For Publication on the Authors' Web Pages

From Mad Men to Maths Men: Concentration and Buyer Power in Online Advertising

Web Appendix

A) Data Details

The data used in the paper come from several sources. First, from Redbooks we obtained data on advertisers, their MAs and their network affiliations. Access to the data is available at Winmo which sells the Redbooks data through the demo request form available here:

https://www.winmo.com/redbooks-agency-and-advertising-database/

In order to benchmark the information on M&A contained in the Redbooks data we relied on the *Zephyr* dataset on M&A, IPO, Private Equity and Venture Capital by Bureau Van Dijk. Data can be purchased through the demo request form accessible at:

https://www.bvdinfo.com/it-it/our-products/economic-and-m-a/m-a-data/zephyr

We complement the data on advertisers with information on bids, keywords and advertisers provided by SEMrush, the most important and renowned provider of SEM data and related services. Access to the data can be purchased at this link:

https://www.semrush.com/prices/

We obtained the data on Click-Through Rate at the industry/month level, position by position, from AdvancedWebRanking by Caphyon a provider of internet data services. The data are freely available at:

https://www.advancedwebranking.com/ctrstudy/

In order to proceed with the thematic clustering we used a pre-trained set of GloVe word vectors—more specifically, we used the Common Crawl, 840B tokens, 2.2 million words, 300d vectors—publicly available and open source at:

https://nlp.stanford.edu/projects/glove/

In this web appendix, we also present results based on two additional data sources. In section D, we report the result of a survey that we run on Amazon Mechanical Turk.[1] In section K, we analyze the ownership structure of the networks using the Refinitiv Eikon. This is a dataset and financial analysis tool provided by Thomson Reuters. Paid and free trial subscriptions are available. Researchers can subscribe to the service at:

https://eikon.thomsonreuters.com/

[1] The complete dataset as well as a sample of the survey is available as part of the replication files provided. See details in section D below.

In Table A.1, we summarize all the main variables used in this study, reporting their source, frequency and their short description. While the Redbook data have been previously used in economics and marketing studies—see Dai [2014] for a recent example—to the best of our knowledge the SEMrush data are new to the literature.

Table A.1: Raw variables' description and sources

Variable Name	Source	Frequency	Definition
Semrush			
keyword	www.semrush.com	year/advertiser	The keyword bringing users to the website via search results – that is, the keyword advertisers bid on
position		year/keyword/advertiser	The position of the domain in paid search for the given keyword at the specified period
searchvolume		year/keyword	Number of search queries for the given keyword in the last 12 months
CPC		year/keyword	Average price advertisers pay for a users click on an ad triggered by the given keyword
traffic		year/keyword/advertiser	The share of traffic driven to the website with the given keyword for the specified period
competition		year/keyword	Competitive density of advertisers using the given term for their ads
results		year/keyword	The number of URLs displayed in organic search results for the given keyword
Redbooks			
enterprise_nbr	www.winmo.com	year/advertiser	Advertiser's ID code
company_name		year/advertiser	Advertiser's business name
website		year/advertiser	Advertiser's website
agency_ID		year/advertiser	Digital Marketing Agency (MA) ID code - possibly with multiple matches per advertiser
agency_name		year/advertiser	Digital Marketing Agency (MA) business name
digital		year/agency	Indicator function for digital agency
parent_ent		year/agency	Agency owner ID code - mainly belonging to 7 networks
industry		year/advertiser	Core business industry of the advertiser
Advanced Web Ranking			
CTR	www.advancedwebranking.com	month/industry/position	Click-through rate: average number of clicks per impression
GloVe			
key_vec	nlp.stanford.edu/projects/glove/	keyword tokens	Set of GloVe vectors pre-trained on Common Crawl, 840B tokens, 2.2 million words, reported in 300 dimensions
Amazon MT			
	https://www.aeaweb.org/journals/aer	keyword/respondent	Responses to the cluster validation task, described in section D fo this appendix
Refinitiv Eikon			
	https://eikon.thomsonreuters.com/	yearly/network	Ownership share of the 5 publicly traded networks in the 2010-2019 period for the 10 largest shareholders.

Notes: summary of the raw variables that we use in the paper. We report the variable name, the data source, the raw frequency—as used for the analysis—and a brief description.

Redbooks data are the digital version of the Standard Directory of Advertising Agencies, better known as the Red Book. This has been the "gold standard" for advertisers and agencies for over 100 years. Over time, it absorbed other directories, like McKittrick's

Directory and the Standard Advertising Register, making it the single, most comprehensive directory of the connections between advertisers and marketing agencies. In 2018, it was acquired by Winmo which currently distributes the Redbooks data among its services. The data contains profiles on the universe of advertising agencies active in the US including their location, corporate contact names, area of specialization and, starting in 2014, the identity of their agency network, if any. Redbooks also links the 6,000 largest advertisers active in the US market to the advertising agencies that work for them. The data is updated annually through a combination of machine learning algorithms scanning over half a million news sources. A specialized content team then verifies, through direct contacts with the companies, the correctness of the information.

SEMrush is a leading provider of sponsored search data and this is why we selected it for this study.[2] Importantly, this implies that the data that we use tend to be the same as that used by many players in this market to set their strategies. Data like those we obtained from SEMrush represent a way to have an overview of the entire market—like those that the internal data from search engines would give—but without the limitations that might be posed by using the internal records of search engines in terms of advertiser identities and prices.[3] A limitation of the data is, however, the non-fully transparent way that the yearly averages are calculated: proprietary algorithms are used to aggregate data from multiple providers and assemble the SEMrush data. As is typical in this industry, Google's Keyword Planner is a key source for accessing CPC data which would otherwise be not observable.[4]. Although Google's Keyword Planner itself does not report the exact algorithms used to calculate the average CPC, its data are accurate and all the rich dynamics that might characterize bidding on a keyword throughout a year should contribute to the formation of the average. Averaging, while leading to some information loss, is needed to form an overall view of such a highly dynamic and fragmented market. In our study, this is made even more necessary by the yearly nature of the Redbooks data. Although these are important limitations, the data that we use are likely representative of those available to many advertisers and intermediaries and are of comparable quality and extent to what might be available from other publicly accessible sources.[5]

Regarding the CTR data that we use, there are a few limitations worth discussing. In particular, since we lack keyword-level click through rates, we impute this from a market

[2]SEMrush was launched in 2008. It gained and maintained a leadership position, frequently winning awards as a top SEO and SEM tool in the last few years, including best SEO suite 2017, "US & UK search awards" and "European Search Awards."

[3]To the best of our knowledge, no published study using internal search engines' data contains this type of information.

[4]This is typically done programmatically using the services of the likes of TargetingIdeaService API, see `https://developers.google.com/adwords/api/docs/guides/targeting-idea-service`. SEMrush's CPC is an average of the past 12 months and is updated on a monthly basis. See: `https://www.semrush.com/kb/162-monthly-numbers`

[5]Indeed, to further ensure that we were not missing some important (and possibly more disaggregated) data, we compared our SEMrush data to what is available from SpyFu, one of the main SEMrush' s competitors. We found that the variables available are essentially the same, but that the CPC data is reported in a more informative way on SEMrush than on SpyFu: SEMrush reports the CPC across all positions, while SpyFu reports that associated with being (on average) in the second position.

average using data from Advanced Web Rankings. However, the research question in this study involves structures that are more aggregate than individual keywords, thus an aggregation is unavoidable.[6] Moreover, keyword-specific CTRs are in most cases useless as they are all just zeros for the obvious reason that most keywords are infrequently searched and even less frequently generate clicks. Hence, using CTRs typically requires substantial aggregation across large sets of keywords and/or over long period of times. In appendix F below, we return to the issue of the reliability of our CTR measure by evaluating the robustness of our estimates to measurement errors in the CTR.

