ACCC Digital Platforms Inquiry Submission

Catherine Tucker, Ph.D.

27/11/2018
1 Introduction

The ACCC’s Digital Platforms Inquiry is exploring, among other things, the ‘impact of digital search engines, social media platforms and other digital content aggregation platforms on the state of competition in media and advertising services markets.’\(^1\) I have been retained by Facebook to provide my perspective on questions raised by the ACCC related to the extent to which digital platform providers are deriving ‘enduring competitive advantage’\(^2\) in advertising markets from their possession of large amounts of data.\(^3\) In considering this question, I have relied on my own academic research on issues related to online platforms, digital advertising and big data. My CV is provided as Attachment A.

In Section 2, I discuss how digital advertising addresses many of the flaws of traditional advertising markets. In Section 3, I consider whether big data in digital advertising confers an enduring competitive advantage to its owner. I separately explore whether big data is rare, whether big data is inimitable, whether big data is independently valuable, and whether big data is sufficient and necessary to ensure success in digital advertising markets. Finally, in Section 4, I describe the relationship between reach—the ability to target many eyeballs, and relevance—the ability to target the right eyeballs, and explain how automated digital advertising means that the traditional reach/relevance trade-off has been diminished, changing competitive dynamics in the industry. Firms can now use data to target the right eyeballs on multiple websites in parallel, allowing them to achieve both reach and relevance. I also discuss the relationship between big data and other traditional sources of market power, such as network effects or switching costs.

\[^2\] ACCC Issues Paper, at 10.
\[^3\] I have consulted for Facebook and other digital platform companies in the past. See ‘Catherine Tucker - Disclosure,’ MIT Sloan School of Management, available at: http://mitmgmtfaculty.mit.edu/cetucker/disclosure/.
2 The Transformation of the Advertising Industry

In 2018, global advertising expenditure is estimated to grow by 4.6% to approximately 800 billion AUD. In Australia, advertisers spent approximately 16 billion AUD in 2017. The advertising market has not only grown in size but also transformed in terms of the value it offers to brands. In the past, advertising was a frustrating part of a firm’s operations, because it was very expensive and yet its direct impact on sales could not be measured. Firms suspected that large amounts of money were wasted because consumers saw ads for products or services they were not interested in. Consumers were annoyed because they were bombarded with ads they did not find interesting or useful. Furthermore, the fact that advertising in general required large amounts of money restricted the ability of smaller firms to grow, as they could not afford to advertise.

At its highest level, digital advertising addresses these flaws:

1. The digital environment allows advertisers to show ads to specific consumers and to measure whether these ads have any effect.

2. The digital environment allows advertisers to identify consumers who are more likely to be interested in their ads. Correspondingly, consumers are less likely to be irritated by irrelevant and annoying ad content.

3. The digital environment makes it easier for advertisers with different levels of experience and revenues to experiment with new services and try to establish which advertising channel delivers the best return on investment (ROI) and allocate their ad dollars accordingly.

The digital advertising world is a complex one, with many firms competing to add value to marketers by the ability to target specific consumers or groups of consumers, present ads in a way that does not lead them to tune out or become irritated, and measure whether their ads actually worked.

---


2.1 An Initial Example

Let us imagine a local Australian store that sells furniture. Its primary marketing challenge is that most of the time people are not that interested in buying furniture, so the store would ideally like to show ads only to people who are moving or redecorating their house. New advertising technology now means that the firm can easily identify people online who are in the ‘moving’ segment.\(^6\) The furniture store could buy ad services from a firm like Unruly, a global data analytics company that has offices in Sydney and Melbourne. Unruly uses data from major publishers (of websites), such as News Corp, to identify people who are interested in content related to moving.\(^7\) In a recent survey of people who viewed ads, Unruly found that by increasing the relevance of the ads, consumers who were moving house or redecorating were more likely to be inspired by the ads (an increase of 50%).\(^8\) This suggests that relative to untargeted advertising, such targeting technology has the potential to be welfare-improving for consumers.

2.2 The Broad Digital Advertising Ecosystem

Unruly is just one example of a firm operating in a large and complex digital advertising ecosystem. Aiding the deployment of targeting technology to help advertisers identify the right customers is a large industry composed of many different players. Increasingly, boundaries between these players are fluid and evolving, but it is worthwhile to explain in broad terms the functions of different categories of players in the industry:

- Demand-side platforms are platforms that help advertisers put in bids to show their ads to a certain pair of eyeballs arriving at a certain website in real time (e.g., MediaMath).\(^9\)

---


\(^7\) Unruly also conducts surveys of people’s intent and offers a testing tool that allows advertisers to test which of their campaigns appears most appealing—such testing is easy in a digital environment where it is easy to run multiple campaigns and measure what works best.

\(^8\) According to Unruly’s Home Mover Study, ‘People in the process of finding a new home were more likely to feel inspired (+50%), exhilarated (+96%) and proud (+69%) when watching ads; Those making home improvements were most likely to feel inspired (+49%), proud (+65%) and happy (+33%) when watching ads.’ See Waterhouse, David, ‘Unruly Launches First-Party Data Segments To Help Advertisers Engage Australian Home Movers,’ Unruly, 26 September 2017, available at: https://unruly.co/news/article/2017/09/26/unruly-first-party-data-australian-home-movers.

• Supply-side platforms allow publishers of websites to receive bids for showing a particular ad to a particular set of eyeballs browsing their website in real time to make sure they can maximise the profitability of their content (e.g., PubMatic).\footnote{PubMatic is currently the largest SSP in Australia. See ‘Supply Side Platforms Market Share Table,’ \textit{Datanyze}, available at: https://www.datanyze.com/market-share/ssps/Australia/. PubMatic recently partnered with Bonzai to improve ease of access to its services via self-service platforms. See ‘PubMatic and Bonzai Partner on Rich Media in Australia and New Zealand,’ \textit{PubMatic}, 23 August 2018, available at: https://pubmatic.com/news/pubmatic-bonzai-partner-rich-media/.}

• Data management platforms sit between these two and allow advertisers to ensure they have the right data to identify whether someone is the right pair of eyeballs for a particular campaign. For example, advertisers might want to focus on people they know have already visited their website or perhaps, in the moving example, they may want to use other data on projected income or taste to work out what kind of furniture might be of interest to that person (e.g., Lotame).\footnote{The biggest DMPs in Australia are Lotame, LiveRamp, Oracle and Adobe. See ‘Top Competitors of Lotame in Datanyze Universe,’ \textit{Datanyze}, available at: https://www.datanyze.com/market-share/dmp/Australia/lotame-market-share.}

• Data brokers collate data about users from a large swathe of offline and online sources. To obtain relevant data in addition to data the advertiser may already own, the advertiser or advertising service-provider can purchase data about a particular pair of eyeballs from a data broker (e.g., Experian).\footnote{Experian, available at: http://www.experian.com.au/}

Together these services allow publishers of websites and advertisers to connect together—usually through means of an auction. In these auctions, an advertiser can bid in real-time to show a certain ad to a set of specific eyeballs—for example, eyeballs that the data suggests may have moved house recently and therefore belong to someone who is in the market for furniture.\footnote{Usually these auctions operate as “second-price” auctions, where the winner pays the price of the second-highest bid. More recently, platforms have experimented with a first-price auction format in which the highest bidder wins the auction and pays the amount that it bids. See Chen, Yuyu, ‘Programmatic advertising is preparing for the first-price auction era,’ \textit{Digiday}, 5 October 2017, available at: https://digiday.com/marketing/programmatic-advertising-readying-first-price-auction-era/.}

Sometimes these services are offered by a single firm—for example Oath (the new name for the combination of AOL, Verizon and Yahoo) offers all these services, while Google, Twitter, and Facebook offer an assortment of these services. Often on top of these basic services, firms in this space offer enhanced services—for example, Rakuten, a firm that has data on over 1.2 billion
consumers worldwide, has supplemented its demand side, data management and data offerings with AI and machine learning to improve data integrity.\textsuperscript{14}

In this paper, I use digital advertising as an umbrella term to refer to the ability of a combination of these technologies to deliver advertisers the right person at the right time for the right ad.

3 Big Data and Enduring Competitive Advantage in Digital Platforms

As a consequence of providing advertisers with the benefits of measurability, relevance and targeting, the digital advertising industry processes and uses large amounts of data. Therefore, a natural question is whether access to large amounts of data confers an enduring competitive advantage for the firms that have access to this data or whether, conversely, a large firm in this industry is likely to have access to large amounts of data to service the needs of clients, but the data does not confer any particular enduring advantage. This is a useful question because, of course, if large amounts of data confer enduring competitive advantage to firms, then lack of access to such data could create a barrier to entry for new firms. It also parallels a discussion about whether big data is an ‘essential facility’ in the digital advertising era.\textsuperscript{15}

When I teach MBAs about how to obtain sustainable competitive advantage, I use a strategy framework that enumerates criteria for whether or not something is indeed going to prove to be a source of long-term competitive advantage, called the ‘resource-based view of the firm’.\textsuperscript{16} Specifically, for a firm resource to be a source of competitive advantage, the resource has to be rare, inimitable, valuable and non-substitutable. I consider each of these in turn.


\textsuperscript{15} For example, some have argued that the essential facilities doctrine should require open access to data. \textit{See} Abrahamson, Zachary, ‘Essential Data,’ \textit{The Yale Law Journal}, Vol. 124, No. 3, December 2014, pp. 867-881, available at: https://www.yalelawjournal.org/comment/essential-data. However, others have called this argument into question. For example, in a recent speech on consumer data access regimes and data sharing obligations, ACCC Chairman Rod Sims acknowledged that data ‘does not display the same characteristics as essential facilities type infrastructure’ because it can easily be duplicated. He noted that the merits of such a regime depend on the ‘extent to which access to data is necessary for effective competition.’ \textit{See also} Sims, Rod, ‘Gilbert & Tobin seminar: the data economy,’ 15 October, 2018, \textit{ACCC}, available at: https://www.accc.gov.au/speech/gilbert-tobin-seminar-the-data-economy.

3.1 Is Big Data Rare?

If a business resource like ‘big data’ is widely available, it is unlikely to confer a competitive advantage.

The extent to which ‘big data’ could be viewed as rare in advertising markets is limited because of the widespread nature of a customer’s digital footprint. Any time I browse or buy a product, online data about my behaviour is observed by multiple different firms. Though the data itself may be obtained from different sources, the key question is whether the data can give similar insight into my likely intent towards a product or service.

As I discuss below, this kind of information is sold to advertisers by multiple firms in the digital space, including data brokers and traditional firms.

3.1.1 What Do Data Brokers Do?

Data brokers (or data aggregators) collect data on individuals (usually identified by a cookie) and resell this information to advertisers, publishers, and firms that provide advertising services. They often use broad data on a consumer’s digital footprint to make inferences about what customer segment that pair of eyeballs may be in. For example, one potentially useful segment to advertisers is the segment of consumers who might be interested in buying a camera. To identify such a segment, a data broker would use data on browsing behaviour to look for signs that that pair of eyeballs is potentially in the market for a camera.

To see what kinds of insights marketers are looking for when it comes to the data they use in advertising, it is useful to turn to a recent survey commissioned by Lotame of 300 brand marketers that use audience data. The most popular data types purchased by these marketers were: Demographic - age and gender (42%), Geographic (34%), Advanced Demographic - income, education, children (28%), Interest (28%), Behavioural (25%), and Social Influencer (24%). Respondents separately ranked their use of demographic audiences. The highest ranked segment types (by percent of respondents who indicated that they ‘usually’ or ‘always’ target based on this segment type) were: Age (76%), Gender (61%), Household Income (50%), Education (40%), and Number of Children in Household (32%).

3.1.2 What Kinds of Digital Data Do Data Brokers Draw Upon?

When a website publisher uses browsing data to target ads that it controls, that data is referred to as ‘first-party’ data. For example, if a travel firm such as Rough Guides uses my browsing data for regions of Malaysia to target ads for related travel products to me, then Rough Guides would be using first-party data.

