<td>20.52</td>
<td>8.12</td>
</tr>
<tr>
<td><strong>Tech</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>All Things Apple</td>
<td>86</td>
<td>1.07</td>
<td>0.16</td>
<td>1.81</td>
<td>0.60</td>
<td>18.82</td>
<td>8.05</td>
</tr>
<tr>
<td></td>
<td>Cryptocurrency</td>
<td>13</td>
<td>1.0</td>
<td>0.0</td>
<td>2.0</td>
<td>0.0</td>
<td>23.5</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>Technology</td>
<td>807</td>
<td>2.29</td>
<td>0.88</td>
<td>2.35</td>
<td>0.91</td>
<td>28.13</td>
<td>24.17</td>
</tr>
<tr>
<td></td>
<td>Virtualization</td>
<td>26</td>
<td>1.03</td>
<td>0.13</td>
<td>2.46</td>
<td>1.66</td>
<td>69.05</td>
<td>174.76</td>
</tr>
<tr>
<td><strong>Web Design</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Visual Arts &amp; Design</td>
<td>146</td>
<td>1.06</td>
<td>0.17</td>
<td>1.87</td>
<td>1.04</td>
<td>27.78</td>
<td>13.10</td>
</tr>
<tr>
<td></td>
<td>Web Design &amp; Development</td>
<td>512</td>
<td>2.34</td>
<td>1.52</td>
<td>2.41</td>
<td>1.55</td>
<td>29.67</td>
<td>24.63</td>
</tr>
</tbody>
</table>
## Summary Statistics on Ad Prices (cpm)

<table>
<thead>
<tr>
<th>Location</th>
<th>Mean</th>
<th>S.D.</th>
<th>Min.</th>
<th>Max.</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>top left</td>
<td>5.51</td>
<td>10.67</td>
<td>0.25</td>
<td>90</td>
<td>88</td>
</tr>
<tr>
<td>top center</td>
<td>6.68</td>
<td>6.23</td>
<td>0.10</td>
<td>35</td>
<td>321</td>
</tr>
<tr>
<td>top right</td>
<td>6.58</td>
<td>6.77</td>
<td>0.27</td>
<td>40</td>
<td>389</td>
</tr>
<tr>
<td>middle left</td>
<td>3.11</td>
<td>2.73</td>
<td>0.50</td>
<td>12</td>
<td>38</td>
</tr>
<tr>
<td>middle center</td>
<td>4.75</td>
<td>4.62</td>
<td>0.25</td>
<td>20</td>
<td>96</td>
</tr>
<tr>
<td>middle right</td>
<td>4.33</td>
<td>5.41</td>
<td>0.25</td>
<td>35</td>
<td>156</td>
</tr>
<tr>
<td>bottom left</td>
<td>2.97</td>
<td>4.00</td>
<td>0.10</td>
<td>15</td>
<td>24</td>
</tr>
<tr>
<td>bottom center</td>
<td>3.16</td>
<td>3.32</td>
<td>0.25</td>
<td>16</td>
<td>85</td>
</tr>
<tr>
<td>bottom right</td>
<td>3.21</td>
<td>3.87</td>
<td>0.30</td>
<td>18</td>
<td>54</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Size</th>
<th>Mean</th>
<th>S.D.</th>
<th>Min.</th>
<th>Max.</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>large (area &gt; 75000 dips)</td>
<td>9.16</td>
<td>7.88</td>
<td>0.25</td>
<td>40</td>
<td>207</td>
</tr>
<tr>
<td>medium</td>
<td>4.98</td>
<td>6.53</td>
<td>0.10</td>
<td>90</td>
<td>585</td>
</tr>
<tr>
<td>small (area ≤ 65520 dips)</td>
<td>4.63</td>
<td>4.74</td>
<td>0.10</td>
<td>30</td>
<td>464</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Shape</th>
<th>Mean</th>
<th>S.D.</th>
<th>Min.</th>
<th>Max.</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>banner</td>
<td>5.49</td>
<td>5.68</td>
<td>0.10</td>
<td>30</td>
<td>494</td>
</tr>
<tr>
<td>square</td>
<td>4.90</td>
<td>6.41</td>
<td>0.10</td>
<td>90</td>
<td>612</td>
</tr>
<tr>
<td>skyscraper</td>
<td>8.32</td>
<td>7.76</td>
<td>0.25</td>
<td>40</td>
<td>150</td>
</tr>
<tr>
<td>Unknown</td>
<td>4.26</td>
<td>4.33</td>
<td>0.10</td>
<td>35</td>
<td>843</td>
</tr>
</tbody>
</table>