Table A.2: M&A Operations across All Networks, 2014-2017

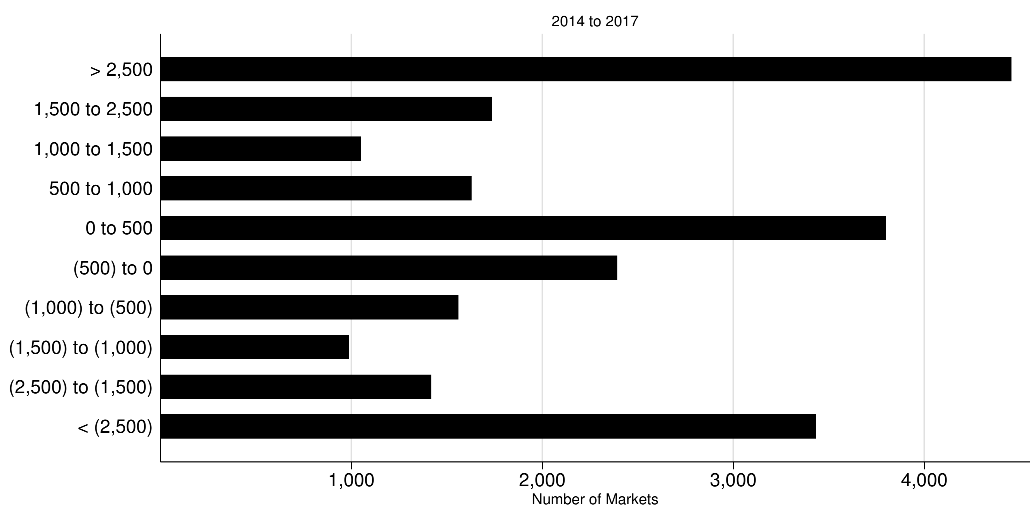
Agency	Acquiring Network	Acquisition year	Number of Advertisers	Number of Industries	Number of Markets
The Brooklyn Brothers	IPG	2016	6	2	23
Essence Digital Limited	WPP	2015	1	1	145
Quirk	WPP	2015	5	2	272
SHIFT Communications	WPP	2017	13	8	1,049
Deeplocal Inc.	WPP	2017	5	1	117
Maruri GREY	WPP	2017	1	1	150
Zubi Advertising Services, Inc.	WPP	2017	3	2	345
Campfire	Publicis	2015	3	1	27
La Comunidad	Publicis	2015	9	5	271
Sapient Corporation	Publicis	2015	17	6	1,038
Blue 449	Publicis	2016	4	2	93
Forsman & Bodenfors	MDC	2017	5	1	315
Formula PR	Havas	2015	6	4	309
FoxP2	Dentsu-Aegis	2015	1	2	42
Rockett Interactive	Dentsu-Aegis	2015	1	1	22
Covario, Inc.	Dentsu-Aegis	2015	3	1	78
Achtung	Dentsu-Aegis	2016	2	1	226
Gravity Media	Dentsu-Aegis	2016	5	3	433
Grip Ltd.	Dentsu-Aegis	2016	3	2	92
Merkle	Dentsu-Aegis	2017	18	7	973
Gyro	Dentsu-Aegis	2017	12	6	363

Notes: the table reports the set of acquisitions in 2014-2017 by the networks. To identify these events, we used Redbooks data and confirmed them through Zephyr data (Bureau Van Dijk). The table only reports acquisition involving at least 51%+ of the acquired agency. Acquisition prices are typically not disclosed. Exceptions are the cases of *Sapient Corporation*, acquired for \$3.7 billion by Publicis Groupe, and *Merkle* acquired for \$1.5 billion by Dentsu-Aegis in 2016. Furthermore, not listed in the table are two divestments: TM Advertising and Moroch returned independent by buying themselves back from the networks.

[6] If different intermediaries were representing clients that, despite operating in the same industry, were facing systematically different CTRs conditional on position-year, then aggregation would be problematic. However, this appears as an unlikely situation because all advertisers, apart from operating in the same industry, are also all large firms active in online advertising in the US market.

The M&A activity is one of the sources of the HHI growth in the data. Figure A.1 shows the dynamics of HHI in our sample; more specifically, in this figure for each market we take the difference between HHI_{2017} and HHI_{2014}. The figure makes it evident that, although many markets experience an HHI decline, the majority experience concentration increases and about 6,000 markets have an HHI increase of more than 1,500 points. In the data, there has been a merger in one-third of the markets where the HHI falls. In more detail, we report in Table A.3 some summary statistics for two subgroups of markets: those where the HHI increases between 2014 and 2017 and those where it falls. The first row, *Subject to Merger*, reports the statistics for a dummy variable equal to 1 if the market ever involved intermediaries taking part in a merger episode. For both markets with positive and negative HHI changes, we find that one-third of them has been exposed to mergers. This evidence is in line with our favoured interpretation that mergers in the agency sector are happening due to reasons that—to a large extent—are exogenous to the features of the keyword search auction markets. This interpretation is also supported by the other observable market characteristics summarized in Table A.3. Indeed, the two groups of markets are close in terms of revenue, HHI, keyword number and characteristics (number of characters, long-tail and branded). The only noticeable difference involves the number of organic results associated with the market keywords: markets with positive HHI change contain keywords that tend to have more organic results than those of the keywords in the other group of markets.

Figure A.1: Change in HHI – 2014 to 2017



Notes: The bars report the number of markets, on the x-axis, grouped according to the differences between the HHI in 2017 and in 2014, clustered in ten classes. The HHI scale ranges from 0 to 10,000.

Finally, we can use the descriptive evidence in Table A.3 to see how the qualitative evidence from the data is broadly consistent with the instrumental variable analysis in the main text. Indeed, in the last table row we report the value of our main dependent variable, ΔR, for the two groups of markets. It is reassuring that, despite the similarity of the two groups of markets along most observables (including the incidence of intermediary mergers), the group of markets where the HHI grows experiences a substantially lower revenue increase

Table A.3: Summary Statistics - Markets with HHI Growth or Decline

	Positive Changes				Negative Changes			
	Mean	SD	Median	Observations	Mean	SD	Median	Observations
Subject to Merger	0.34	0.48	0.00	15,615.00	0.35	0.48	0.00	13,961.00
$log(R)$	11	2	11	15,615	11	2	11	13,961
HHI	2,864	2,309	2,126	15,615	2,417	1,774	1,978	13,961
# of words	2.89	1.01	2.90	15,615.00	2.87	0.99	2.83	13,961.00
# of characters	17.93	5.92	17.40	15,615.00	17.96	5.82	17.42	13,961.00
Long-tail Keywords	0.24	0.35	0.00	15,615.00	0.24	0.34	0.01	13,961.00
Branded Keyword	0.15	0.30	0.00	15,615.00	0.14	0.29	0.00	13,961.00
Organic Results (million)	78.94	220.60	21.72	15,615.00	65.55	168.11	18.29	13,961.00
ΔR	0.15	1.43	0.11	14,939.00	0.54	1.46	0.39	13,269.00

than the other group.

B) Redbooks Industries and Imputation

In the Redbooks data, advertisers are associated to one out of 23 different macro-sectors, with the three largest ones being *Media*, *Industrial* and *Financial services*. In each sector, the number of advertisers ranges from a handful (*Tobacco* and *Telecom*) to several hundred. For a third of its advertisers, however, Redbooks does not report the information on industry affiliation; hence, we exploited SEMrush data to impute it. In particular, we matched all keywords by advertisers without a reported industry with the keywords by all advertisers for which this information is available: advertisers with a missing industry are then assigned to the industry with which they share most keywords. The industries most affected by cases of imputation are: *Media*, *Apparel*, *Technology*, *Financial Services* and *Industrial*; the least affected are: *Tobacco*, *Telecom*, *Food Retail*, *Restaurants*, *Utilities* and *Food and Beverage*.

C) Vector Representation and Clustering

We proceed in generating vector representations of the keywords by splitting the keywords in our sample, term by term, and by merging them with the GloVe pre-trained set of words. More specifically, we split each keyword $k \in [1, .., K]$ into its consitutent terms $t_k \in [1, .., T_k]$, where T_k is the number of terms in the k^{th} keyword. After stemming we then matched each term with the corresponding GloVe term $t_g \in [1, .., G]$, in our application $G \approx 2.2$ million, and each t_g is a vector in $J = 300$ dimensions. Each vector locates the term/keyword into the GloVe vector space, which is a sub-structure of the classic word-word co-occurrence matrix ([Pennington, Socher and Manning, 2014]). For each keyword, we generate a single vector in J dimensions by summing up all the T_k vectors. If any term was not matched with the GloVe pre-trained sample (it covers $\approx 80\%$ of the terms in our sample), we input a vector of zeros, which does not impact the total sum.