When a website shares browsing data with another website or firm that knows its exact providence, then that data is referred to as ‘second-party’ data. For example, Lastminute.com (a travel website) and Rough Guides (travel guide books) explicitly share data in order to promote cross-selling opportunities.  

When a data broker buys browsing data and collates this data as part of a broader data collection exercise about a particular individual, that data is referred to as third-party data. There are many businesses who create value for advertisers by combining these different data sources and partnering with other organisations. For example, Rough Guides might buy data from a data broker such as Experian, Acxiom, or Epsilon to identify segments of people who recently purchased travel products or travel frequently. What is key about the services of a data broker is that they allow an advertiser, a website or a new firm offering advertising services to match their client list with this data, easing the ability of advertisers to place ads on new websites or use new advertising technologies.

To see how these sources of data translate into multiple options for advertisers who want to use data, let us suppose, for example, that I am thinking about buying a camera. Websites (perhaps review websites or websites with professional content) that I visit are aware of my intention and can record the fact that I was browsing articles that were giving advice about new cameras. They can use this data themselves to target ads (first-party data), share this data with a second-party they know (second-party data) or resell this data to data brokers who create audience segment profiles (third-party data). The market for this data is large, and advertisers report that it is not expensive to obtain such data from publishers.

---

As well as data from websites, there are also other online players providing other services that might have access to this insight. For example, a browser, VPN service or toolbar might record information relating to any product service I might pursue. Anti-virus software can also track browsing history and searches.\textsuperscript{21} If I use my mobile phone, again, a variety of apps might collect data about my sudden interest in cameras. In addition, internet service providers themselves might also have access to such data.\textsuperscript{22} All of these websites and services would be aware that it would be timely to show me ads for a camera. Data brokers help the process of collating this data, aggregating it and then reselling it to an advertiser, website or advertising services firm to aid in showing a camera ad only to individuals who are likely to be interested in it.


\textsuperscript{22} At the moment, the U.S. is leading this experiment but experts in Australia believe that this may be a possibility in Australia as well. O’Brien, Jennifer, ‘ISPs could sell your browsing history: Aussie expert weighs in on US proposal,’ \textit{CIO}, 12 April 2017, available at: https://www.cio.com.au/article/617575/isps-could-sell-your-browsing-history-aussie-expert-weighs-us-proposal.
3.1.3 Data Brokers Offer Enriched Customer Segment Data

Data brokers don’t just use data from browsing behaviour to identify what ad may be relevant but also enrich their data with demographic data and spending behaviour offline. Figure 1 gives an idea of the demographic data that is available for each person on Acxiom’s ‘InfoBase’ database. As a marketer for cameras, I might believe parents of toddlers will be more likely to respond to a camera advertisement, and therefore I may use ‘InfoBase’ to filter my audience data based on the ‘Child at Home 0-4 Years Old’ variable. If I am particularly interested in targeting parents of toddlers that also go on European holidays and have already demonstrated an interest in cameras, I can combine my selected demographic variable with the purchase-intent segment variables ‘Take European Holidays’ and ‘Digital Camera’. This means that an advertiser, publisher or advertising services provider doesn’t independently have access to this kind of demographic or interest-segment data—but instead they can purchase it from a provider such as Acxiom. Figure 2 shows these variables, along with the wide range of other variables advertisers can access. Acxiom describes its product ‘InfoBase’ in the following manner for the UK:

‘Infobase Enhancement—the leading consumer data-append product, InfoBase Enhancement supplies consumer descriptive data for use in analytic, segmentation and targeting applications. Hundreds of demographic, homeowner, buying behaviour, financial, motoring and interest variables enable you to segment, analyse and model consumer data, resulting in accurate targeting and more predictive modelling. With multiple data sources and sophisticated build logic, available across more than 90% of UK households, our enhancement capability is the most complete, comprehensive and accurate source of consumer data available.’

---


Figure 1

Demographic Data Available in Acxiom’s Infobase

Figure 2

Purchase Intent Segments Available in Acxiom’s Infobase


Acxiom also purchases offline transaction data to calculate a customer’s ability to spend money on a variety of categories. Figure 3 below is a screenshot from their Infobase affordability product estimates. Projections of spending on consumer electronics is useful for understanding not only whether someone is interested in buying a camera but also whether they can actually afford to buy it and are therefore worth advertising to. Acxiom is not the only data broker that provides this information—Epsilon (linked to Equifax, and which has operations in Australia) offers a similar service, which it describes in the following manner: ‘Our MarketView™ product offers exclusive access to a multi-sourced transactional dataset capturing $2T of consumer spend across hundreds of leading merchants. Identify and reach your most valuable customers, and learn what they spend with both you and your top competitors.’

**Figure 3**

Spending Projection Categories Available in Acxiom’s Infobase

**RECREATION AND LEISURE**
- Eating and Drinking Out: Household’s likely spend on eating and drinking out index to UK average.
- Holidays Abroad: Household’s likely spend on holidays abroad.
- Holidays in the UK: Household’s likely spend on holidays in the UK.
- Holidays: Household’s likely total spend on holidays.
- Betting and Gambling: Household’s likely spend on betting and gambling.
- Entertainment and Recreation: Household’s likely spend on entertainment and recreation.
- Gardening: Household’s likely spend on gardening.
- Pets: Household’s likely spend on pets (excluding food).
- Sports and Leisure: Household’s likely spend on sports admissions, leisure classes and equipment hire.
- Recreation and Leisure: Household’s likely total spend on other leisure and recreation activities.

**CONSUMER PACKAGED GOODS**
- Food: Household’s likely spend on food.
- Alcohol at Home: Household’s likely spend on alcohol at home.
- Consumable Household Products: Household’s likely spend on consumable household goods and cleaning products.
- Personal Goods: Household’s likely spend on personal goods, toiletries etc.
- Pet Food: Household’s likely spend on pet food.
- Consumer Packaged Goods: Household’s likely total spend on food, drink, pet food, toiletries, cleaning products, tobacco etc.

**CLOTHING, FOOTWEAR AND PERSONAL EFFECTS**
- Women’s Clothing, Footwear and Personal Effects: Household’s likely spend on women’s clothing, footwear and personal effects.
- Men’s Clothing, Footwear and Personal Effects: Household’s likely spend on men’s clothing, footwear and personal effects.
- Children’s Clothing, Footwear and Personal Effects: Household’s likely spend on children’s clothing, footwear and personal effects.
- General Personal Effects: Household’s likely spend on general personal effects.
- Total Clothing, Footwear and Personal Effects: Household’s likely total spend on clothing, footwear and personal effects.

**EXPENDITURE MEASURES**
- Income by Outgoings: Segments based on level of outgoings within net household income decile.
- Proportion of Income Spent: Proportion of net household income spent (total outgoings as proportion of net income).
- Proportion of Income Fixed: Proportion of income committed.
- Proportion of Income Discretionary: Proportion of income discretionary.
- Committed Spend Index: Food/drink Spend Index.
- Total Spend Index: Household’s Standard of Living.
- Indulgence Rank: Asset Rank.
- Income by Asset level: ILIU segmented by income and relative asset level within the income band.

Experian also gathers other online and offline data to combine with its own consumer credit and transaction data to create a database with more than 500 demographic and behavioural variables on customers to identify relevant marketing segments through its customer segmentation tool Mosaic. Australian households are then organised into 49 Mosaic types (within 13 groups) such as ‘Inner City Aspirations,’ ‘Coastal Contentment,’ and ‘Sensible Seniors’. Figure 4 below is a screenshot highlighting the key features of the ‘Suburban Elites’ segment from Mosaic Australia, many of which would be knowable to the credit agency without purchasing third-party data. Figure 5, however, presents browsing information purchased by Experian that shows changes in interest in the ‘Photography’ website category by segment, where ‘Suburban Elites’ are represented by the left-most purple bar. In effect, the integration of Experian’s own credit data with purchased third-party data about consumers’ interest in photography enables camera marketers to narrowly target a particular Australian demographic that has already demonstrated an interest in cameras.

Figure 4

Experian Mosaic: Example of an Australian Customer Segment

Figure 5
Experian Mosaic: Change in Interest in Photography by Segment

Source: ‘Mosaic Australia Interactive Guide,’ Experian Australia, available at:
Oracle’s Data Cloud, known as Datalogix before its acquisition in 2014, is another data broker that combines online and offline data to offer its own Oracle-branded consumer data products to marketers, as well as access to ‘5 billion global IDs, $3 trillion in consumer transactions, and more than 1,500 data partners.’ Figure 6 shows a list of Oracle Technology audiences, including a ‘Cameras & photography’ audience. This audience is curated by combining data on ‘offline transaction history’ with ‘intent signals based on users searching for consumer technology products on e-commerce sites or conducting product reviews’ and ‘online behavior including search browse.’ In 2017, Oracle announced that it had added 400 Australia-specific customer segments covering 64 million devices to Oracle Audiences.

---


**Figure 6**  
Oracle: Technology Audiences

In the past the ability to directly link offline behaviour with online data for individual consumers has been limited, but this is changing. For example, many firms that collect customers’ email addresses are then able to use that information to match a loyalty card used offline with an online profile. There are other firms that allow companies to explicitly enrich their own email data. For example, Towerdata’s ‘Email Intelligence’ product allows firms to understand better the characteristics of their customers who have shared their emails with the firm. It allows a firm to buy data to determine demographics for their customers such as age, location, income or gender, but also, crucially, the data show the customer’s current purchase intent. For example, one could use ‘Email Intelligence’ to determine not only the income, age, gender and address of a customer, but also whether they matched the description of ‘Consumers that are actively shopping online for dog supplies.’ This could be useful to an advertiser who is selling dog food as there is no point trying to sell dog food to someone who does not have (or plan to have) a dog.


Figure 7

Examples of Demographic, Household, and Purchase Intent Segments Offered by Towerdata

<table>
<thead>
<tr>
<th>Category</th>
<th>Field</th>
<th>Possible values</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic</td>
<td>Age</td>
<td>18 - 20; 21 - 24; 25-34; 35 - 44; 45 - 54; 55-64; 65+</td>
<td>Age Range</td>
</tr>
<tr>
<td>Demographic</td>
<td>Gender</td>
<td>Male, Female</td>
<td>Gender</td>
</tr>
<tr>
<td>Demographic</td>
<td>Postal Address</td>
<td>Street, city, state and zip</td>
<td>Address where person lives or works</td>
</tr>
<tr>
<td>Demographic</td>
<td>First Name</td>
<td>&lt;First&gt;</td>
<td>First name</td>
</tr>
<tr>
<td>Demographic</td>
<td>Last Name</td>
<td>&lt;Last&gt;</td>
<td>Last name</td>
</tr>
<tr>
<td>Household</td>
<td>Household Income</td>
<td>0-19k; 15-25k; 25-35k; 35-50k; 50-75k; 75k-100k; 100-125k; 125-150k; 150-175k; 175-200k; 200-250k; 250k+</td>
<td>Income of household by range</td>
</tr>
<tr>
<td>Household</td>
<td>Marital Status</td>
<td>Single; Married</td>
<td>Marital status</td>
</tr>
<tr>
<td>Household</td>
<td>Presence of Children</td>
<td>Yes; No</td>
<td>Indicates whether there are 1 or more children in the household</td>
</tr>
<tr>
<td>Household</td>
<td>Home Owner Status</td>
<td>Own; Rent</td>
<td>Whether the person owns or rents their home</td>
</tr>
<tr>
<td>Household</td>
<td>Home Market Value</td>
<td>1k-25k; 25k-50k; 50k-75k; 75k-100k; 100k-150k; 150k-200k; 200k-250k; 250k-300k; 300k-350k; 350k-500k; 500k-1mm; 1mm+</td>
<td>Market value of person's home. In ranges of $25K and $50K increments</td>
</tr>
<tr>
<td>Household</td>
<td>Length of Residence</td>
<td>&quot;Less than 1 year&quot;, &quot;1 Year&quot;, &quot;2 Years&quot;, &quot;3 Years&quot;, &quot;4 Years&quot;, &quot;5 Years&quot;, &quot;6 Years&quot;, &quot;7 Years&quot;, &quot;8 Years&quot;, &quot;9 Years&quot;, &quot;10 Years&quot;, &quot;11-15 years&quot;, &quot;16-19 years&quot;, &quot;20+ years&quot;</td>
<td>Number of years spent in the current residence. Reported as number; not range.</td>
</tr>
<tr>
<td>Household</td>
<td>Home Property Type</td>
<td>Single Family; Multifamily</td>
<td>The type of building the person resides in</td>
</tr>
<tr>
<td>Household</td>
<td>Net Worth</td>
<td>0-5k; 5k-10k; 10k-25k; 25k-50k; 50k-100k; 100k-250k; 250k-500k; 500k-750k; 750k-1mm; 1mm+</td>
<td>The approximate net worth of the household</td>
</tr>
<tr>
<td>Household</td>
<td>Occupation</td>
<td>Blue Collar Worker; Business Owner; Civil Service; Technology; Executive/Upper Management; Health Services; Homemaker; Middle Management; Military Personnel; Nurse; Part Time; Professional; Retired; Secretary; Student; Teacher; White Collar Worker</td>
<td>Occupation</td>
</tr>
<tr>
<td>Household</td>
<td>Education</td>
<td>Completed High School; Attended College; Completed College; Completed Graduate School; Attended Vocational/Technical</td>
<td>Completed High School; Attended College; Completed College; Completed Graduate School; Attended Vocational/Technical</td>
</tr>
</tbody>
</table>
3.1.4 Traditional Firms Also Offer Data Services