*The size of digital ads are measured in density-independent pixels (dips).*
Research Plan and Difficulty

\[ H \rightarrow D \rightarrow V \]

- **\( H \)**: average multi-homing level of a publisher, from comScore
- **\( D \)**: unwanted duplication, unobserved
- **\( V \)**: advertiser’ WTP for ad slots. Ad prices are scraped from BuySellAds
- Goal: identify the marginal effect \( V_H \)
- Endogeneity concerns:
  - Unobserved publisher characteristics (e.g. quality)
  - Unobserved reader characteristics (e.g. demographics, engagement)
  - Publisher strategic pricing
  - Advertiser selection
Quasi-experiment

The experiment set-up:
- Treated group: “top” ad slots (denote T)
- Control (or less treated) group: other ad slots (denote B)
- Continuous treatment: multi-homing level $H$
- Outcome variable: ad prices
- $\Delta V_H = V^T_H - V^B_H$ is expected to $< 0$
- Essentially a diff-in-diff approach

Identification: top and bottom ad slots are balanced
Assessment on the Identification Strategy

- Most websites have both top and bottom ads, hence both groups contain similar publishers.
- Within each publisher, the reader profiles of top and bottom ads are roughly balanced.
- Publishers may employ different pricing strategies for the top and bottom slots, but the price difference due to the strategic behavior is unlikely to be systematically correlated with their multi-homing levels.
- Advertisers may select different ads to show on top and/or bottom slots... but how does this affect the estimation of multi-homing effect?
Illustrative Example

Website 1

Top Slot
(Probability for an effective view = 0.8)

Bottom Slot
(Probability for an effective view = 0.4)

Website 2

Top Slot
(Probability for an effective view = 0.8)

Bottom Slot
(Probability for an effective view = 0.4)

Exp. Impressions for a Multi-homer

1.6

0.8
Limitations: Advertiser Selection

- Advertisers may endogenously choose fewer publishers when having all ads on the top slots, or they may pair top slots in some websites with bottom slots in other websites, to balance the effective exposures
  - The heterogeneous treatment is attenuated

- Advertisers may endogenously pair “more durable” ad content to “more viewable” slots (top) to prevent over exposures
  - Those more durable ads are less responsive to multi-homing – this effect works against the hypothesis

- Bottom line: my result will be a conservative estimation
Econometrics

\[ P_{ijt} = \beta_0 + \gamma T_{ijt} + f^n(H_{it} | \delta) + f^n(T_{ijt} \times H_{it} | \lambda) + S'_{ijt} \beta_1 + X'_{it} \beta_2 + \nu_i + \tau_t + \epsilon_{ijt} \]

- Publisher \( i \), ad slot \( j \), and month \( t \)
- \( P_{ijt} \): cpm prices
- \( T_{ijt} = 1 \) if treated (top ad)
- \( H_{it} \): multi-homing level
- \( f^n(H_{it} | \delta) = \delta_1 H_{it} + \delta_2 H_{it}^2 + ... + \delta_n H_{it}^n \)
- \( f^n(T_{ijt} \times H_{it} | \lambda) = \lambda_1 (T_{ijt} \times H_{it}) + \lambda_2 (T_{ijt} \times H_{it}^2) + ... + \lambda_n (T_{ijt} \times H_{it}^n) \)
- \( S_{ijt} \): vector of ad specs (e.g. size, shape)
- \( X_{it} \): vector of publisher characteristics (e.g. reader demographics)
- The key parameters of interest is \( \lambda \)
## Results for Various Multi-homing Measures