The resulting vector representation $(\overrightarrow{\boldsymbol{d_k}})$ of the K keywords reads:

$$\overrightarrow{\boldsymbol{d_1}} = (d_{1,1}, d_{2,1}, ..., d_{J,1}),$$
$$\vdots$$
$$\overrightarrow{\boldsymbol{d_k}} = (d_{1,k}, d_{2,k}, ..., d_{J,k}),$$
$$\vdots$$
$$\overrightarrow{\boldsymbol{d_K}} = (d_{1,K}, d_{2,K}, ..., d_{J,K}),$$

Step 1: for each industry defined by Redbooks we run a spherical k-means algorithm ($k = 1,000$ in the baseline model) on the matrix of vectorized keywords in order to group them according to their Euclidean distance.[7] Hence, through the first layer of the algorithm we are able to capture the similarities between keywords (i.e., their "distance" in GloVe terms) and make the underlying semantic themes emerge from the data structure itself. The well-known drawback of the k-means algorithm, though, is that the number of clusters is pre-specified and might not reflect the "real" number of topics; in order to address the issue, we run several checks on clustering quality—and we show the robustness of the results to different choices of K.

Step 2: we add a second clustering layer exploiting the structure of the competition within the thematic clusters. More specifically, for each cluster c, we build a $K_c \times N_c$ sparse matrix, whose rows correspond to the keywords in the cluster, and whose columns match the advertisers which, at least once in the data, have participated in one of those keyword auctions—panel B in table C.1. The resulting row vectors, akin to term vectors in text analyses, are projections of the keywords in the space spanned by the advertisers—i.e., the competitive structure space. The underlying assumption is that keywords showing similar patterns of bidders are more likely to belong to the same competitive space, and that the latter has substantial overlaps with the—unobserved—product space. In order to exploit the keyword similarity, we build a matrix of pairwise Euclidean distances among the keywords, in terms of co-occurrency, panel C in table C.1. Each non-diagonal cell a_{ij} represents the distance between keywords i and j, computed with the L2 norm $d()$, that is

$$a_{i,j} = d(\overrightarrow{\boldsymbol{k^i}}, \overrightarrow{\boldsymbol{k^j}}) = \sqrt{\sum_{v=1}^{N_c} (k_v^i - k_v^j)^2}$$

where N_c is the number of advertisers in cluster c.[8]

Finally, we select the best-fitting definition of competitive clustering through a hierarchical clustering algorithm run on the distance matrix. In order to optimally prune the cluster tree we employ the Kelley, Gardner and Sutcliffe [1996] penalty function. A

[7]In the code, we use the standard python libraries `nltk` [Bird, Klein and Loper, 2009] and `sklearn` [Pedregosa et al., 2011], which feature functions for NLP and unsupervised clustering.

[8]In the code, we use the `R` base functions `dist` and `hclust` (package `stats`).

Table C.1: Layer 2 clustering: data preparation

Keyword	Advertiser
key 1	Adv 1
key 1	Adv 2
key 1	Adv 3
key 2	Adv 2
key 3	Adv 2
key 3	Adv 3

A. Actual Data

⇒

	Adv 1	Adv 2	Adv 3
key 1	1	1	1
key 2	0	1	0
key 3	0	1	1

B. Advertisers' co-occurrency

⇒

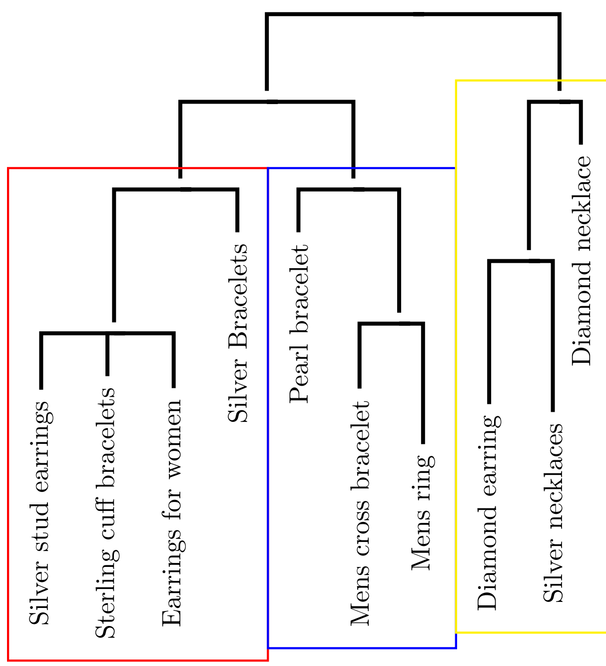
	key 1	key 2	key 3
key 1	0	$\sqrt{2}$	1
key 2	$\sqrt{2}$	0	1
key 3	1	1	0

C. Keyword distance metric

Notes: data preparation for layer 2 clustering. For each thematic cluster, from the keyword auction data listing keywords and advertisers (panel A) we build a matrix of the co-occurrency of advertisers (panel B). Through that, we can compute the pairwise Euclidean distance between keyword vectors in the advertisers' space and build the distance matrix (panel C).

random set of the resulting clusters is available for download and inspection at `https://github.com/GabrieleRovigatti/adNets_clusters`.

Figure C.1: Hierarchical clustering



Notes: Graphical representation of the structure of competitive clusters. The three clusters (red, blue, and yellow boxes) are identified by optimally pruning the thematic clusters through the Kelley, Gardner and Sutcliffe [1996] penalty parameter applied to the keyword distance matrices.

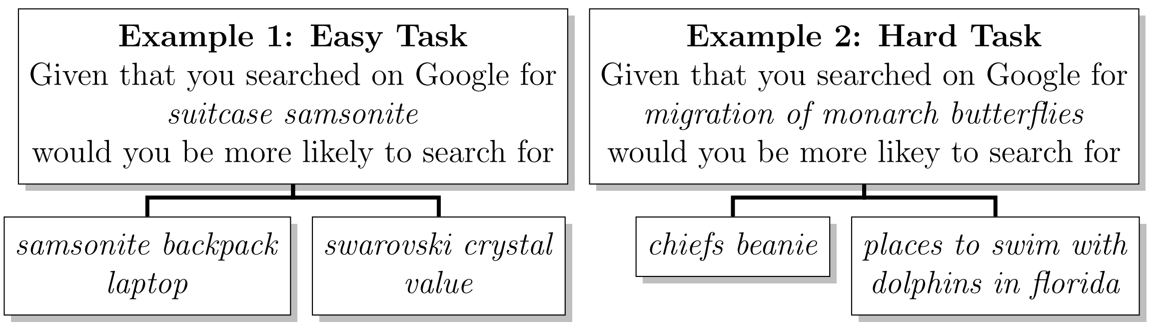
We conclude this section by discussing why starting from demand (i.e., thematic clustering based on a keyword vectorization) is preferable than starting from supply (i.e., matrix factorization with partitioning). The reason is based on the need to develop a method that is robust to keyword splitting strategies (discussed in the main text and further explored in appendix H) that the agency networks might follow. In fact, if the common intermediary splits keywords so that its clients never compete, by starting from the matrix factorization we would tend to assign to different markets advertisers that instead belong to the same competitive space. If, instead, we begin by demand, then we obtain "thematic clusters" that are valid regardless of any keyword splitting strategy by the intermediary. This implies that we can perform a robustness analysis of our findings about the effects of increasing concentration on the search engine revenue: we can look at it—as we do in our baseline estimates—through the "competitive clusters" (that incorporate both demand and supply) or—as we do in one of the robustness checks—through the "thematic clusters" (that incorporate demand only). In our analysis, the fact that under the latter type of clustering we find results that are qualitatively similar to those obtained with the former type of clustering is reassuring that our findings are not distorted by the strategic behaviour of the networks.[9]

[9]By keeping the ordering of the steps as in our study, it would also be possible to integrate additional information in the definition of the clusters that would otherwise get lost. Suppose for instance that the researcher knows that in a certain industry keyword splitting strategies are more frequently used than in another industry. In principle, one could then adapt our procedure by first obtaining the thematic clusters and then setting up the definition of the competitive clusters in a way that accounts for the extra piece of information: for the industry where market splits are deemed more common, we would require less overlap between the advertisers in order to pool the respective keywords within the same competitive cluster relative to what we would do for the other industry where keyword splitting is less common.