In addition to more obvious sources of online browsing data, many traditional companies also sell data (collected both online and offline) to advertisers. This more traditional data can also establish that I may be a good target for a camera ad. For example, it could be that my credit card
data makes me a likely prospect for a camera purchase. In Australia several banks also sell transaction information. Westpac, NAB and CBA already sell aggregated transaction data to their clients and NAB is considering selling ‘aggregated insights’ in the future. Qantas, which co-owns a data sharing platform called Data Republic, along with Westpac, NAB and ANZ, sells data gathered from over 10 million members of its loyalty program, purchases made through its Qantas Shop, and other sources.

The reuse of data that resides in existing Australian firms is also thriving. For example, Quantium, leveraging a partnership with Woolworths Targeted Media, states, ‘We help brands target up to 10.5 million shoppers representing more than 80% of Australian households one-on-one, with the right message, in the right place and at the right time direct or via their preferred channel.’ Data brokers—whose business model depends on the resale of data about individual eyeballs online—are key in this industry, and I discuss them in more detail below.

Another alternative source of data is credit reporting agencies such as Dun & Bradstreet, which offers up to 600 data attributes about each potential customer. Other traditional credit reporting agencies such as Experian also have advertising businesses where they help their clients identify

---

32 Kaye, Kate, ‘Mastercard, Amex Quietly Feed Data to Advertisers,’ *AdAge*, 16 April 2013, available at: https://adage.com/article/dataworks/mastercard-amex-feed-data-marketers/240800/. Credit card data is usually used to create data segments of ‘electronic intenders’ which is then sold back to data brokers.


36 Qantas sells profiles for targeted marketing saying ‘We provide de-identified data to help with targeted marketing (e.g., 30-year-old males from Sydney who regularly travel domestically and therefore might be interested in car rental or hotel offers).’ Fernyhough, James, ‘Big banks earn cash from selling data about your spending,’ *The New Daily*, 6 April 2018, available at: https://thenewdaily.com.au/money/finance-news/2018/04/06/big-banks-sell-data-customer-spending-habits.

37 ‘FMCG / CPG,’ *Quantium*, available at: https://www.quantium.com/fmcg-cpg. For more on the Woolworth’s data source, see also ‘WOW Personalisation,’ *Quantium*, *YouTube*, 3 November 2016, available at: https://youtu.be/HPi2-Q6oZUQ, which explains that it is equivalent to asking 70 million questions each week.

38 ‘D&B DataVision,’ *Dun & Bradstreet*, available at: https://www.dnb.com/content/dam/english/dnb-solutions/sales-and-marketing/dnb-datavision-brochure-2017.pdf. Dun & Bradstreet has built up a database of 250 million businesses in 220 countries. They also sell their data through a variety of data brokers. See also ‘Q&A with Dun & Bradstreet on Data Quality and Fraud,’ *Lotame*, 27 February 2017, available at: https://www.lotame.com/qa-with-dun-bradstreet-on-data-quality-and-fraud. See also, Liyakasa, Kelly, ‘Dun & Bradstreet aims to be the de facto B2B data shop,’ *Ad Exchanger*, 30 June 2016, available at: https://adexchanger.com/ad-exchange-news/dun-bradstreet-aims-de-facto-b2b-data-shop. Though their focus has been on B2B they have also discussed how their data can be used for B2C marketing—for example, identifying individuals who work in firms that specialise in Audio Visual services who may be interested in cameras.
new audience segments through their shopping data.\textsuperscript{39} For example, Experian’s ConsumerView database covers more than 14 million Australian adults, 10 million residential households, and 51 million linkage records.\textsuperscript{40}

In recent years, many media companies have agreed to share audience data with each other so that they can offer cross-platform audiences to marketers. In 2017, two of France’s largest publishing companies, Le Monde and Le Figaro, formed the Skyline alliance.\textsuperscript{41} Another French publishing alliance called Gravity pools data from 15 publishers that together reached approximately 44% of the French population every day in the summer of 2017.\textsuperscript{42} These alliances can also be multinational. One prominent example is EBX (European Broadcaster Exchange), a partnership between Germany’s ProSiebenSat.1, the UK’s Channel 4, France’s TF1, and Italy and Spain’s Mediaset.\textsuperscript{43}

Such alliances are popular in Australia, too. In 2015, Fairfax Media and Nine formed a premium publishing alliance and mobile advertising marketplace called Australian Premium Advertising Exchange (APEX).\textsuperscript{44} The alliance offered advertisers real-time bidding across over 120 mobile sites and apps. Prior to the announcement of the Fairfax-Nine merger in 2018, Fairfax, Nine, and News Corp had agreed to explore the possibility of a partnership that would allow advertisers to target specific users as they moved between the companies’ various platforms.\textsuperscript{45}

\begin{itemize}
\item \textsuperscript{39} In the UK, Experian emphasises its reach saying, ‘By linking our ConsumerView database of 49m individuals and 25m households to client or publisher 1st party data, we can reach 99% of the UK’s targetable population.’ ‘Experian Marketing Services,’ \textit{Experian}, 2016, pp. 1-40, at 33, available at: https://www.experian.co.uk/assets/marketing-services/brochures/experian-marketing-services-brochure.pdf.
\item \textsuperscript{41} Davies, Jessica, ‘French publishers are joining forces to take on Google and Facebook,’ \textit{Digiday}, 10 July 2017, available at: https://digiday.com/media/french-publishers-joining-forces-take-google-facebook.
\item \textsuperscript{42} Davies, Jessica, ‘French publishers are joining forces to take on Google and Facebook,’ \textit{Digiday}, 10 July 2017, available at: https://digiday.com/media/french-publishers-joining-forces-take-google-facebook.
\item \textsuperscript{43} Pidgeon, David, ‘EBX ready to execute first campaigns,’ \textit{Mediatel}, 20 March 2018, available at: https://mediatel.co.uk/newsline/2018/03/20/ebx-ready-to-execute-first-campaigns.
\end{itemize}
3.2 Are the Benefits Conferred by Big Data Imitable?

If a competitor can easily imitate the benefits conferred by a ‘resource’ such as big data, then it is unlikely to be a source of sustainable competitive advantage. The major benefits of having access to data for a firm selling products or services in the digital environment is that the data can be used for targeting and measurement. Therefore, to understand whether a large digital platform having data is likely to confer enduring competitive advantage, it is useful to try to understand what alternative exists for firms in terms of targeting, personalisation and measurement.

If we think of the key benefits of data in digital advertising markets—which are providing measurability, personalisation and targeting—then these benefits are broadly available from many different datasets.

3.2.1 Alternative Targeting Options

In Section 3.1, I discussed the variety of data sources that might be available to an advertiser. As I just illustrated, this means that a firm, for example, selling cameras has several alternative sources of data that they can use to pinpoint potential customers. They could start off with their own customer lists, and then use one or more of a variety of these data services either to better track and target these customers, or to identify other consumers who are either in the market for cameras or have a similar behavioural profile to their existing customers. Demand-side platforms and data management platforms can integrate the advertiser’s customer lists with their own data on consumer targeting when submitting bids on supply-side platforms. Though this can be a somewhat complex process technically, the key is that these platforms work together seamlessly and in real time to identify when the advertiser should bid for a pair of eyeballs, and what ad the advertiser should show them. A wide variety of data can be used to potentially establish whether someone should be shown, for example, a camera ad. This data spans interest by people browsing a camera website, or intent by checking out prices of a camera on Amazon, or even past purchase of a camera (as recorded in loyalty program transactional data). My purchase intent for cameras will be clear to whichever ecommerce sites I browse while looking for a camera, such as Amazon, dicksmith.com.au, jbhifi.com.au, digitalecamerawarehouse.com.au, or camerahouse.com.au. Amazon, for example, collects information on me whenever I search for a
product, place an order, submit a review, compile a Wishlist, etc. Amazon also has the data I supplied when setting up my account, such as my name, address, and phone number.

As well as data-brokers, who I discuss at length above, another source of online purchase-intent data is through social media platforms and search engines. Social media websites like Twitter, Pinterest, or Facebook will know if I post that I am excited about buying a new camera or even a specific camera. Search engines such as Google or Bing have access to my search history, which might also reveal my intent to buy a camera, for example, if I use their platform to search for a ranking of the best cameras, a review of a specific camera model, or a site where I can actually buy a camera.

However, these different types of data can all be used to achieve the same goal of potentially selecting the right set of eyeballs to show a camera ad to. Though these eyeballs may be closer to purchase or further from purchase, there is still the potential for advertising to have a positive effect.

To understand how easy it is for an advertiser to use targeting technology, it is worthwhile to consider the example of Pinterest. This website allows advertisers to target consumers by precise locations, interests, and keyword searches. Figures 8 to 10 show screenshots from Pinterest that illustrate how an advertiser might target potential camera buyers using Pinterest’s data. For example, using Pinterest’s data on user location, which might be provided by the user when creating an account, an advertiser can target specific regions of Australia, such as ‘Melbourne - Outer East’ and ‘Melbourne - South East’. Moreover, an advertiser can leverage Pinterest’s data on customers’ demonstrated interests (perhaps gathered from browsing history) to target potential customers that have demonstrated an interest in ‘[c]ameras and [a]ccessories’ or have searched for keywords related to cameras, such as ‘polaroid camera.’ Like many other sites, Pinterest also allows advertisers to import and use their own data, such as customer lists, to create and modify target audiences.

46 ‘What data does Amazon collect and use?’ Amazon, available at: https://www.amazon.co.uk/gp/help/customer/display.html?nodeId=G6RZ4RMNMLUQRLY2.

47 ‘What data does Amazon collect and use?’ Amazon, available at: https://www.amazon.co.uk/gp/help/customer/display.html?nodeId=G6RZ4RMNMLUQRLY2.

There are many other websites with similar information. For example, Figures 11 to 12 show how this can be done with Twitter. The first screenshot shows how an advertiser might use Twitter’s data on users’ demonstrated interest in ‘[c]ameras and camcorders’ to target people that may be more receptive to an advertisement for a new camera. The second screenshot shows an advertiser’s ability to use ‘lookalike’ targeting—that is, for example, to target someone who is the kind of person who has followed or engaged with Leica on Twitter.