<table>
<thead>
<tr>
<th></th>
<th>Narrow $ln(cpms)$</th>
<th></th>
<th>Broad $ln(cpms)$</th>
<th></th>
<th>Universal $ln(cpms)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>0.151</td>
<td>0.331***</td>
<td>0.0727</td>
<td>0.227</td>
<td>-0.0686</td>
</tr>
<tr>
<td></td>
<td>(1.38)</td>
<td>(3.09)</td>
<td>(0.57)</td>
<td>(1.29)</td>
<td>(-0.12)</td>
</tr>
<tr>
<td>$ln(H)$</td>
<td>0.195</td>
<td>0.855**</td>
<td>0.0552</td>
<td>0.581</td>
<td>0.00970</td>
</tr>
<tr>
<td></td>
<td>(0.97)</td>
<td>(2.38)</td>
<td>(0.30)</td>
<td>(1.39)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>$ln(H)^2$</td>
<td>-0.389***</td>
<td></td>
<td>-0.294*</td>
<td></td>
<td>-0.499**</td>
</tr>
<tr>
<td></td>
<td>(-2.67)</td>
<td></td>
<td>(-1.74)</td>
<td></td>
<td>(-2.17)</td>
</tr>
<tr>
<td>$T \times ln(H)$</td>
<td>0.00848</td>
<td>-0.782***</td>
<td>0.130</td>
<td>-0.373</td>
<td>0.0711</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(-3.23)</td>
<td>(0.81)</td>
<td>(-0.99)</td>
<td>(0.40)</td>
</tr>
<tr>
<td>$T \times ln(H)^2$</td>
<td>0.527***</td>
<td></td>
<td>0.304*</td>
<td></td>
<td>0.347</td>
</tr>
<tr>
<td></td>
<td>(4.52)</td>
<td></td>
<td>(1.86)</td>
<td></td>
<td>(1.66)</td>
</tr>
<tr>
<td>Ad Specs</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Month FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Publisher FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Exact Match</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>User base</td>
<td>$\geq 20$</td>
<td>$\geq 20$</td>
<td>$\geq 20$</td>
<td>$\geq 20$</td>
<td>$\geq 20$</td>
</tr>
<tr>
<td>Publishers</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>Observations</td>
<td>360</td>
<td>360</td>
<td>360</td>
<td>360</td>
<td>360</td>
</tr>
</tbody>
</table>
Estimated $\Delta V_H$
**Table:** Robustness on Publisher Selection and Match

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(cpm)</td>
<td>0.442***</td>
<td>0.279**</td>
<td>0.322***</td>
<td>0.337***</td>
<td>0.388</td>
</tr>
<tr>
<td></td>
<td>(4.35)</td>
<td>(2.56)</td>
<td>(2.83)</td>
<td>(3.21)</td>
<td>(1.49)</td>
</tr>
<tr>
<td>ln(H)</td>
<td>0.396</td>
<td>0.592*</td>
<td>0.924**</td>
<td>0.759**</td>
<td>1.48</td>
</tr>
<tr>
<td></td>
<td>(1.06)</td>
<td>(1.79)</td>
<td>(2.37)</td>
<td>(2.22)</td>
<td>(1.60)</td>
</tr>
<tr>
<td>ln(H)^2</td>
<td>0.0722</td>
<td>-0.274**</td>
<td>-0.426***</td>
<td>-0.365***</td>
<td>-0.901**</td>
</tr>
<tr>
<td></td>
<td>(0.34)</td>
<td>(-2.08)</td>
<td>(-2.70)</td>
<td>(-2.62)</td>
<td>(-2.02)</td>
</tr>
<tr>
<td>T \times ln(H)</td>
<td>-0.846***</td>
<td>-0.456*</td>
<td>-0.663***</td>
<td>-0.778***</td>
<td>-0.678</td>
</tr>
<tr>
<td></td>
<td>(-3.20)</td>
<td>(-1.80)</td>
<td>(-2.80)</td>
<td>(-3.33)</td>
<td>(-0.95)</td>
</tr>
<tr>
<td>T \times ln(H)^2</td>
<td>0.575***</td>
<td>0.343***</td>
<td>0.451***</td>
<td>0.526***</td>
<td>0.630</td>
</tr>
<tr>
<td></td>
<td>(3.63)</td>
<td>(2.67)</td>
<td>(4.23)</td>
<td>(4.72)</td>
<td>(1.41)</td>
</tr>
</tbody>
</table>