D) Cluster Validity

In order to test the reliability of the clustering exercise, we implemented a task to validate them. With no training samples needed, we relied on human intervention only at the very end of the funnel—i.e., we checked the "quality" of the clusters ex-post by designing a series of simple tasks that we submitted to human testers.[10] More specifically, for each cluster $c \in [1, .., 1,000]$, within industry i, we randomly picked a reference keyword $refK_{ci}$ and two test keywords $testK_{ci}$ and $testK_{-ci}$, from c and from one of the other clusters in i, respectively. Figure D.1 is a graphical representation of the task we submitted to the human testers: the user is asked whether, given that she searched for $refK_{ci}$, she would be more likely to search for $testK_{ci}$ or $testK_{-ci}$, or neither of them. The task yields three potential outcomes: i) the user chooses $testK_{ci}$ (*success*), ii) the user chooses $testK_{-ci}$ (*failure*), and iii) the user cannot choose either option (*no answer*).

Figure D.1: Cluster Quality Checking Task



Notes: Amazon Mechanical Turk task representation. First, the user is given a reference keyword belonging to cluster c (*suitcase samsonite* in Example 1) which is supposed to have been searched for on Google. Then, the user is asked to identify out of two additional keywords which of the two is considered more likely to be searched for given the initial search. One of the two keywords proposed belongs to clusters c (*samsonite backpack laptop* in Example 1), while the other belongs to the same industry but to a different cluster (*swarovski crystal value* in Example 1). Example 2 is analogous, but representative of a more difficult case for the tester.

The question is designed to check whether the keyword links emerging from the thematic clustering are effectively mimicking the user behavior when surfing the web. In the figure, example 1 is an "easy task"—from the *Apparel* industry—and had a very high hit rate in the test: the presence of the brand name within both $refK_c$ and $testK_c$ helps to delimit the market (and enhances the similarity, too). Example 2 is relative to the *Travel & Leisure* industry, and experienced a high rate of non-respondence by the testers: the underlying theme linking *migration of monarch butterflies* and *places to swim with dolphins in Florida* is the Florida Keys, which are both one of the destinations of monarch butterfly migrations and a renowned place to swim with dolphins. While this theme is known to real users, it was not identified by most of our human testers, nonetheless GloVe correctly highlighted their similarity. We submitted the tests to *Amazon Mechanical Turk*, a marketplace for work that

[10]This experiment was conducted with the IRB approval for the Amazon Mechanical Turk survey from Bocconi University ECR (SA000267).

requires human intelligence. In table D.1 we report the share of successes, failures and no answers in a sample of industries. Our initial design of the test did not allow the user to skip answers (i.e., *No answer = 0* by design for the first five industries in the table); however, when subsequently we introduced the option we recorded an average of one third of non-responses. The success rate is consistently high and evenly distributed among industries. Moreover, it does not appear to be influenced by the rate of non-response.

The complete dataset as well as a sample of the survey is available as part of the replication files provided. Please cite the data as: Decarolis, Francesco and Gabriele Rovigatti. 2021. "From Mad Men to Maths Men: Concentration and Buyer Power in Online Advertising: Dataset"American Economic Review. `https://www.aeaweb.org/journals/aer`.

Table D.1: *Amazon Mechanical Turk* Test

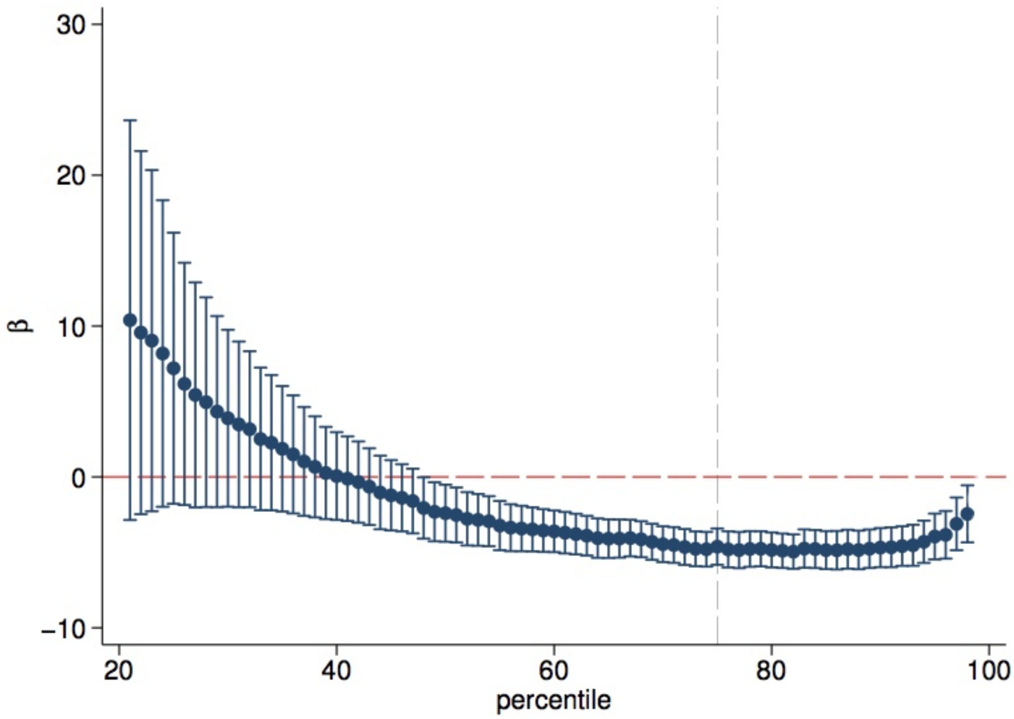
Industry	Answer		No Answer
	Success	Failure	
Technology	.80	.20	0
Travel & Leisure	.85	.15	0
Media	.84	.16	0
Food Processing	.92	.08	0
Miscellaneous	.59	.32	.09
Utilities	.82	.15	.04
Apparel	.78	.15	.06
Retail	.83	.12	.05
Industrial	.84	.11	.05

Notes: clustering test results on a subset of industries. For *Technology*, *Travel & Leisure*, *Media* and *Food Processing* we did not allow the user to leave the question blank.

E) Sample Selection

Following on from the discussion in section 6, we report in Figure E.1 how β^{IV} changes with the dimension of the analysis sample. Indeed, among the competitive clusters, many are composed of keywords that contribute very little, or not at all, to the search engine's revenue and are never involved in any of the M&As that we exploit for the IV strategy. Therefore, we keep in the baseline analysis sample only markets that either experience variation in the instrument at least once during the sample period or, for the remaining ones, those that are in the top quartile of revenue. This leads us to drop markets that represent between 1% and 2% of the total yearly revenue. In Figure E.1, the baseline sample—with the corresponding IV estimate—is marked by the vertical dashed line. As this figure illustrates, after we drop less than 50% of the lowest revenue markets, the IV estimates become fairly constant and similar to the baseline ones. Dropping 50% (or less) of the lowest revenue markets corresponds to dropping less than 1% of the total yearly revenue. Thus, for the purpose of our analysis, we consider these small markets not to be a valuable source of variation in the data, but rather to be a source of noise that makes it impossible to detect the causal association between demand concentration and revenue. This is especially the case because these zero (or nearly so) revenue markets are often very small, possibly made up of one or very few keywords and, crucially, with a single advertiser bidding on them.

Figure E.1: Effects of Sample Selection on the IV Estimates



Notes: points estimates (blue dots) and their confidence intervals (blue caps) on samples of different sizes. The dotted grey line at the 75^{th} percentile marks the sample used in the baseline analysis.

F) Robustness Checks

In this section, we analyze the robustness of the baseline estimates presented in the text to several modifications. First, to ensure the reasonableness of the IV approach, we repeat the analysis looking exclusively at the largest mergers. We perform this analysis separately for each one of the four largest mergers, involving four different networks. The results reported in Table F.1 are broadly consistent with the baseline estimates presented in the main text. The top panel reports reduced form and first stage estimates, while the bottom panel reports OLS and IV estimates. In all cases the model specification is that of the baseline estimates (model (9) in the previous table). For the mergers involving Sapient, Merkle and Forsman & Bodenfors, both the significance and the magnitude of the estimates track closely what is reported in Table **??** (although the IV estimates are smaller for the Forsman & Bodenfors merger). For the Shift merger, however, the reduced form is not statistically significant. Thus, while the OLS estimates are in line with those of the other mergers, this is the only IV estimate that is not significant. Possibly this is because WPP never fully integrated Shift into its systems as this company entered the WPP network indirectly through an acquisition by a large Canadian affiliate of WPP, National Public Relations, that maintained Shift as its agency for its US clients. Despite some heterogeneity across the cases, the overall takeaway is that, even narrowing down the analysis to the subset of the data where the IV strategy is the most reasonable, the results are qualitatively close to those of the baseline estimates.