**Figure 8**

Pinterest: Customisable Customer Attributes Include Gender, Location, and Language

Figure 9

Pinterest: Pick Topics Related to Your Brand to Reach Your Target Audience

Add interests
Pick topics related to your brand so you can reach people based on other Pins they engage with.
Learn more

Figure 10
Pinterest: Extend Your Reach by Selecting Relevant Keywords

Figure 11
Twitter: Targeting Consumers by Interest in Cameras

Retailers also create and have access to data that can be used for targeted advertising. For example, Woolworths’ Media Hub states to prospective advertisers that it is able ‘to talk to Woolworths customers across a variety of media touch points including TV, Radio, Press, Point of Sale, Online, Social Media, Fresh Magazine, Marketing Activations, Category Events, In store
radio and Baby & Toddler Club’. Moreover, Woolworths works with advertisers throughout the life cycle of a marketing campaign, including ‘planning, development, implementation, [and] analysis.’

Other advertisers with first-party data have developed or acquired data analytics skills to better understand potential consumers and implement targeted marketing strategies. For example, the insurance company IAG recently acquired Ambiata, a data analytics firm specialising in artificial intelligence and machine learning, to better leverage their own data for both marketing and pricing strategies.

In 2015, Australia’s largest airline Qantas leveraged its own Frequent Flyer program data to build an audience data management and media buying business called Red Planet. The business uses this data and data from other sources, including a 100,000-person research panel, to target ads for Qantas Group and its clients.

An advertiser can use different types of data on different media to achieve similar goals. A good way to see this is to consider what data and insights an advertiser could have through access to a data management platform. Figure 13 below is a screenshot from Experian, which offers a data management program in addition to being a data broker. When discussing these data, they state:

‘[T]he media buying aspect is just the first benefit that data-informed decision making can deliver. The insights a DMP provides can be actioned across channels – driving efficiencies in not just display advertising but also in social media advertising, video on demand and - at some point in the near future - Addressable TV. And this is not restricted to just paid media – insights can as easily be used in other execution channels, opening the door to more accurate email,

---


organic social, SMS campaigns and of course how you manage customer journeys on your own website.'

**Figure 13**

A Typical Data Management Platform Used by an Advertiser


Figure 13 shows how Experian can connect a user’s profile across platforms and channels through a wide range of data sources to provide more valuable insights to advertisers. The increase in data availability also has the potential to change how ads are delivered in other more traditional channels. Competition from digital providers in social media have challenged existing incumbents and forced them to innovate to the benefit of consumers. For example, firms are developing technologies in Australia that allow TV advertisers to buy targeted TV ads—that is, establishing
fine-grained micro-segments and showing relevant ads to that segment online. In Australia, Nine Network offers marketers access to 6.5 million customer IDs through direct on-demand video advertising segmented by age, sex, and location. The company has the broadest reach of on-demand broadcasting in the country, accounting for 2.6 million viewers averaging 2.25 hours spent watching per day, according to Nielsen. Similarly, an Australian radio station recently announced the launch of one-to-one personalised ads.

3.2.2 Means of Tracking or Measuring Advertising Success

Though there has been much discussion of the role of data in transforming digital advertising, there has been a parallel shift in industry capacity in terms of the tools available to participants to allow them to measure the effectiveness of any one advertising channel. This is important because advertisers use a variety of media channels when reaching a customer. The advent of tools that quickly steer advertising dollars away from less effective channels has itself changed the competitive dynamics of this industry and rendering ad-performance and return of investment the most important consideration when advertisers are deciding how to allocate advertising dollars.

As Mark Ritson, adjunct professor of marketing at Melbourne Business School, explained: ‘The more channels a campaign uses, the more effective it appears to become. We keep setting media against each other when, in truth, the right approach says yes to both.’ For example, a recent successful campaign from the Australian Defence Force used ten distinct platforms,

---


including TV, online video, social, cinema, out-of-home, digital display, mobile, search, Spotify, and radio.\textsuperscript{58}

The development of cross-channel attribution software allows advertisers to allocate their spend across these options. Cross-channel attribution software facilitates this comparison process by enabling advertisers to reallocate their advertising dollars to the channels that are offering the highest return on investment in real time (whether it be an offline channel, display or search). Today, there are many competitive software options for cross-channel marketing campaigns, including Adobe Campaign, IBM Watson Marketing, Oracle Marketing Cloud, Salesforce Marketing Cloud, Conversion Logic, and SAS Customer Intelligence (all of which are available to Australian companies).\textsuperscript{59} Importantly, the success and proliferation of these platforms suggest that advertisers are actively employing these products to measure and optimise ROI, revealing that these tools facilitate advertisers’ switching between channels.

Figure 14 presents an example of cross-channel analytics available through Conversion Logic, Figure 15 presents an example from SAS Customer Intelligence, and Figure 16 presents an illustration of Nielsen’s cross-channel campaign management tool. In a Q1 2018 report from Forrester, 40% of the 32 marketers surveyed reported using their cross-channel marketing software to manage at least 14 different digital and offline channels.\textsuperscript{60}

\begin{flushleft}
\footnotesize


\end{flushleft}
Figure 14

An Example of Cross Channel Attribution Software Offered by Conversion Logic

![Image](image)

Figure 15

SAS Customer Intelligence: Cross-Channel Campaign Attribution

Figure 16

Nielsen’s Media Impact Solution

THE FIRST TOTAL AUDIENCE VIEW OF CONSUMER MEDIA BEHAVIOR FOR PLANNING

Total Media Fusion incorporates TV, smartphone, tablet, computer, video streaming, magazine, cinema and digital place-based media. It’s the fuel for our planning system, designed specifically for media planning and analytics.

Many new forms of technologies have evolved that allow for better measurement of marketing performance across multiple different channels. An example of this is Datorama (recently purchased by Salesforce) which uses AI to allow firms to track cross-channel marketing performance with many different key performance indicators (KPIs) designed to measure campaign metrics like volume, efficiency, pacing, effectiveness, impact, and brand health.\(^{61,62}\) Datorama describes its service as follows:

‘Your sites, SEO strategy and social pages form a critical set of connection points for customers to find you, interact, purchase, and share their voice. With Datorama you don’t have to look at your organic marketing as a silo, separate from your advertising. With Datorama, you can connect all of your marketing holistically– so you can see how your paid and organic programs interact, boost your best content, optimize for quality traffic and tie your dwell times and buzz to the business.’

Figure 17 shows some of the digital marketing channels available on Datorama, including Adobe Analytics, Bing Ads, Facebook Ads, Google Ads, LinkedIn Ads, Mail Chimp, and Pinterest Ads. Figure 18 shows Datorama’s cross-channel ROI trend and forecast calculation, along with other relevant KPIs for marketers.


Figure 17

Datorama: Select Campaign Channels Using Datorama

Figure 18

Datorama: Track Spending and KPI Performance Across Multiple Channels

With so many different media channels, tracking consumers across different media is also a significant challenge. Two firms that offer technology to address this challenge are ‘Epsilon’ and ‘TapAd’. Epsilon is also a data broker; it describes its process as follows:63

‘First, we combine your customer data with Epsilon’s unparalleled data hygiene to enable the highest quality identity matching. Then we augment your known customers with online data and behavioural activity on your properties across all their devices, giving you a more complete picture of how they are interacting with your brand. This process enables immediate and persistent customer identification across devices and channels better than ever before.

Founded in 2010, TapAd provides cross-device marketing technology, analysing data points across multiple screens to identify consumer activity online, then selling that information to media buyers and other advertisers in order to create better targeted advertising.64 TapAd describes its product, the ‘TapAd Graph,’ as ‘enabling marketers to capture a wealth of consumer touch points across devices and channels, resolving them back to an individual.’65 In 2018, TapAd launched a Customer Data Platform (CDP), which is a self-service platform that lets marketers upload their own first-party data and match customers to TapAd’s third-party data on devices.66

More traditional players have also started offering cross-channel management solutions. For example, Nielsen now provides marketers the ability to track ad campaigns across television and online platforms.67

3.3 Is Big Data Valuable?

One key question of course with a resource is the extent to which it is valuable. As I discuss above, digital data is valuable in advertising not only as a means of ensuring the right eyeballs see the right ad at the right time, but also as a means of ensuring that consumers don’t see ads they find uninteresting and advertisers can measure the success of their ads. In other words, it is evident that

64 Lunden, Ingrid, ‘Telenor Jumps Into Ad Tech, Acquires Tapad For $360M,’ Techcrunch, available at: https://techcrunch.com/2016/02/01/telenor-jumps-into-ad-tech-acquires-tapad-for-360m/. TapAd was acquired by Telenor, the Norwegian telecommunications company in February 2016.
in this ecosystem that data is valuable if analysed and applied in a productive way. Perhaps the more useful question from a strategic point of view though is whether big data by itself is always going to be valuable or whether it requires complementary assets to ensure that it is valuable.

This is something I studied in recent research into the data broker industry. In this research, we examined data brokers’ ability to correctly identify whether or not a set of eyeballs were female based on browsing behaviour. We found surprising variation in their ability to do so, with success rates ranging from 27.5% to 62.7%. Furthermore, this ability was not related to the amount of data eyeballs they had access to—the data broker with the data about the most eyeballs (nearly three times as much as its closest competitor) only had a success rate of 42.4%.

This research highlights that by itself big data is not inherently valuable. What is valuable is the ability to make the right inferences based on the data that a firm has access to. This is often a matter of deploying the right algorithms as well as an understanding about what may lead to errors in classification.

### 3.4 Is Big Data Non-Substitutable? Is Big Data a Sufficient or Necessary Condition for Success in Digital Advertising Markets?

To analyse whether big data is a sufficient or necessary condition to ensure success in digital markets, I proceed as follows. First, I consider whether there are examples where firms have possessed a lot of data but not succeeded—if such examples exist, it is evidence against the idea that big data is a sufficient condition for success. Second, I consider whether there are examples of firms that have grown successful in digital markets despite not having access to big data—if such examples exist, it is evidence against the idea that big data is a necessary condition for success.

#### 3.4.1 Is Big Data Sufficient?

The first set of examples where big data turned out not to be sufficient are in the social networking space. As I have discussed in some of my other work, possessing data about a large...
number of users has been repeatedly insufficient to overcome the great fragility exhibited by social networking websites.\textsuperscript{69}

Friendster was founded in 2002 and at one point was the largest social networking website in the United States, reaching 20 million users.\textsuperscript{70} As such, it had access to 20 million U.S. user profiles and data about their connections. However, this proved neither sufficient as a basis for building a business model nor sufficient for retaining users. Instead, the fact that Friendster’s user base knew each other offline made it easy for them to coordinate and leave when superior alternative offerings were found. Indeed, the underlying network effects which led the website to grow so quickly proved a destabilising influence when users experienced their friends leaving.\textsuperscript{71}

Myspace largely replaced Friendster because Myspace offered better personalisation options and the potential to hear new music.\textsuperscript{72} Myspace also pursued a strategy of enticing ‘influencers’ who were on Friendster to migrate to its network.\textsuperscript{73} Myspace itself was the most visited social networking website between 2005 and April 2008.\textsuperscript{74} However, Myspace’s attempt to be too many things to too many people, its intrusive advertising such as the infamous ‘punch the monkey ad’ and the failure to innovate new features led Myspace to lose users to Facebook.\textsuperscript{75}

Outside of social networking, there are other examples of digital firms who appeared to have access to large amounts of monetisable user data that ended up failing in the digital advertising space. Flickr, the image and video hosting service acquired by Yahoo! in 2005 reported having 90


However, when Yahoo! attempted to convert this user data into the foundation of an advertising-based business model, it failed. This is because Flickr had previously been a service that was free for most (and free from ads) while a few people paid for a subscription. Data by itself was not sufficient to convince users to accept the presence of ads by their photos, highlighting that one of the big barriers in advertising markets is not data—instead, it is user acceptance of advertising and their willingness to accept their attention being monetised. As I have documented in my own research, these natural constraints push advertisers to insert smaller and less intrusive ads than they may have done otherwise.

Yahoo! made other acquisitions at this time that in some sense offered data that were even more promising from an advertising perspective. del.icio.us was a social bookmarking site which initiated the idea of users tagging content to aid sharing. From an advertising perspective, such data is remarkably valuable because to tag something and share it with one’s social network is a useful indicator of ‘intent’ or interest towards that content. However, as a service it stumbled because it was easy for users to transition to other services, which were perhaps more convenient as they were linked to a browser. Its lack of stickiness—something I discuss below as a hallmark of many cloud-based services—meant that its large number of users and the usefulness of the associated data was not sufficient to build a successful advertising platform.