Ad Specs: Yes, Yes, Yes, Yes, Yes
Month FE: Yes, Yes, Yes, Yes, Yes
Publisher FE: Yes, Yes, Yes, Yes, No
Exact Match: Yes, Yes, Yes, No, Yes

User base: ≥ 3, ≥ 10, ≥ 30, ≥ 20, ≥ 20
Observations: 823, 496, 291, 425, 360
Publishers: 209, 113, 62, 84, 75
How to Interpret the Positive Trend?

\[
H \rightarrow D \rightarrow V
\]

- Expect that $\Delta V_H < 0$, but it is the case only when $\ln(H^{\text{narrow}}) < 0.5$ (about 50% of the data)

- Advertiser selection on the “durable” content leads to a stronger counter effect when the multi-homing level gets higher
  
  \[
  \Delta V_H = \frac{\partial D_T}{\partial H} \frac{\partial V_T}{\partial D} - \frac{\partial D_B}{\partial H} \frac{\partial V_B}{\partial D}
  \]

  \[
  \frac{\partial D_T}{\partial H} \quad \text{and} \quad \frac{\partial D_B}{\partial H}
  \]
  
  are positive and remain approximately constant

  \[
  \frac{\partial V_B}{\partial D}
  \]
  
  becomes relatively more negative than $\frac{\partial V_T}{\partial D}$

- The multi-homing effect is strong enough to overcome the advertiser selection initially – closest to the unbiased estimation
Auxiliary Models

- Two reasons to include the publisher fixed effect
  - Remove the time-invariant endogenous error for identifying the main effect
  - Improve precise matching on the comparison of multi-homing effect
  - Problem: it could remove too much information...

- I explore two auxiliary models
  - Two-sided market model: I utilize the detailed consumer-side information to retrieve necessary control variables
  - Instrumental variable model: I use the multi-homing pattern far from the examining genre as the instrument
  - Both models report similar results as the main model
Trade-off between Advertising Effectiveness and Privacy Protection

- Privacy regulation particularly affects the online display ad market (Budak et al. 2016)

- Previous papers
  - Goldfarb and Tucker (2011) found that after the “EU Privacy Directive” implemented in 2002, online display advertiser may need to spend 2.85 times as much advertising to reach the same purchase of intent
  - Johnson (2013) assessed the economic impact of tracking ban for the RTB market – revenue would drop by about 40% for both publishers and advertisers

- My result can reflect the impact in the “programmatic direct” market via the channel of frequency management... let’s push it further...
Back-of-the-envelope Calculation (1)

- If perfect tracking were allowed, the multi-homing effect will be nullified
  - just like browsing different webpages within a huge website...