Table F.1: Individual Mergers

Panel a): Individual Mergers – Reduced Forms and First Stages

	Sapient		Merkle		Shift		Forsman & Bodenfors	
	RF	FS	RF	FS	RF	FS	RF	FS
sim$\Delta\hat{HHI}$	-4.911*	1.026***	-5.981***	1.388***	4.536	0.707***	-16.30**	6.357***
	(2.882)	(0.387)	(1.181)	(0.0386)	(2.998)	(0.230)	(6.388)	(0.159)
Observations	4,776	4,776	3,047	3,047	3,013	3,013	981	981

Panel b): Individual Mergers – OLS and IV Estimates

	Sapient		Merkle		Shift		Forsman & Bodenfors	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
$\hat{HHI}$	-5.302***	-4.786*	-4.516***	-4.308***	-3.823***	6.415	-5.236***	-2.563**
	(0.208)	(2.547)	(0.293)	(0.871)	(0.175)	(4.963)	(0.672)	(0.999)
Observations	4,776	4,776	3,047	3,047	3,013	3,013	981	981
Industry FE	✓		✓		✓		✓	
Year FE	✓		✓		✓		✓	
Organic Results	✓		✓		✓		✓	

Notes: the dependent variable is the (log) revenue, R_{mt}. For each reported M&A action (*Sapient, Merkle, Shift and Forsman & Bodenfors*), the estimation sample amounts to all markets involved, i.e., all markets in which at least one of an agency's clients was bidding before the merger. In panel a) odd columns report the reduced form and even columns the first stage estimates, respectively. In panel b), odd columns report the OLS and even columns the IV estimates. All models feature controls for the average number of organic results, industry and year fixed effects, and the standard errors are clustered at the thematic clusters level.

We consider next five sets of robustness checks presented in Table F.2. All estimates reported in this table are the IV estimates of the baseline model specification. In the first two columns, we explore the effects of using alternative definitions of "markets." In column (1), markets are defined as the industries of the advertisers. Earlier we discussed why this is likely to be problematic, as industries are an excessively broad category and, indeed, the estimates in column 1 indicate a very unreasonable IV estimate. In the following column, we thus return to a definition of market based on the 2-layer keyword clustering procedure, but we use as markets the thematic clusters. The qualitative insight of a negative and significant β is maintained, but the magnitude is substantially larger, which is reassuring with regards to the fact that our baseline is a conservative estimate of the true effect. The following three columns explore the robustness of the estimates to the details of the proposed 2-layer approach. In column (3), instead of using the term-by-term sums of GloVe vectors, the thematic clusters are built by averaging GloVe vectors within keywords. Intuitively, averaging the vectors attenuates the effects of "topical" terms, whose weight is instead amplified by the sum; moreover, the latter method tends to isolate long tail keywords—keywords with more terms face a higher likelihood of being positioned "far away" in the vector space. As a result, the averaged GloVe keywords are less sparse, and possibly harder to cluster. Despite this, the estimates are very close to the baseline ones. The next two columns, (4) and (5), explore related modifications of the clustering approach involving the number of centroids of the k-means algorithm, using either 500 centroids or the number of keywords in the industry divided by 30. Again, the baseline estimates appear robust to these modifications.

Table F.2: Robustness Checks

	Market Definition		Two-layers Clustering		
	Industry Level	Thematic Clusters	GloVe mean	500K	N/30K
	(1)	(2)	(3)	(4)	(5)
HHI	9063.3	-3.353**	-4.537***	-5.493***	-2.820*
	(1528504.4)	(1.586)	(1.360)	(1.151)	(1.538)
Observations	68	16,959	52,237	41,966	40,572
Industry FE	✓				
Cluster FE		✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓

Notes: the dependent variable is the (log) revenue, R_{mt}. The definition of m changes across models. In column (1), we do not perform any clustering exercise, and m is the industry level (there are up to 23 industries per year). In column (2), m is the thematic clusters level. In columns (3) to (5) m is a competitive cluster, but the clustering algorithm used is not the same as in the baseline estimates. In column (3), we average over GloVe-vectorized terms—instead of summing up the vectors—before performing the step-1 clustering exercise, column (4) features 500 clusters *per industry* in step-1, while in column (5) we repeat the exercise with a size-dependent number of clusters, i.e., with 1 cluster for every 30 unique keywords in the sample. All models feature controls for the average number of organic results, industry—(1)—or thematic clusters and year fixed effects. Standard errors are clustered at the industry or thematic clusters level.

The last set of robustness checks involve the CTR. The CTR measure that we use presents a measurement error problem, as discussed both in the text and in section A of this appendix. In this section, we explore the robustness of our baseline estimates to this problem. In particular, we consider two sets of robustness checks: first, we exclude the CTR from the analysis by setting all CTRs to 1 and, second, we randomly re-match CTRs to keywords. The first exercise consists of estimating the same regression models presented in Table **??** using modified versions of the main variables: in the case of Table F.3, the CTRs are set to 1 only for the dependent variable, while in the case of Table F.4, they are set to 1 for all variables whose calculation involves the CTR. To distinguish these modified variables from those used earlier, we indicate the former with an upper bar: $\bar{R}_{mt}$ is thus R_{mt} recalculated without CTRs.[11] The reason why it is interesting to present the two sets of estimates in Table F.3 and F.4 is that estimating the effect of HHI_{mt} on $\bar{R}_{mt}$ can also serve as a check of the robustness of our analysis to an alternative measure of the revenue: an upper bound on the revenue attainable when all ads generate the same number of clicks per time. In any case, the estimates in both Table F.3 and F.4 are quite close to each other and also close to the estimates in Table **??** in the text, although systematically smaller. For instance, relative to our benchmark estimated effect of an 11.32 percent drop in revenue, the corresponding estimate in Table F.3 indicates a drop of 8.54 percent and that in Table F.4 a drop of 8.77 percent. Thus, the qualitative implications of our analysis are robust to this type of alternative use of CTRs.

In the second set of robustness checks involving the CTR, we consider a different approach

[11] $\bar{R}_{mt} = \sum_{k \in K_m} CPC_{kmt} * Volume_{kmt}$, $\bar{s}^i_{mt} = \frac{1}{\bar{S}_{mt}} \sum_{a \in A_i} \sum_{k \in K_m} Volume_{kt}$, and $H\bar{H}I_{mt} = \sum_{i=1}^{I} (\bar{s}^i{}_{mt})^2$.

aimed at assessing how the variation of our CTR measure across markets might impact our findings. We proceed by setting up a bootstrap procedure that, at the beginning of each repetition, randomly assigns to each keyword a vector of industry-year CTRs (i.e., the CTRs of positions 1 to 11 for the specific industry-year, from AWR data) that is drawn (with replacement) from the whole set of industry-year CTR vectors in the data. Then we calculate the baseline estimate (corresponding to the model of column 9 of Table **??**). Figure F.1 reports the IV estimates obtained on 500 samples with the block-bootstrapped CTR data. The figure reports each repetition on the x-axis. On the y-axis, it reports the estimates: the point estimate (red solid square), and the 95% confidence interval (blue spikes). The dashed white line marks the baseline $\hat{\beta}^{IV}$ (from column 9 in Table **??**), whereas the dashed grey line, $\hat{\beta}^{boot}$, reports the average bootstrapped $\hat{\beta}^{IV}$. Although there is variability in the estimates across the 500 repetitions, all point estimates are close to the baseline estimate. The average estimate across the samples, $\hat{\beta}^{boot}$, is in fact very close to $\hat{\beta}^{IV}$. Furthermore, $\hat{\beta}^{IV}$ always falls within the 95% confidence interval of the bootstrapped parameter, while zero (or positive values) are never contained. Therefore, the results in Figure F.1 confirm the robustness of the main estimates in the text.