These examples suggest that firms need more than ‘data’ to succeed in advertising. Rather, firms need to develop advantages in letting advertisers directly target their consumers, to retain users and their attention, and to present advertising unobtrusively. None of these advantages are related to the size of the user base or the volume of data collected. If another platform is better at retaining users’ attention (either with better content or by making the ads themselves less distracting), one would expect users to switch away from a competitor to enjoy the better experience, and for advertisers to follow.

---


For Facebook and other social network providers, much attention and the data generated through user use of the platform is not monetisable. For example, take the ritual of posting photos of kids on the first day of their school year. Such posts are popular among users but there is little that is monetisable about this practice. When users are focusing on photos of their friends’ kids they are not clicking on ads and don’t particularly wish to see ads. People posting these photos are not in the market for a specific type of goods—they have already bought their school supplies. And there is little new information about their long-term purchasing habits, given that the fact they have kids is knowable from a long history of web browsing and kid-related purchases across the entire internet.

Facebook and other digital advertisers face the trade-off between investing in maintaining the relevance of the platform with features and content of interest to users, while ensuring there is enough advertising inventory to monetise successfully and support those investments. This trade-off characterises the challenges that Yahoo! had with shifting Flickr’s business model in 2012. They had a large user base, and even offered more storage space for free to attract more users and expose them to advertisements. But these capabilities were not enough. Flickr’s users may have found additional storage for free attractive, but it did not encourage them to spend time and attention on Flickr’s website looking at advertisements.

In fact, the failure of the Flickr acquisition by Yahoo! illustrates the perils of assuming that merely having users and data about those users is sufficient to build a successful advertising model. When Yahoo! acquired Flickr in 2004, Flickr was recognised for having strong social connection and search tools. Yahoo! sought to exploit Flickr’s data associated with users and their photographs—tags, labels, and categorisations assigned by users. Yahoo! executives focused on monetising that data, but failed to build connections among users, or add new features that would keep them involved with Flickr, generating and updating new data and delivering more attention.80 Users shifted their time and their data to platforms like Facebook, YouTube, Instagram, and Dropbox.81

---


Despite also having strong capabilities to interact on mobile devices, Flickr also failed to adapt to the era of ‘apps’ on smart phones like iPhones and Google Android devices. Despite the introduction of the iPhone in 2007, Flickr did not introduce a mobile app until 2009, and it was generally derided by users for being of poor quality. Instagram’s success, and Flickr’s failure, is not related to the number of users and amount of data they captured from photos, but to Instagram’s innovation surrounding its user interface and innovative filters that kept users connected and interested.

Furthermore, the advantages of data are not persistent. The large amount of data that Flickr acquired from millions of users over many years is no longer of much value to advertisers. Flickr was recently purchased by SmugMug, an older photo sharing service that largely focused on the value of its current users’ willingness to pay for photo storage and related features and eschewed online advertising.

3.4.2 Is Big Data Necessary?

Each example of a larger platform failing despite having more user data than its much smaller competitors is also an example of a smaller platform succeeding without access to big data. In a similar manner to the offline world, what determines success online is a superior ability to understand and meet customer needs. The number of changes that have taken place in the relatively short history of digital advertising platforms have shown repeatedly that smaller platforms can first offer a better value proposition to consumers and they can subsequently acquire big data as a natural consequence of their success.

In the last section, I discussed how Myspace was able to supplant Friendster as the largest social networking site in 2005 and Facebook grew significantly from 2008 as Myspace lost users. In each of these cases, a company with less user data managed to win users’ attention. More recently, Snapchat has been very successful in competing in this space without access to big data because it recognised that people wanted to share personal information more privately. In 2016, Snapchat


surpassed Twitter in number of daily active users.\textsuperscript{84} That same year, four years after its launch, Snapchat began running advertisements.\textsuperscript{85}

Similarly, possession of a large amount of data does not explain the rise of Australia’s Travello, a travel social network application. Rather, the platform’s success can be attributed to the fact that it has managed to attract users with the ability to connect them with other travellers and locate points of interest. It now allows companies to geo-target users to display their ads and even add their locations to users’ maps.\textsuperscript{86} Liven, another Australian start-up in the digital advertising space that did not need vast amounts of user data to get started, charges restaurants to be featured on their app and pays users to dine at participating restaurants. The company closed a 10 million AUD funding round in 2017.\textsuperscript{87}

These new platforms may acquire user data in the course of developing their user base and providing their services, and this data may allow them to target advertisements more effectively, offering value to advertisers. However, this data was not a necessary precondition of their successful entry, but rather an outcome of their operations.

4 Does Big Data in Digital Platforms Lead to Switching Costs or Network Effects?

In its Digital Platform Inquiry, the ACCC posits that ‘data-driven network effects and economies of scale may give established digital platforms that possess large amounts of data an enduring competitive advantage’ (ACCC, 2018). The relevant question here is whether big data can augment or be the root cause of either substantial network effects or switching costs.

4.1 Network Effects, Big Data and Online Advertising

Network effects occur when a service becomes more useful as others use it. This is related to potential competition concerns because it suggests an iterative feedback loop which may guarantee the entrenchment of large firms if markets are characterised by network effects.


\textsuperscript{85} Vincent, James, ‘Snapchat will start showing ads between your friends’ stories,’ \textit{The Verge}, 14 June 2016, available at: https://www.theverge.com/2016/6/14/11930386/snapchat-ads-api-stories.


4.1.1 Relevance, Reach and Network Effects

To understand the potential for network effects in digital advertising markets, it is useful to introduce some advertising terminology. In advertising, reach and relevance are often two stated goals. Reach refers to the ability to target a large proportion of eyeballs. Historically, reach has mattered because advertisers prefer to reach more new sets of eyeballs rather than showing the same ad multiple times to the same eyeballs.\textsuperscript{88} Relevance refers to the ability to target the right eyeballs of those that are more likely to be influenced by an ad.

For example, suppose a new online gourmet pet food company called YumTreats is trying to acquire new customers for its premium dog food product line. When selecting a platform to place advertisements, YumTreats is unlikely to care about how many eyeballs an advertising platform has in total (\textit{i.e.}, a platform’s potential to reach a consumer). Instead YumTreats cares about whether it can isolate people who own dogs and are willing to pay for expensive dog food, and potentially whether there is enough data to personalise its ad content to adjust for the different dog-food demands of different dog breeds (\textit{i.e.}, a platform’s ability to target a relevant consumer).

In the past, while reach could be achieved through mass-media channels such as television advertising, relevance was harder to achieve because there was little data in an analogue world, meaning that attempts at advertising were often scattershot.

As a goal, reach is heavily related to network effects because if an advertiser wants reach then that suggests it desires many users to join its platform.

However, in the digital world, reach can now be achieved in a variety of ways, undermining the potential for network effects in advertising markets. This is partly because it is possible to use digital techniques such as ‘frequency capping’ to ensure that no set of eyeballs sees an ad too many times.\textsuperscript{89} Moreover, it is possible to use software to ensure that this frequency capping software is used across multiple ad platforms. Continuing with our example, YumTreats can utilise Sizmek to manage frequency capping across a marketing campaign that runs across multiple demand-side


platforms. This is important because it means that if YumTreats values reach as a way of avoiding showing the same pair of eyeballs a dog food ad multiple times, then the availability of these new digital technologies means that reach is no longer a potential explanation of why YumTreats might value the size and audience of an advertising platform.

The link between relevance and platform size is less important in digital platform markets due to the advent of automated digital advertising. In the past, YumTreats might have concluded that if it wanted to target dog owners, then it might make sense to go to Yahoo!, Microsoft, or AOL and advertise next to their content about dogs. This is because relevance as a goal was still dependent on reach—and therefore potentially subject to network effects. In other words, YumTreats’ only way of finding enough dog owners to advertise to was to go to a website with many users and then use content on that website to determine that these people were dog owners.

However, the automated digital advertising system has changed relevance’s dependence on reach. Now it is possible to use demand-side platforms and data broker services to identify a pair of eyeballs that might be interested in dog food on any website. This means that even the smallest of websites can benefit from this ability to aggregate data about an individual and use that to target ads even on completely unrelated websites.

4.1.2 Economics of Scale and Network Effects

An additional argument that has appeared in the literature regarding network effects in advertising markets is the argument that there can be ‘data-driven’ network effects. This refers to the idea that if a firm has data then there can be a positive feedback loop where it uses this data to improve its operations and in doing so, because it offers a better service, it will attract more customers and more data. The issue with this definition is that it really refers to a classical economics idea of economies of ‘scope’ or ‘scale’, not network effects. There is no reason to think such a process will be limited to ad platforms. For example, a meat pie shop could use its data on purchase decisions to understand better customers’ preferences, and use this to refine and hone its
pie recipes and offerings, and this in turn use this superior offering to attract more customers and more data. This is commonly thought of as being an economy of scope or scale. Similarly, it would be misleading to suggest that the data collected from the sale of billions of hamburgers by McDonald's have provided the fast-food chain a data-driven network effect that affects Hungry Jack’s ability to compete.

One recent study explored explicitly the question of whether there were economies of scope and scale in the use of data for forecasting. Researchers used Amazon data to explore whether having additional data from multiple products helped improve forecasting product demand. They found that though additional data about a single product helped Amazon's ability to forecast demand (with diminishing returns to scale), there were no gains from more data from other products for the purposes of forecasting.

4.2 Switching Costs in Digital Ad Platforms

Switching costs occur when customers find it costly to switch between services of platforms. The idea is commonly referred to as capturing the ‘stickiness’ of a platform.

Real-time dashboards and other cross-channel attribution technologies allow advertisers to invest their advertising dollars instantaneously into the campaign that is offering the largest ROI. In effect, the advent of these technologies has lowered switching costs. This kind of switching is something that Facebook itself has experienced—when advertisers discovered that costs for showing ad impressions were getting higher, they transferred their advertising dollars away from the platform.

4.2.1 When Are There High or Low Switching Costs in Data-Driven Industries?

The question is then whether data itself can lead to switching costs. In general, I identify two drivers of switching costs in data-driven markets in my teaching:

(1) The ease with which data can be transferred between platforms

---


94 As discussed above in section 3.2.2, there are a plethora of real-time dashboard and cross-channel attribution technologies available to Australian companies and it would be reasonable to assume advertisers in Australia make use of them to switch their advertising spend to the channel delivering the largest ROI.

(2) The need to transfer data or the extent to which historical data is valuable.

For example, imagine a platform that helped a hospital manage digital patient histories. The hospital needs to have access to these historic data to provide the right care. However, it is difficult for users to transfer and export these data. This is because of a variety of privacy regulations that are intended to protect this data but also make it difficult to transfer. This makes it challenging for hospitals to switch to new providers of health records systems, because in doing so they have to reobtain patients’ permissions to export the data. In such cases, where there are high costs to transferring historical data bases and historical data bases are valuable, we might expect switching costs and little switching across platforms.

By contrast, large amounts of historical advertising data does not appear valuable to advertisers or ad platforms. In Chiou and Tucker (2017), we examine what happened to search engine accuracy when the search engines changed the length of time they retained data as a response to requirements in the European Union. We found no effect on search engine accuracy as measured by whether a consumer felt the need to repeat the search. We discuss a potential explanation, which is that search engines operate a ‘long tail’ business where many search queries are actually unique. As a result, more data on what consumers did in the past is not useful. Instead, what is useful is smart predictive algorithms that help interpret what a user might mean or want by their search. Those algorithms require investments and innovation by firms and are the basis for effective competition by platforms to attract advertisers; put differently, a platform that builds its models on more old data will not do better than an ad platform that builds better models on existing data.