- The diff-in-diff result indicates that at least the price for top slots will increase

- The price for top ad slots will increase by 21.5% of the bottom price
  - Use the first-order interaction term as the diff-in-diff result ($-0.782$)
  - Use the multi-homing level of an average website in my sample (1.88)
  - $\ln(cpm)$ of top slots would increase by roughly $0.782 \times \ln(1.88) = 0.215$ more than bottom slots for an average website
Back-of-the-envelope Calculation (2)

Status quo:
- Avg. cpm of bottom slots: $3.9
- Avg. cpm of top slots: $6.5

Perfect tracking:
- Bottom slots: $3.9 (to be conservative)
- Top slots: $6.5 + 0.215 \times $3.9 = $7.34

Simple add-up to the revenue of the “programmatic direct” display ad market
- Assume the total impressions remain equal for top and bottom slots
- Ad revenue in 2016 could increase from $14 billion to $15.13 billion
- 7% increase
Media Content Strategy

- Each publisher focuses on different niche content – to attract single-homing consumer (Anderson et al. 2015)...
  - Contrast to classic argument “popular contents is often overproduced” in Steiner (1952)

- Publishers may have the incentive to attract single-homing advertiser by producing wide (but shallower) contents (Athey et al. 2016)...
  - My result suggests that the width due to consumer multi-homing is limited to a narrow genre

- If privacy regulation becomes stricter, the analyses becomes more relevant
Conclusion

- Examine whether consumer multi-homing lowers ad prices due to difficult frequency management across websites

- Use data from BuySellAds and comScore 2016, and SimilarWeb

- Employ a quasi-experiment (diff-in-diff method) based on the webpage location of ad slots to find the (conservative) causal estimation

- Results are robust to publisher selection and empirical models
  - Multi-homing within a narrow genre lowers ad prices – corresponding to the rationale and analyses in literature
  - Behavioral tracking in my sample publishers (middle class) is imperfect
  - Perfect tracking increases ad revenue (program. direct) by 7% ($1 billion)

- Potential extension
  - Empirically examine the non-market impact of privacy regulation (e.g. media content and consumption pattern)
Appendix
FTC’s Self-Regulatory Principle of Online Behavioral Advertising

Staff agrees that “first party” (or “intra-site”) behavioral advertising practices are more likely to be consistent with consumer expectations, and less likely to lead to consumer harm...

By contrast, when behavioral advertising involves the sharing of data with ad networks or other third parties, the consumer may not understand why he has received ads from unknown marketers based on his activities at an assortment of previously visited websites...
Interpretation

Given assumptions:

1. $\frac{\partial D^T}{\partial H} \geq \frac{\partial D^B}{\partial H} > 0$ or $\frac{\partial D^T}{\partial H} = \frac{\partial D^B}{\partial H} = 0$
2. $\frac{\partial V^T}{\partial D} \geq \frac{\partial V^B}{\partial D}$ due to advertiser selection
3. $V(D)$ is unimodal for both groups (Vakratsas et al. 2004)
4. $D(H)$ is approximately linear

$\Delta V_H = \frac{\partial D^T}{\partial H} \frac{\partial V^T}{\partial D} - \frac{\partial D^B}{\partial H} \frac{\partial V^B}{\partial D} < 0$ implies

1. $\frac{\partial D^T}{\partial H} > \frac{\partial D^B}{\partial H} > 0$, heterogeneous treatment exists and consumer multi-homing causes duplication
2. $\frac{\partial V^B}{\partial D} < \frac{\partial V^T}{\partial D} < 0$, Advertisers indeed endogenously pair durable ads to top slots to avoid over-exposure duplication decreases advertiser valuation
3. The multi-homing effect is strong enough to overcome the advertiser selection initially, but the selection effect gets stronger as multi-homing level increases
Two-sided Market Model

Linear demand system in the advertising market ($t$ suppressed)

\[ P_{ij} = \beta_0 + \gamma T_{ij} + f^n(H_i | \delta) + f^n(T_{ij} \times H_i | \lambda) + S'_{ij} \beta_1 + X'_i \beta_2 + \pi a_i + \pi q_i \] + \epsilon_{ij}