Table F.3: Effect of Concentration on Search Engine Revenues - $log(\bar{R})$

	OLS					IV				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\hat{HHI}$	-2.314***	-2.278***	-2.228***	-2.220***	-2.215***	-12.07***	-5.012***	-3.495***	-3.483***	-3.456***
	(0.0646)	(0.0529)	(0.0525)	(0.0524)	(0.0525)	(3.682)	(0.846)	(1.075)	(1.079)	(1.079)
Organic Results (billion)				0.261***	0.252***				0.238***	0.232***
				(0.0582)	(0.0564)				(0.0595)	(0.0577)
Keywords Characteristics										
Branded Keyword					-0.0406					-0.00740
					(0.0534)					(0.0608)
Long-tail Keywords					-0.121***					-0.0993**
					(0.0357)					(0.0403)
Observations	52,476	52,476	52,476	52,476	52,476	52,476	52,476	52,476	52,476	52,476
Cluster FE		✓	✓	✓	✓		✓	✓	✓	✓
Year FE			✓	✓	✓			✓	✓	✓

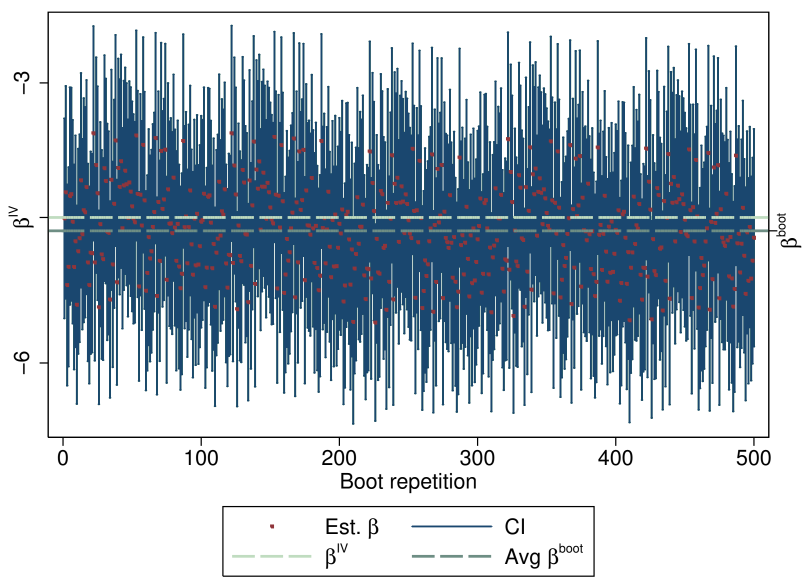
Notes: the dependent variable is the (log) revenue, $\bar{R}_{mt}$. Columns (1) to (5): OLS estimates, with an increasing number of fixed effects and controls. Columns (6) to (10): IV estimates, where we instrumented HHI_{mt} with the merger-induced change in concentration. In all models the standard errors are clustered at the thematic clusters level.

Table F.4: Effect of Concentration on Search Engine Revenues - $log(\bar{R})$ on $H\bar{H}I$

	OLS					IV				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$H\bar{H}I$	-2.347***	-2.252***	-2.188***	-2.181***	-2.176***	-11.45***	-4.931***	-3.592***	-3.580***	-3.544***
	(0.0648)	(0.0536)	(0.0533)	(0.0533)	(0.0533)	(3.303)	(0.825)	(1.081)	(1.084)	(1.081)
Organic Results (billion)				0.262***	0.253***				0.236***	0.232***
				(0.0577)	(0.0561)				(0.0586)	(0.0570)
Keywords Characteristics										
Branded Keyword					-0.0203					0.0296
					(0.0533)					(0.0662)
Long-tail Keywords					-0.119***					-0.0921**
					(0.0357)					(0.0412)
Observations	52,476	52,476	52,476	52,476	52,476	52,476	52,476	52,476	52,476	52,476
Cluster FE		✓	✓	✓	✓		✓	✓	✓	✓
Year FE			✓	✓	✓			✓	✓	✓

Notes: the dependent variable is the (log) revenue, $\bar{R}_{mt}$. Columns (1) to (5): OLS estimates, with an increasing number of fixed effects and controls. Columns (6) to (10): IV estimates, where we instrumented $H\bar{H}I_{mt}$ with the merger-induced change in concentration. In all models the standard errors are clustered at the thematic clusters level.

Figure F.1: CTR Bootstrap repetitions



Notes: IV estimates obtained by estimating the baseline IV model (column 9 in Table **??**) on 500 samples with block-bootstrapped CTR data. For each industry-year, we draw (with replacement) the distribution of CTR - positions 1 to 11 - from AWR data, then randomly merge them to the SEMrush data before aggregating at the market level, and run the estimation. For each repetition, reported on the x axis, we plot the point estimate (red solid square), and the 95% confidence interval (blue spikes). The dashed white line marks the baseline $\hat{\beta}^{IV}$, whereas the dashed grey line reports the average bootstrapped $\hat{\beta}^{IV}$.

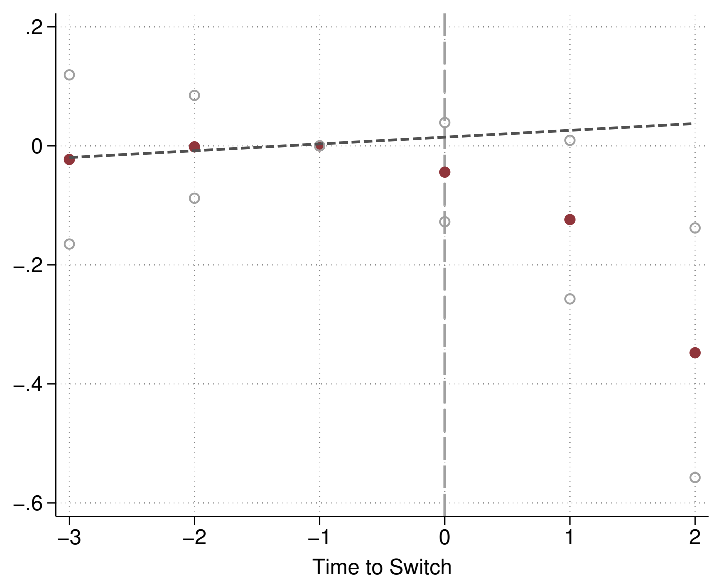
G) Falsification

In a setting like the one analyzed, it seems useful to visualize changes in the outcome variable before and after an acquisition-driven change in concentration. In figure G.1, we report a graph built as in Dobkin et al. [2018].[12] Specifically, in order to show the impact of the mergers on total revenue, we first build indicator variables for time relative to the event at the market level (i.e., time from the first M&A which involved any MAs in the competitive cluster), then we estimate a nonparametric event study of the form:

$$log(R_{mt}) = \alpha + X_{mt}\gamma + \sum_{r=-3}^{-2} \mu_r + \sum_{r=0}^{2} \mu_r + \varepsilon_{mt}$$

where X_{mt} are market-level controls and μ_r are the coefficients on the relative time indicators (i.e., the key coefficients plotted in the figure, alongside their pre-merger linear trend, the dotted line). The vertical, dashed grey line indicates the first year after the merger. The upward sloping, dashed black line is the linear fit in the pre-merger period (as the figure suggests, the fit approximates these data quite well). The full dots are the period averages, while the hollow dots indicate the standard errors. There is rather clear graphical evidence: the drop in the average revenue post-merger indicates a negative association between the post-merger period and the log revenue, which is consistent with the estimates in the paper.

Figure G.1: Impact of Mergers on $log(R)$



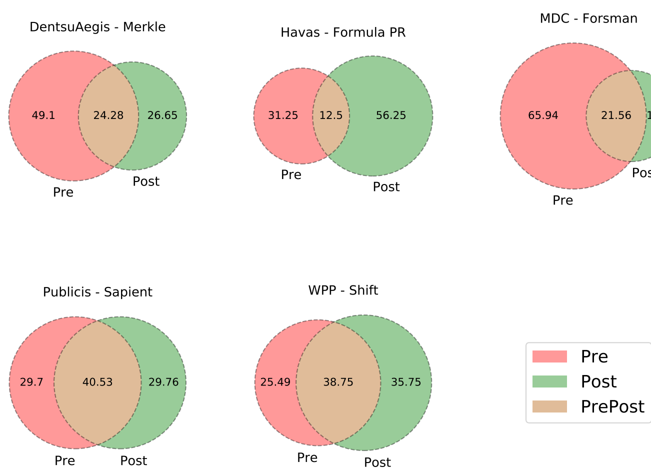
Notes: full dots are the averages of the time indicator point estimates in a nonparametric event study estimation of $log(R)$ on M&A events, while the hollow dots indicate their confidence intervals. The upward sloping, dashed back line is the linear fit of the pre-merger years, projected on the post-merger period.