In other research, I also emphasise the limited usefulness of historical data in advertising. One useful illustration of this is a type of display advertising called ‘retargeting’ or ‘remarketing’. This is a highly profitable form of advertising, where after a user has browsed a product—for example, a pair of shoes—these shoes (and alternatives) are featured in ads shown to the user. However, my research suggests that usefulness of such browsing data is short lived. For example, imagine someone who is searching for a bunch of flowers. Typically, they will buy those flowers within a few days, if not a few hours meaning that the value of their prior product browsing data as a signal of intent is short-lived, and long histories of historical advertising data is less useful.


Furthermore, the cloud-based nature of advertising data facilitates the transfer of data across platforms. Indeed, the key functionality of data-management platforms is to allow advertisers to combine and store data about potential customers from a variety of sources such as the firm’s own website and third-party data vendors. For example, Adobe’s Audience Manager states that ‘marketers can combine data from mass media activities, direct marketing, and social media conversations to reveal individuals and audience segments who can be targeted for specific messaging. Audience Manager allows you to monitor anonymous customer data from channels such as ad display, search, and video.‘

Tools also make it easy to switch campaign data (such as the specifications of who is targeted in a campaign and its creative) across platforms. Bing, for example, makes it easy to import Google Search Ad settings into Bing with a couple of clicks. Yahoo! allows advertisers to easily import Native Ad settings from their Google accounts with a few clicks. Facebook allows the export of all campaign settings in a csv file.

A natural corresponding question is whether there are switching costs on the consumer side of digital platforms. Again, many digital tools exist to make it easier for consumers to switch platforms. For example, Apple provides a downloadable app on the Google Play store which potential iPhone buyers can use to seamlessly transfer their data and settings from an Android phone.

Like advertisers, consumers can also ‘multi-home’ across apps or websites. A classic example of multi-homing is the Uber, Taxify, Ola and 13Cabs apps. Many users have multiple apps on their phone and choose which platform to use each time based on the expected cost and wait. In attention markets such as the market for digital content, the main constraint on multi-homing is how much attention consumers have to spare, as there are no physical barriers to switching apps or switching

102 ‘Switch from Android to iPhone,’ Apple, available at: https://www.apple.com/uk/switch.
websites. Recent theoretical work has suggested that multi-homing in social networks\textsuperscript{103} has positive implications for consumer welfare where platforms can differentiate their offerings, such as with Facebook (for maintaining social and news communications), and LinkedIn (for professional interactions).\textsuperscript{104}

4.2.2 The Relationship between Low Switching Costs and Innovation in Digital Advertising

The fact that cross-channel attribution tools and the general use of aggregate dashboards to facilitate cross-channel comparison reduce switching costs across different advertising platforms has implications for innovation in advertising markets. Digital advertising platforms need to constantly invest to attract and retain advertisers to their platforms because, first, advertisers face low switching costs and, second, instantaneous feedback on campaign performance provides an easy tool to assess if they should leave an advertising platform. This intense competition to invest in innovation, which increases advertising’s return on investment, benefits both audiences and advertisers because consumers receive more relevant ads and advertisers show their ads to audiences who are more likely to find the information useful.

The increasing popularity of self-serve platforms has further reduced switching costs and increased competition. Nine recently announced a self-serve trading platform called 9Voyager to allow Australia’s small to medium size enterprises (SMEs) to buy video ads on its various platforms.\textsuperscript{105} Self-serve advertising platforms such as Epom Markets, SmartyAds, and DanAds also offer advertisers access to low cost ways to implement and evaluate their campaign strategies.\textsuperscript{106}

The fact that advertisers shift their dollars to whatever platform offers the most return on investment can be seen in the mobile advertising space, which has seen a growth of over 20,000%\textsuperscript{103}


\textsuperscript{105} Mason, Max, ‘Nine fights tech giants for slice of SME pie,’ The Australian Financial Review, 18 October 2018.

in the last decade.\textsuperscript{107} The increasing use of mobile devices and the expansion of apps in both the Google Android and Apple iOS environments has led to rapid growth in ‘in-app’ advertising, which would include advertising on any of the more than 4.8 million apps currently in circulation.\textsuperscript{108} This reflects the migration of users to mobile devices—a recent report found that the average Australian has just under 100 apps installed on their phone, and on average uses 36 apps per month.\textsuperscript{109}

5 Concluding Thoughts

Advertising is a dynamic industry that has been at the forefront of the digital revolution. This dynamism is particularly striking, because traditionally advertising has been the most unmeasurable and imprecise part of marketing strategy. However, advertising has transformed itself into an industry which offers advertisers considerable advantages in terms of the ability to show the right ad to the right consumer at the right time, and then to measure the effectiveness of that ad.

This report was written in response to the ACCC’s Digital Platforms Inquiry which is, among other things, exploring whether digital platform providers are deriving ‘enduring competitive advantage’ from their access to data. I use a framework named a ‘resource-based view of the firm’ from the strategy literature, to consider whether big data is a potential source of competitive advantage in the digital advertising industry. Specifically, I ask whether big data is rare, inimitable, valuable and non-substitutable. For each of these criteria, I describe why there are reasons to doubt that data itself is satisfying any of these key criteria for being a source of enduring competitive advantage or market power for digital advertising platforms.

Recent improvements in targeting have allowed advertisers to more effectively identify and reach potential consumers at lower cost. Advertisers also have access to different types of online and offline data that can be used together with cross-channel management tools. These developments have led to reduced switching costs as advertisers can more easily allocate spending across channels and platforms in response to changes in relative returns of investment for that


particular advertising channel. This reduction in switching costs has also destabilised the potential for network effects to develop in the digital advertising industry.

Looking forward, I would expect that not only is the technology of ad delivery likely to change, but the data used to target ads is also likely to change. For example, the advent of digital assistants, devices placed within the home to aid the automation of home life, is likely to provide new sources of data and new ways of targeting ads. Outside the home, as automobiles become increasingly digital and connected, our use of vehicles is also likely to generate new sources of data and targeting opportunities. Similarly, though ‘wearable’ devices have (as yet) largely failed to find a mainstream place in the marketplace for consumer electronics, they may in the future replace the current mobile ecosystem as a means of delivering ads.

This report has largely focused on the question of whether at the current time, data offers a source of sustainable competitive advantage in the online advertising space. In the long run, the key question for all players currently in this space is whether they will be able to keep pace with such shifts in consumer use of technology, and also understand consumer needs sufficiently to continually improve the process of providing relevant ads to consumers.
References

Academic Publications


Other Publicly Available Documents


‘Cross-channel marketing unifies your brand voice,’ *Adobe*, available at: https://www.adobe.com/experience-cloud/articles/cross-channel-marketing.html.


*DanAds*, available at: https://danads.com/.

Davies, Jessica, ‘French publishers are joining forces to take on Google and Facebook,’ *Digiday*, 10 July 2017, available at: https://digiday.com/media/french-publishers-joining-forces-take-google-facebook.


Lekach, Sasha, ‘Say goodbye to bookmarking service Delicious,’ *Mashable*, 2 June 2017, available at: https://mashable.com/2017/06/02/rip-delicious-bookmarking-site/?europe=true#LL5EGy4Zb8q2.


‘Supply Side Platforms Market Share Table,’ Datanyze, available at: https://www.datanyze.com/market-share/sspss/Australia/.

‘Switch from Android to iPhone,’ Apple, available at: https://www.apple.com/uk/switch/.


Vincent, James, ‘Snapchat will start showing ads between your friends’ stories,’ The Verge, 14 June 2016, available at: https://www.theverge.com/2016/6/14/11930386/snapchat-ads-api-stories.


‘What data does Amazon collect and use?’ Amazon, available at: https://www.amazon.co.uk/gp/help/customer/display.html?nodeId=G6RZ4RMNMLUQRLY2.


‘WOW Personalisation,’ Quantum, YouTube, 3 November 2016, available at: https://youtu.be/HPi2-Q6oZUQ.
CATHERINE TUCKER

MIT Sloan School of Management
100 Main St, E62-536
Cambridge MA 02142

http://mitmgmtfaculty.mit.edu/cetucker/

EDUCATION

Stanford University, Ph.D. in Economics (Advisor: Tim Bresnahan), 2005
Oxford University, BA in Politics, Philosophy and Economics, 1999

APPOINTMENTS

MIT Sloan, Sloan Distinguished Professor of Management Science, September 2015 –
MIT Sloan, Chair MIT Sloan PhD Program, July 2015 –
MIT Sloan, Professor of Management Science, July 2015 –
MIT, Co-Founder of the MIT CryptoEconomics Lab, 2018 -
National Bureau of Economic Research (NBER), Research Associate, September 2012 –
MIT Sloan, Mark Hyman Jr. Career Development Professor (with tenure), July 2012 –
MIT Sloan, Associate Professor of Management Science, July 2011 – July 2015
National Bureau of Economic Research (NBER), Faculty Research Fellow, May 2011 –
September 2012
MIT Sloan, Douglas Drane Career Development Chair in IT and Management, July 2006 –
MIT Sloan, Assistant Professor of Marketing, July 2005 – June 2011
HONORS AND AWARDS

2018 ISMS Long Term Impact Award
2018 O’Dell Award
2018 MSI Scholar
2017 Nominated for Teacher of the Year award (Also in 2012, 2010 and 2009)
2015 Erin Anderson Award
2014 Paul E. Green Award
2013 Teacher of the Year Award, MIT Sloan
2013 Jamieson Prize for Excellence in Teaching
2012 Garfield Economic Impact Award for Best Paper in Health Economics
2011 WHITE Award for best paper in the Economics of Healthcare IT
2011 Public Utility Research Prize for the best paper in regulatory economics
2011 NSF CAREER Award
2011 MSI Young Scholar
2010 Management Science Distinguished Service Award
2004 Koret Foundation Scholar, Stanford Institute for Economic Policy Research Fellowship
2004 Fourth Annual Claire and Ralph Landau Student Working Paper prize

PUBLISHED/ACCEPTED PAPERS


   • Republished as part of INFORMS ‘Healthcare in the Age of Analytics’ series


   - Nominated for John D. C. Little Award
   - Nominated for Long Term Impact Award 2017
   - Long Term Impact Award 2018


17. ‘Active Social Media Management: The Case of Health Care’ with Amalia Miller. *Information Systems Research* Vol. 24, No. 1, March 2013, pp. 52-70
   - Republished as part of Inform’s ‘Healthcare in the Age of Analytics’ series


   - Paul E. Green Award for the ‘Best article in the Journal of Marketing Research that demonstrates the greatest potential to contribute significantly to the practice of marketing research.’
   - William O’Dell Award. This award award honors the JMR article published in 2013 that has made the most significant, long-term contribution to marketing theory, methodology, and/or practice


   - Citation of Excellence Award Emerald Publishing


27. ‘Standardization and the Effectiveness of Online Advertising’ with Avi Goldfarb. *Management Science* Vol 61, No. 11, 2015, pp 2707-2719


31. ‘Network Stability, Network Externalities, and Technology Adoption’ in *Advances in Strategic Management*, Volume 37, pp.151 - 175


33. ‘Digital Content Aggregation Platforms: The Case of the News Media.’ with Lesley Chiou - Forthcoming at *Journal of Economics & Management Strategy*

34. ‘Privacy Protection, Personalized Medicine and Genetic Testing’ with Amalia Miller. Forthcoming at *Management Science*

35. ‘Digital Economics’ with Avi Goldfarb. Forthcoming at *Journal of Economic Literature*

36. ‘Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads ’ with Anja Lambrecht. Forthcoming at *Management Science*


38. ‘Search Engine Advertising - Examining a profitable side of the long tail of advertising that is not possible under the traditional broadcast advertising model’ with Avi Goldfarb, *Communications of the ACM*, Vol. 51 No. 11, November 2008, pp. 22-24

A-5


48. ‘Privacy and the Internet’ Chapter 11, Handbook of Media Economics, 2016, Edited by Simon Anderson and Joel Waldfogel

49. Frontiers of Health Policy: Digital Data and Personalized Medicine, Innovation Policy and the Economy, Vol. 15, 2016, Josh Lerner and Scott Stern (Eds), NBER