- $q_i \leftarrow$ market share in the consumer market, e.g. $q_i = log\left(\frac{n_i}{m_g - n_i}\right)$
  - $n_i$ number of unique visitors to publisher $i$ in comScore
  - $m_g$ number of unique visitors to genre market $g$ in comScore
- $a_i \leftarrow$ average visiting duration, e.g. $a_i = \frac{1}{\alpha_g}(duration_i - \phi_g q_i - \omega_g)$
  - $\alpha_g$: genre-specific ad preference parameter (Kaiser and Song, 2009)
  - $\phi_g$: genre-specific quality parameter
  - $\omega_g$: genre-specific time constraint parameter
Instrumental Variable Model

▶ IV1: the multi-homing level outside the broad genre market
\( H_{out} = H_{universal} - H_{broad} \) as the instrument for \( H_{narrow} \)
  ▶ Relevance: general tendency of multi-homing for a consumer
  ▶ Exclusion: publisher characteristics, content quality, or other strategies do not affect the multi-homing pattern in the almost irrelevant genre

▶ IV2: total time spent online
  ▶ Relevance: more time online may increase multi-homing level
  ▶ Exclusion: single publisher cannot affect the total time spent online
<table>
<thead>
<tr>
<th></th>
<th>ln(cpm)</th>
<th>OLS</th>
<th>IV-2SLS</th>
<th>IV-LIML</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>0.438***</td>
<td>0.329***</td>
<td>0.300**</td>
<td>0.697***</td>
</tr>
<tr>
<td></td>
<td>(4.70)</td>
<td>(2.86)</td>
<td>(2.55)</td>
<td>(3.82)</td>
</tr>
<tr>
<td>$\ln(H)$</td>
<td>1.154***</td>
<td>1.508***</td>
<td>2.202***</td>
<td>3.757***</td>
</tr>
<tr>
<td></td>
<td>(3.62)</td>
<td>(3.01)</td>
<td>(4.05)</td>
<td>(3.11)</td>
</tr>
<tr>
<td>$\ln(H)^2$</td>
<td>-0.568***</td>
<td>-0.687***</td>
<td>-0.892***</td>
<td>-1.591**</td>
</tr>
<tr>
<td></td>
<td>(-3.71)</td>
<td>(-2.90)</td>
<td>(-3.27)</td>
<td>(-2.54)</td>
</tr>
<tr>
<td>$T \times \ln(H)$</td>
<td>-0.396**</td>
<td>-0.429</td>
<td>-0.912***</td>
<td>-1.574**</td>
</tr>
<tr>
<td></td>
<td>(-1.99)</td>
<td>(-1.44)</td>
<td>(-2.76)</td>
<td>(-2.06)</td>
</tr>
<tr>
<td>$T \times \ln(H)^2$</td>
<td>0.379***</td>
<td>0.417**</td>
<td>0.757***</td>
<td>1.104*</td>
</tr>
<tr>
<td></td>
<td>(3.31)</td>
<td>(2.10)</td>
<td>(3.13)</td>
<td>(1.92)</td>
</tr>
</tbody>
</table>

|                | Yes         | Yes         | Yes         | Yes         | Yes         |
| Ad specs       |             |             |             |             |             |
| Demographics and priority | Yes         | Yes         | Yes         | Yes         | Yes         |
| Month FE       | Yes         | Yes         | Yes         | Yes         | Yes         |
| Genre FE       | Yes         | Yes         | Yes         | Yes         | Yes         |
| Genre FE × quality | Yes         | Yes         | Yes         | No          | No          |
| Genre FE × duration | Yes         | Yes         | Yes         | No          | No          |

| Cragg-Donald Wald F Value | . | . | . | 11.29 | 11.29 |
| Hansen’s J Statistic     | . | . | . | 0.10  | 0.13  |

| User base | ≥ 3 | ≥ 10 | ≥ 20 | ≥ 3  | ≥ 3  |
| Publishers | 300 | 172  | 116  | 300  | 300  |
| Observations | 1585 | 986  | 714  | 1585 | 1585 |