[12] https://pubs.aeaweb.org/doi/pdfplus/10.1257/aer.20161038

H) Mechanisms: Segmentation by Keywords and by Branded Keywords

In this section, we present additional material regarding the issue of the mechanisms through which concentration of intermediaries lowers the CPC. In particular, we explore two aspects related to market segmentation via the division of keywords. The first result that we present is the representation in Figure H.1 and complements the keyword-level descriptive evidence presented in section IV. As explained there, for six large mergers involving different networks, we look at whether, after the acquisition of an agency by a network, there is any change in the overlap in the sets of keywords of the clients of either the network or the acquired MA. If the overlap declines, it might indicate that the intermediary splits the market by keywords, while if it stays identical (or grows) it might indicate that most of what the intermediary does takes place within-auctions. As discussed in the text, the evidence in Figure H.1 suggests that both strategies are adopted, although to different extents across the seven networks.

Figure H.1: Venn Diagram: all mergers



Notes: share of coalition keywords—i.e., keywords bid by both the advertisers in the acquired agency and those in the acquiring network—before and after the merger. Shares are computed on the overall number of coalition keywords. "Pre" is the share of keywords in coalition in the year before the merger only; similarly, "Post" refers to the share of keywords in coalition only in the year after the merger, and "PrePost" are keywords in coalition both before and after.

The second result is more specific and concerns branded keywords. Advertisers spend significant portions of their marketing budgets on branded keywords: these are related to both their own brand and to the brand of their rivals'. Among the feasible coordination strategies, keyword splitting represents the easiest way to fully segment the market. Explicit coordination by advertisers to stop bidding on each others brands, though, is unlawful. But the same bidding pattern would be legal if autonomously implemented by a network representing rival advertisers. Hence, from an advertiser's viewpoint, coordination through network intermediaries might be two-fold optimal: on the one hand, it lowers keyword-level costs by decreasing price competition in the auctions; on the other hand, it guarantees lower marketing costs by preventing brand competition.

Regarding branded keywords, we first offer additional details regarding the construction of Figure **??** in the main text, then we present a numerical example. Recall that in Figure **??** we assess the change in probability for both the other branded and own branded. Regarding this figure, we formally define the brand bidding, and we build a *branded* indicator variable for all keywords that contain one or more words related to a brand (e.g., the keyword "Volkswagen beetle" would be branded, given that it contains the brand "Volkswagen"). We also define a few additional variables at different aggregation levels:
1. Keyword-auction level. Within the *branded* keywords we further distinguish two subdomains, depending on an advertiser's identity: *own_branded* is an indicator for advertisers bidding on keywords related to their own brand ("Volkswagen beetle" when the advertiser is Volkswagen); *other_branded* indicates whether an advertiser bids on a keyword whose related brand is not its own ("Volkswagen beetle" when the bidder is Ford Motor Company);
2. Market-year level. For each agency-year pair (j, t) we define the variable *timetoswitch* as the distance—in years—to the relative M&A event (t^*). When aggregating at the market level, we use the *first* recorded event as the reference point in the definition of *timetoswitch*. We also define the indicators for the presence of branded, own branded and other branded at the market/year level (*dbranded*, *dother_branded* and *down_branded*, respectively);
3. Time-to-switch level. We aggregate the probability of being branded (total, own and other) at the timetoswitch level. We also compute the yearly change in probability as $\Delta P(branded)_t = \frac{(shareBranded_t - shareBranded_{t-1})}{shareBranded_{t-1}}$. Finally, in order to ensure the comparability of all measures, we de-meaned them. Finally, provided with these variables, we apply the method by Dobkin et al. [2018] to produce the outcomes reported in Figure **??**.

We conclude this section with a numerical example showing through a case of keyword segmentation the reason why even small increases in HHI can cause large drops in revenue. Indeed, crucial to the understanding of the magnitudes in the main text is knowing how the GSP auction system works when bids are coordinated. Suppose that there is a market composed of 2 keywords, $k1$ and $k2$. Both keywords have the same number of available slots, the same set of advertisers $(a_1, ..., a_5)$ and the same CTRs associated with the different slots, which are equal to 20 clicks for the top position, 10 clicks for the second, 5 clicks for the third, 2 clicks for the fourth and 0 clicks for the fifth. The only element along which $k1$ and $k2$ differ is that advertiser a_3 values 3 dollars a click on keyword $k1$ and 2 dollars a click on keyword $k2$. The reverse is true for advertiser a_4. This situation is illustrated in the first three columns of the Table H.1.

Under the EOS-equilibrium characterization typically used in the literature (i.e., the envy-free Nash equilibrium of Edelman, Ostrovsky and Schwarz [2007], Varian [2007]), the bids for $k1$ would be those reported in the fourth column and the total revenue for the search engine would be equal to 96 dollars. Everything is identical for $k2$ except that the position and payments of a_3 and a_4 are flipped relative to $k1$. Assume now that there is a merger of the intermediaries bidding on behalf of a_3 and a_4 and that the resulting intermediary decides to have a_3 exiting $k2$ and a_4 exiting $k1$. This could be, for instance, the case of $k1$ being

Table H.1: Example: 2-Bidder Merger in a 2-Keyword Market

Advertiser	Valuations $k1$	Valuations $k2$	Bids for $k1$	Bids for $k1$ post merger	Market shares	Market shares post merger
a_1	5	5	b_1	b_1	0.541	0.541
a_2	4	4	3.15	2.90	0.270	0.270
a_3	3	2	2.30	1.80	0.095	0.068
a_4	2	3	1.60	–	0.095	0.068
a_5	1	1	1.00	0.60	0.000	0.054
			Tot.Rev=96	Tot.Rev.=79	HHI=3,831	HHI=3,864

Notes: the market has 2 keywords ($k1$ and $k2$) and 5 advertisers ($a_1, ..., a_5$). The table reports for each of the five advertisers (column 1), their valuations for $k1$ (column 2) and for $k2$ (column 3), bids for $k1$ both before the merger (column 4) and after it (column 5). The bids are identical for $k2$ but with the order of a_3 and a_4 switched. The last two columns report market shares, i.e. the share of clicks associated with the slots occupied relative to the total number of clicks in the market. The last row reports total revenue and the market-level HHI. The merger is between the intermediaries bidding for a_3 and a_4. Post merger a_3 is assumed to exit keyword $k2$ and a_4 is assumed to exit keyword $k1$.

a branded keyword of a_3 and $k2$ being a branded keyword of a_4. Under this scenario, the new EOS-equilibrium bids would be those reported in the fifth column of the table. Not surprisingly, the search engine revenue drop after the merger: from 96 dollars to 79 dollars (for each keyword), an 18 percent drop.[13] What is remarkable in this example is how small the HHI increase is: a mere 33-point increase.

The reason why the HHI change is so limited is that the advertisers involved in the merger occupy slots that are worth few clicks. These slots correspond to a small market share. However, since within the GSP auction all bids are interlinked in equilibrium, even bid changes by bidders occupying slots far from the top one can trigger a chain reaction of bid changes causing large shifts in revenue. Stated differently, bid changes are not limited to the advertisers directly involved in the merger (*direct effect*), but also those advertisers placed above them (*indirect effect*). Notice, for instance, the drop in a_2's bid after the merger: from 3.15 dollars to 2.90 dollars, despite this advertiser not being directly affected by the concentration.[14]

I) Industry Heterogeneity

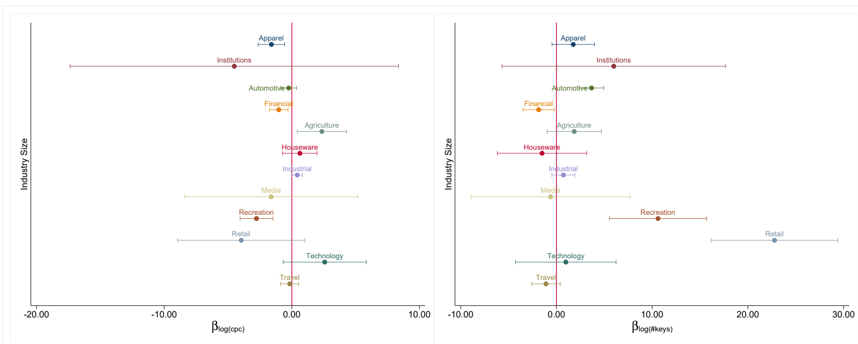
In Figure I.1, we explore differences among industries by showing the distribution of $\hat{\beta}_{IV}$

[13] On the contrary, all advertiser payoffs increase. For both a_3 and a_4 the payoff goes from 9 dollars to 12 dollars. We shall also remark that for the post merger scenario we assume there is a sixth advertiser with a valuation (and a bid) of zero.