51. ‘Field Experiments in Marketing,’ with Anja Lambrecht, Handbook of Marketing Analytics, Forthcoming

52. ‘Can Big Data Protect a Firm from Competition?’, CPI Chronicle, January, 2017 with Anja Lambrecht


---

**BOOKS EDITED**

55. Economic Analysis of the Digital Economy, University of Chicago Press, 2015, with Avi Goldfarb and Shane Greenstein

56. The Economics of Digitization, Edward Elgar Publishing, 2013., with Avi Goldfarb and Shane Greenstein

---

**POLICY WRITING**


58. Written Congressional Testimony on ‘Internet Privacy: The Impact and Burden of European Regulation,’ U.S. House Energy and Commerce Committee, September 2011


---

**PAPERS UNDER REVIEW**

60. ‘Social Advertising’. Revise and resubmit at *Management Science*

61. ‘How Do Restrictions on Advertising Affect Consumer Search?’ with Lesley Chiou. Revise and resubmit at *Management Science*

62. ‘Patent Trolls and Technology Diffusion: The Case of Medical Imaging’ Revise and resubmit at *RAND Journal of Economics*

63. ‘Third-Party Certification: The Case of Medical Devices’ with Cristina Nistor Revise and resubmit at *Management Science*
64. ‘Guns, Privacy and Crime’ with Alessandro Acquisti Revise and resubmit at
*Information Systems Research*

65. The Surprising Breadth of ‘Harbingers of Failure’ with Duncan Simester and Clair
Yang. Revise and resubmit at *Journal of Marketing Research*

66. ‘Tensile Promotions in Displays Advertising’ with Anja Lambrecht Revise and resubmit
at *Quantitative Marketing and Economics*

67. ‘A New Method of Measuring Online Media Advertising Effectiveness: Prospective
Meta-Analysis in Marketing’ with Gui Liberali, Glen L. Urban, Benedict G. Dellaert,
Yakov C. Bart, and S. Stremersch.

68. ‘Personalizing mental fit for online shopping applications - How the success of
recommendations depends on mental categorization and mental budgeting’ with Oliver
Emrich and Thomas Rudolph

69. ‘The Digital Privacy Paradox: Small Money, Small Costs, Small Talk’ with Susan Athey
and Christian Catalini

70. ‘Information Shocks and Internet Silos: Evidence from Creationist Friendly
Curriculum’ with Ananya Sen

71. ‘Government Surveillance and Internet Search Behavior’ with Alex Marthews

72. ‘Does IT Lead to More Equal or More Unequal Treatment? An Empirical Study of the
Effect of Smartphone Use on Social Inequality in Employee-Customer interactions’
with Shuyi Yu

73. ‘Antitrust and Costless Verification: An optimistic and a pessimistic view of the
implications of blockchain technology’ invited at ‘Antitrust Law Journal: Innovative
Antitrust with Christian Catalini

74. ‘How Effective Is Black-Box Digital Consumer Profiling And Audience Delivery?:
Evidence from Field Studies’ with Nico Neumann and Tim Whitfield

---

**WORK IN PROGRESS**

*Manuscripts*

75. Health IT and Ambulatory Care Quality with Carole R. Gresenz, Scott Laughery, and
Amalia R. Miller

77. ‘Testimonial Advertising on Social Networks to Existing Customers and New Customers’ with Shuyi Yu

Data Analysis


79. ‘Big Bad Data: The Case of For-Profit College Advertising’ with Avinash Gannamaneni and Avi Goldfarb

80. ‘Policing and Social Media: How police response times vary with YouTube postings’ with Arvind Karunakaran

81. ‘The Circularity of Marketing Communications in the Marketing Funnel: Evidence from a Field Experiment’ with Anja Lambrecht

82. ‘Nationalism, Xenophobia, Globalization and Global Brand Reach’ with Willem Smit

83. ‘Sexism, Ageism and Social Media Usage’ with Willem Smit

84. ‘Spillovers from Product Failure’ with Amalia Miller

85. ‘The Role of Marketing in ICOs’ with Christian Catalini

86. ‘The Shifters and Virality of Hate Speech Online’ with Uttara Ananthakrishnan

Data Collection

87. ‘Mergers and Big Data: Evidence from Healthcare’ with Amalia Miller

88. ‘The Lack of Appeal of Cross-Partisan Appeals: Evidence from an Experiment on Facebook’ with Christina Tucker

89. ‘Can the way someone interacts with a new technology predict their future career?’ with Christian Catalini

INVITED SEMINARS

Universities
1. November 2018, Marketing Group, HEC Paris, France
2. November 2018, Cass Business School, City University of London, UK
3. October 2018, Marketing Group, University of Amsterdam, Netherlands
4. October 2018, Marketing Group, King’s Business School, King’s College, London
5. September 2018, Marketing Group, University of Frankfurt, Germany
6. June 2018, Harbin Institute of Technology, China
7. February 2018, IS/OM Group, New York University, NY
8. November 2017, Marketing Group, Rochester University, NY
9. October 2017, Marketing Group, Maryland University, MD
10. May 2017, Marketing Group, Old Dominion University
11. April 2017, Marketing Group, University of Southern California
12. March 2017, Marketing Group, Arison School of Business, IDC, Israel
13. January 2017, Distinguished Speakers Series, McGill University, Canada
15. August 2016, Southern Jiatong University, Sichuan, China
16. May 2016, Chapman University, Marketing Group
17. April 2016, Carnegie Mellon University, Public Policy Group
19. March 2016, INSEAD, Marketing Group
20. March 2016, University of Paris-Sud, Privacy Research Group
21. March 2016, Vienna University of Economics and Business, Marketing Group
22. September 2015 University of Maryland, IS Group
23. June 2015, Marketing Group, University of Cambridge, UK
24. May 2015, Marketing Group, University of Texas at Dallas, TX
25. March 2015, Health Policy Group, Georgia State University, GA
26. March 2015, Marketing Group, University of Colorado, CO
27. February 2015, Strategy Group, University of North Carolina, NC
28. January 2015, Marketing Group, Emory University, GA
29. December 2014, OPIM, Wharton School of Management, PA
30. October 2014, Economics Department, Yale University, CT
31. September 2014, Marketing Group, Boston University, MA
32. March 2014, Technology Group, University of California at Berkeley, CA
33. January 2014, Marketing Department at Texas A&M
34. November 2013, Marketing Group, University of California at Berkeley, CA
35. October 2013, Marketing Group, Tulane University, LA
36. October 2013, Marketing Group, University of Houston, TX
37. May 2013, Tuck School of Management, Dartmouth University, NH
38. March 2013, Economics Department, University of Toulouse
39. March 2013, Marketing Group, Rotterdam University
40. March 2013, Economics Department, University of Zurich
41. March 2013, Marketing group, Georgia Tech
42. January 2013, Anderson School, UCLA
43. January 2013, Marketing Group, CMU
44. October 2012, Marketing Group, Stanford University
45. October 2012, Marketing Group, Columbia University
46. October 2012, Marketing Group, University of Texas at Austin
47. September 2012, Marketing Group, Harvard Business School
49. March 2012, Marketing Group, Cornell
50. February 2012, IS Group, Indian School of Business
51. February 2012, Marketing Group, Wharton
52. January 2012, Marketing Group, UCLA
53. November 2011, Marketing Group, University of Rochester
54. October 2011, Marketing Group, University of Zurich
55. October 2011, Department of Law and Economics, Swiss Federal Institute of Technology, Zurich
56. May 2011, Marketing Group, National University of Singapore
57. May 2011, IS Group, National University of Singapore
58. May 2011, Strategy Group, LMU Munich
59. May 2011, Marketing Group, New York University
60. March 2011, Marketing Group, Florida University
61. February 2011, IS Group, New York University
62. November 2010, European School of Management and Technology
63. October 2010, Marketing Group, Yale University
64. October 2010, Networked Business Group, Harvard Business School
65. September 2010, TIES Group, MIT Sloan
66. July 2010, Department of Economics, University of Mannheim
67. March 2010, Marketing Group, Wharton School, University of Pennsylvania
68. January 2010, Marketing Group, University of Michigan
69. November 2009, Marketing Group, University of California at Berkeley
70. October 2009, Digital Business Seminar, MIT Sloan
71. December 2008, Marketing Group, MIT Sloan
72. November 2008, Marketing Group, Rady School of Business, UCSD
73. September 2008, Strategy Group, MIT Sloan
74. May 2008, Digital Strategy Group, Tuck School of Business, Dartmouth University
75. April 2008, Kellogg Management and Strategy Group, Northwestern University
76. March 2008, Marketing Group, Duke University
77. March 2008, Strategy Group, Chicago GSB
79. April 2007, Marketing Group, Chicago GSB
80. March 2007, Marketing Group, Rotman School, University of Toronto
81. November 2005, Economics Department, Harvard University

Other
83. January 2018, IMF
84. December 2017, Technology Policy Institute
85. October 2016, Federal Communications Commission
86. April 2015, Federal Communications Commission
87. November 2014, Office of Research at the Consumer Financial Protection Bureau
89. February 2014, Main Street Patent Coalition, Panel hosted at the Senate by Senator Orrin Hatch
90. July 2013, Federal Communications Commission
91. August 2012, DG Competition, European Commission, Brussels
92. August 2012, Technology Policy Institute Conference, Aspen
94. June 2011, Eneca
95. September 2010, Federal Trade Commission
96. September 2010, Google European Public Policy Unit, Paris

PRESENTATIONS OF RESEARCH AT CONFERENCES

1. June 2018, Antitrust and Big Data, Penn Wharton China Center Conference, Beijing
2. June 2018, Marketing Science
3. May 2018, Boston College Digital Innovation Workshop
4. December 2017, Mobile Marketing and Big Data Conference, NYU
5. September 2017, NBER Economics of AI Conference
7. July 2017, NBER Digitization Meetings
8. June 2017, Marketing Science
10. May 2017, Boston College Digital Innovation Workshop
11. November 2016, ICANN Public Meetings
12. October 2016, Conference on Digital Experimentation, Cambridge, MA
15. May 2016, Competing with Big Data, Brugel, Brussels, Belgium
16. April 2016, NBER Innovation and Policy, Washington DC
17. April 2016, Financial Services Roundtable, NYC
18. March 2016, Digitization Tutorial, NBER
20. July 2015, NBER Law and Economics (co-author presented), Cambridge, MA
23. June 2015, Marketing Science, Baltimore, MD
25. March 2015, IP Leadership Conference, Washington, DC

A-12
27. June 2014, Marketing Science, Atlanta, GA
28. May 2014, Boston College Social Media Workshop, Boston, MA
30. July 2013, Marketing Science, Istanbul, Turkey
31. June 2013, Searle Center Conference on Internet Search and Innovation, Chicago, IL
32. April 2013, Brown University Mini-Networks Conference
33. February 2013, WSDM 2013 Conference (Keynote Speaker), Rome, Italy
34. January 2013, American Economic Association Meetings, San Diego, CA (Co-author presented)
36. November 2012, Search and Competition Conference, Melbourne Australia
37. October 2012, Economics of Personal Data, (Keynote Speaker), Amsterdam
38. August 2012, Amsterdam Symposium on Behavioral and Experimental Economics
39. July 2012, Fudan University Marketing Research Symposium, China
40. June 2012, Searle Center Conference on Internet Search and Innovation, Chicago, IL
41. June 2012, Innovation, Intellectual Property and Competition Policy Conference, Tilburg, Netherlands
42. June 2012, Marketing Science, Boston, MA
43. June 2012, Social Media and Business Transformation, Baltimore, MD
44. May 2012, The Law and Economics of Search Engines and Online Advertising, George Mason University, Arlington, VA
45. February 2012, NBER Economics of Digitization (co-author presented), Cambridge, MA
46. January 2012, Symposium on Antitrust and High-Tech Industries, George Mason University, VA
47. January 2012, Patents, Standards and Innovation, Tucson, AZ
48. January 2012, Econometric Society Meetings, Chicago, IL
49. January 2012, AEA Meetings (2 papers), Chicago, IL
50. December 2011, Economics of Privacy Workshop, Boulder, CO
52. November 2011, HBS Strategy Research Conference, Boston, MA
53. November 2011, The Law and Economics of Internet Search and Online Advertising Roundtable, George Mason University, Arlington, VA
55. October 2011, Workshop on Health IT and Economics, Washington, DC
56. October 2011, Innovation, Organizations and Society, University of Chicago, IL
57. October 2011, Direct Marketing Research Summit, Boston, MA
59. July 2011, NBER Economics of Digitization, Cambridge, MA
60. July 2011, SICS, Berkeley, CA
61. June 2011, The Law and Economics of Search Engines and Online Advertising, George Mason University, Arlington, VA
63. June 2011, Marketing Science (3 papers), Houston, TX
64. June 2011, Searle Center Conference on Internet Search and Innovation, Chicago, IL
65. May 2011, Boston College Social Media Workshop, Boston, MA
66. May 2011, Technology Pricing Forum, Boston, MA
67. April 2011, NBER Innovation Policy and the Economy, Washington, DC
68. April 2011, International Industrial Organization Conference (3 papers), Boston, MA
69. March 2011, Technology Policy Institute, Washington, DC
70. February 2011, NBER Economics of Digitization (co-author presented), Palo Alto, CA
71. January 2011, Sixth bi-annual Conference on The Economics of Intellectual Property, Software and the Internet (2 papers, plenary speaker), Toulouse, France
72. January 2011, MSI Young Scholars Conference, Park City, UT
73. December 2010, Workshop on Information Systems and Economics, Washington University of St. Louis (co-author presented), St. Louis, MO
74. December 2010, OECD Economics of Privacy Roundtable, Paris, France
75. November 2010, Net Institute Conference, New York, NY
76. October 2010, Workshop on Media Economics and Public Policy (co-author presented), New York, NY
77. October 2010, Workshop on Health IT and Economics, Washington DC
78. September 2010, ITIF and CAGW Privacy Working Group Meetings, Washington, DC
79. September 2010, Medical Malpractice Conference, Mohegan, CT
80. September 2010, Search and Web Advertising Strategies and Their Impacts on Consumer Workshop, Paris, France
81. July 2010, NBER Meetings (IT), Cambridge, MA
82. July 2010, NBER Meetings (Healthcare and IT), Cambridge, MA
83. July 2010, SICS, Berkeley, CA
84. July 2010, Keynote Speaker, 8th ZEW Conference on the Economics of Information and Communication Technologies, Mannheim, Germany
85. June 2010, American Society of Health Economists Conference, Cornell, NY
86. June 2010, Marketing Science (2 papers), Koeln, Germany
87. June 2010, Workshop on the Economics of Information Security (2 papers), Harvard, MA
88. January 2010, AEA Meetings, Atlanta, GA
90. November 2009, WPP/Google Marketing Awards, Cambridge, MA
91. July 2009, NBER meetings (IT), Cambridge, MA
92. June 2009, IHIF Debate on Privacy, Washington, DC
93. June 2009, Marketing Science, Ann Arbor, MI
94. April 2009, International Industrial Organization Conference, Boston, MA
96. January 2009, Modeling Social Network Data Conference, Philadelphia, PA
97. July 2008, NBER Meetings (Productivity), Cambridge, MA
98. July 2008, SICS, Berkeley, CA
100. June 2008, Marketing Science, Vancouver, BC
102. April 2008, Net Institute Conference, New York, NY
103. November 2007, NBER Health Meetings (Co-author presented), Boston, MA
104. July 2007, SICS, Berkeley, CA
Professional Service

- **Associate Editor:** Management Science, Marketing Science, Journal of Marketing Research, International Journal of Research in Marketing
- **Associate Editor:** Information Systems Research, Special Issue on Social Media and Business Transformation
- **Departmental Editor:** Quantitative Marketing and Economics
- **Editor:** The Economics of the Internet, Palgrave Dictionary of Economics
- **Co-Editor:** NBER: The Economics of Digitization - An Agenda
- **Co-Editor:** Information Economics and Policy, Special Issue on Economics of Digital Media Markets
- **Editorial Review Board:** Journal of Marketing, ISR Special Issue on Managing Digital Vulnerabilities, Journal of Economic Literature

- **Conference Program Committees**
  - 2018 Co-organizer, NBER Conference on the Economics of Artificial Intelligence
  - 2018 Scientific Committee: ZEW Conference on the Economics of Information and Communication Technologies
  - 2018 Program Committee: Workshop on the Economics of Information Security
  - 2019 Scientific Committee: IP Statistics for Decision Makers
  - 2017 Scientific Committee: IP Statistics for Decision Makers
  - 2017 Scientific Committee: ZEW Conference on the Economics of Information and Communication Technologies
  - 2017 Program Committee: Workshop on the Economics of Information Security
  - 2016 Program Committee: Workshop on the Economics of Information Security
  - 2016 Scientific Committee: ZEW Conference on the Economics of Information and
Communication Technologies
- 2015 Scientific Committee: Competition, Standardization and Innovation
- 2015 Associate Editor: ICIS 2015, Healthcare track
- 2015 Scientific Committee: European Association for Research in Industrial Economics
- 2015 Program Committee: ACM Conference on Economics and Computation
- 2015 Program Committee: Workshop on the Economics of Information Security
- 2015 Chief-Organizer: Quantitative Marketing and Economics Conference
- 2015 Scientific Committee: ZEW Conference on the Economics of Information and Communication Technologies
- 2014 Scientific Committee: European Association for Research in Industrial Economics
- 2014 Scientific Committee: Conference on the Economics of Information and Communication Technologies
- 2014 Program Committee: International Conference on Big Data and Analytics in Healthcare
- 2013 Program Committee: Quantitative Marketing and Economics
- 2013 Scientific Committee: European Association for Research in Industrial Economics Conference
- 2013 Scientific Committee: Conference on the Economics of Information and Communication Technologies
- 2013 Program Committee: Workshop on the Economics of Information Security
- 2013 Associate Editor of Personal Data Markets Track: ECIS 2013
- 2012 Program Committee: European Association for Research in Industrial Economics Conference
- 2012 Program Committee (Conference Organizer) NBER: The Economics of Digitization Pre-Conference, June 2012
- 2012 Scientific Committee: Conference on the Economics of Information and Communication Technologies
- 2012 Senior Program Committee: 13th ACM Conference on Electronic Commerce
- 2012 Program Committee: Workshop on the Economics of Information Security
- 2011 Scientific Committee: European Association for Research in Industrial Economics Conference
- 2011 Scientific Committee: Conference on the Economics of Information and Communication Technologies
- 2011 Program Committee: Ad Auctions Workshop
- 2011 Program Committee: Workshop on the Economics of Information Security
- 2010 Program Committee: Workshop on IT and Economic Growth
- 2010 Program Committee: Conference on Health IT and Economics
- 2010 Program Committee: Workshop on the Economics of Information Security
- 2009 Program Committee: Workshop on the Economics of Information Security
- 2008 Program Committee: Workshop on the Economics of Information Security
- 2008 Program Committee: Ad Auctions Workshop

External Affiliations
- **Affiliate:** CESifo Research Network
• **Advisory Board:** Future of Privacy Forum

---

**MIT Service**

- 2015: Faculty Chair, PhD program
- 2015: EMBA Committee
- 2015: ASB Committee
- 2014: MIT Sloan Gender Equity Committee
- 2013-2014 Group Head, Marketing Group
- 2013-2014 Chair, Marketing Faculty Search Committee
- 2013-2014 MIT Committee on Undergraduate Admissions and Financial Aid
- 2011 North East Marketing Conference Coordinator
- 2011 MIT Sloan Marketing Conference, Panel Moderator
- 2011 Sloan Women in Management Conference, Panel Moderator
- 2005, 2008, 2012 Marketing Faculty Search Committee

---

**Advising**

- 2019: Shuyi Yu, PhD Thesis supervisor
- 2016: Abhishek Nagaraj, PhD Thesis advisor
- 2012: Cristina Nistor, PhD Thesis advisor
- 2010: Katherine Molina, Masters Thesis
- 2008: Dinesh Shenoy, Masters Thesis
- 2007: James Kelm, Masters Thesis
GRANTS AND SUPPORT

Academic Research Grants

2018 Sloan Foundation Grant (2018-2021), ‘NBER Project on the Economics of Artificial Intelligence’. Principal Investigator: $914,250
2017 Net Institute Grant: $3,000
2016 Net Institute Grant: $6,000
2013 MSI research Grant 4-1840: $10,200
2011 Tilburg Law and Economics Center (TILEC) IIPC grant: $21,000
2011 Google Grant: $50,000
2011 Junior Faculty Research Assistance Program: $30,000
2011 Net Institute Grant: $6,000
2011 NBER Digitization Grant: $20,000
2011 NSF CAREER Award: $502,000
2010 Time-Warner Research Program on Digital Communications: $20,000
2010 Net Institute Grant: $6,000
2009 Net Institute Grant: $6,000
2009 The James H. Ferry, Jr. Fund for Innovation in Research Education: $50,000
2009 Google/WPP Grant: $55,000
2008 Net Institute Grant: $15,000
2007 Net Institute Grant: $8,000
2006 Net Institute Grant: $8,000

Industry Research Grants

2015 CCIA Research: Research into Sustainable Competitive Advantage and Big Data: $60,000
2015 E-Logic: Research into Vertical Mergers and Patent Litigation: $60,000
2014 CCIA Research: Research into Patent Litigation and Entrepreneurship: $100,000
2012 Google Australia: Research into Measurement and Attribution: $50,000

EXPERT TESTIMONY

- In Re Appraisal of AOL Inc: Consolidated C.A. No. 11204-VCG. Chancery Court of Delaware
  - Expert Report, Deposition and Trial Testimony (2017)
- Michael Edenborough v. ADT, LLC, d/b/a ADT Security Services, Inc. Case No: 3:16-cv-02233-JST United States District Court, Northern District of California, San Francisco Division
  - Declaration (2017).
- Red Online Marketing Group LP, d/b/a 50onRED v. Revizer Ltd., d/b/a Ad Force Technologies, Ltd., and Revizer Technologies, Ltd. United States District Court, Eastern District of Pennsylvania Civil Action No. 14-1353
- Matthew Campbell and Michael Hurley et al. v. Facebook, Inc. Case No. C 13-05996 PJH. United States District Court Northern District of California
- GO Computer, Inc. et al. v. Microsoft Corporation Case No. CGC-05-442684 Superior Court of the State Of California for the City and County of San Francisco
- In re: Chapter 11, Nortel Networks, Inc., et al., Debtors, U.S. Bankruptcy Court, District of Delaware, Case No. 09-10138(KG) (Jointly Administered), Re Dkt No. 13208. Deposition and Trial Testimony (2014)
- Angel Fraley, et al., Plaintiffs, v. Facebook, Inc., a corporation; and DOES 1-100, Defendants, U.S. District Court, Northern District of California, Case No. 5:11-cv-01726-LHK. Deposition Testimony (2012)

TEACHING

- 15.818, Pricing (MBA Elective) 2006-
- 15.732, Marketing Management for Senior Executives 2012-
- 15.726, Pricing (EMBA Elective) 2012-
- 15.838, Doctoral Seminar, Spring 2006, Fall 2007, Fall 2013
- Marketing Management, Asian School of Business, 2016
- Guest Lecturer: HST.936: Health information systems to improve quality of care in resource-poor settings, 2014
- Executive Education: Blockchain Technologies: Business Innovation and Application, 2018-
- Executive Education: Marketing Innovation, 2016-
- Executive Education: Pricing 4dX, 2016-
- Executive Education: Strategic Marketing for the Technical Executive, 2012-2015
- Executive Education: Systematic Innovation of Products, Processes, and Services, 2013-
- Executive Education: Platform Strategy: Building and Thriving in a Vibrant Ecosystem, 2014-
- Executive Education: Global Executive Academy (multi-language), 2013, 2014
- Executive Education: Entrepreneurship Development Program, 2012-
- Faculty Coach, Takeda Leadership Academy, 2016-18