[14] For the case of bid coordination within a single keyword auction, Decarolis, Goldmanis and Penta [2020] formalize this logic of *direct* vs *indirect* bid reduction effects. They also prove why a strategy proof mechanism, like the VCG, would limit the revenue loss by preventing the revenue loss from the *indirect effect*. Notice that for the example in Table H.1 we are not resorting to the equilibrium characterization of Decarolis, Goldmanis and Penta [2020], but to the standard notion of EOS-equilibrium.

estimated at the industry level, for $log(cpc)$ (left panel) and $log(\#keywords)$ (right panel). Although negative on average, the former features positive values for one sector, *Agriculture*. The estimated effect of concentration on changes in the number of keywords shows a higher degree of noise, with most industries characterized by an imprecisely estimated zero effect. A positive impact, however, is clear for three important industries, *Automotive*, *Recreation* and *Retail*. The resulting picture suggests that networks, and MAs, follow different strategies depending on the market structure and competitive pressures within industries. The overall effect on revenue hence emerges from multiple, different paths.

Figure I.1: Industry-level IV estimates distribution



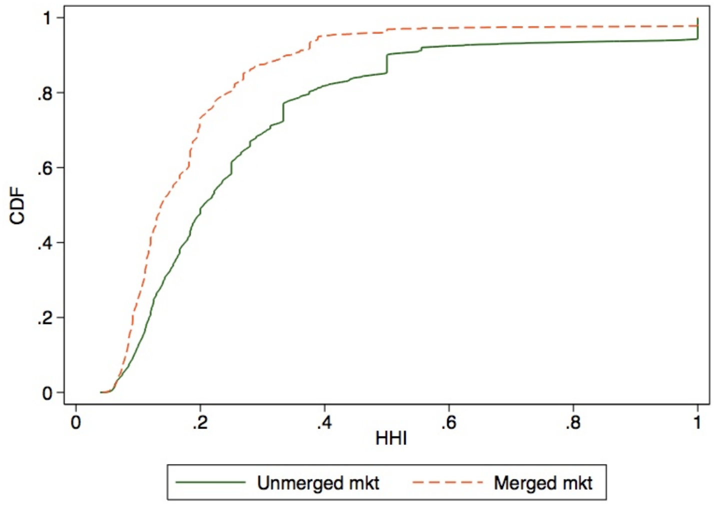
Notes: Industry-level IV estimates of $\hat{\beta}_{IV}$ with different outcomes: $log(cpc)$ (left panel) and $log(number\ of\ keywords)$ (right panel). Industries are ranked according to their size in terms of total search volume, and each point estimate is reported alongside its standard error. Industries with estimates 30 times bigger than the average point estimate have been excluded in order to ensure plot readability.

J) Monotonicity Test of the Instrument

In this section, we report the results of the instrument's monotonicity test proposed by Angrist and Imbens [1995]. Verifying that monotonicity holds is important because the sign of the first stage regression is theoretically unclear and, also, because splitting the market may create a negative relationship between HHI and simulated HHI over some of the latter's range. In fact, by instrumenting the HHI ($S^{\tilde{Z}}_{mt}$, $\tilde{Z} = [0, 1]$) with the merger-induced change in HHI (Z_{mt}), we are implicitly assuming that the merger effect is monotone—that is, either $S^1_{mt} \leq S^0_{mt}$ or $S^0_{mt} \geq S^1_{mt}$, $\forall m, t$. The assumption is not verifiable, but has testable implications on the CDFs of HHI for merged ($\tilde{Z}_{mt} = 1$) and unmerged markets ($\tilde{Z}_{mt} = 0$)—that is, they should never cross. In fact, if $S^1_{mt} \geq S^0_{mt}$ with probability 1, then $Pr(S^1_{mt} \geq j) \geq Pr(S^0_{mt} \geq j), \forall j \in \text{supp } S$. Figure J.1 plots the CDFs of markets subject to a merger (dashed red line) and not subject to any merger (solid green line). Since the two CDFs never cross, the instrument passes the test.

K) Competition among Intermediaries and Network Common Ownership

Figure J.1: Instrument Monotonicity Test

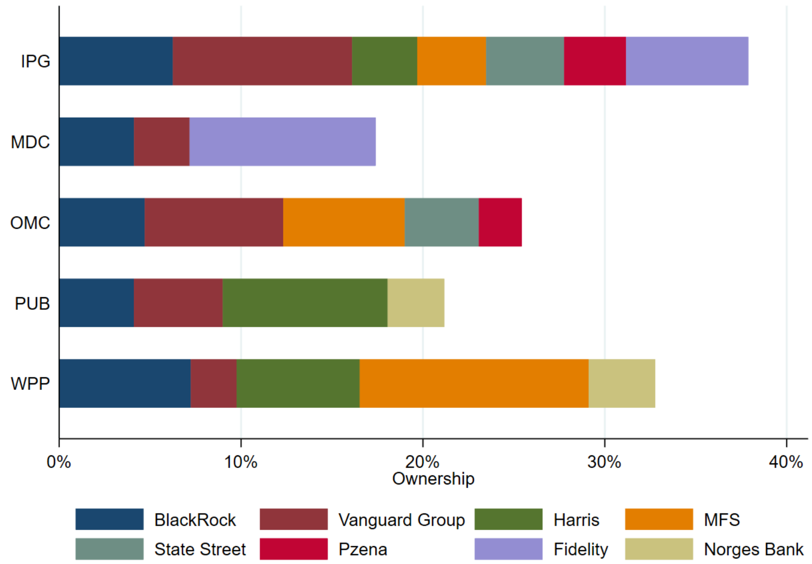


Notes: Instrument Monotonicity Test (Angrist and Imbens [1995]). By instrumenting the HHI ($S^{\tilde{Z}}_{mt}$, $\tilde{Z} = [0, 1]$) with the merger-induced change in HHI (Z_{mt}), we are implicitly assuming that the merger effect is monotone—that is, either $S^1_{mt} \leq S^0_{mt}$ or $S^0_{mt} \geq S^1_{mt}$, $\forall m, t$. The assumption is not verifiable, but has testable implications on the CDFs of HHI for merged ($\tilde{Z}_{mt} = 1$) and unmerged markets ($\tilde{Z}_{mt} = 0$)—that is, they should never cross. In fact, if $S^1_{mt} \geq S^0_{mt}$ with probability 1, then $Pr(S^1_{mt} \geq j) \geq Pr(S^0_{mt} \geq j), \forall j \in \text{supp } S$. The plot reports the CDFs of markets subject (dashed red line) and not subject to any merger (solid green line): indeed, they never cross.

In this section, we present a series of caveats to our discussion in the final section of the papers about the intermediary sector being reasonably competitive. In particular, there are at least five features that limit the extent of competition among intermediaries and that would deserve further research. First, intermediary competition is limited by the well-known difficulties in measuring the returns to advertising, which Lewis and Rao [2015] indicate to be severe for online advertising. Second, a closely related feature regards the lack of transparency in intermediary reports to advertisers about how their budget is spent [ISBA, 2020]. While the former issue relates to an intrinsic difficulty in advertising, the second relates to the behavior of intermediaries who typically report very aggregated measures of how they allocated client money. This contributes to explaining why advertisers may fail to optimize their bidding campaigns; something that powerfully emerges from Blake, Nosko and Tadelis [2015]. Third, the exact same features mentioned in our study for why bid coordination by a common intermediary can be valuable, all imply that advertisers might become locked in. This is because obtaining the same benefits would require a joint deviation by competing advertisers, from their current intermediary toward a different one. Fourth, some industry observers suggest even more complex forms of collaboration with Google having found ways to cooperate with the intermediaries in order to ensure its long run dominance, even at a short term cost. In recent years, there have been multiple revelations about agency kickbacks. An investigation by the US Association of National Advertisers, [ANA, 2016], states that "numerous non-transparent business practices, including cash rebates to media agencies, were found to be pervasive in the US." Nevertheless, some of the networks, like WPP, have responded by saying that they did not take part in the Google's US media rebate program, while others, like Omnicon, admitted to being part of it but argued that the rebate was passed down to clients. Overall, monitoring these five areas of concern would be an essential component of a policy intervention that seeks to make good use of advertising intermediaries as a remedy to the dominance of the largest online platforms. Fifth, a collusive conduct between some of seven agency networks might be aided by some features, like their common ownership. Following Azar, Schmalz and Tecu [2018], we look for the presence of owners that are in common between the 5 publicly listed networks. In Figure K.1, for each of these 5 networks, we report the average ownership share in the 2010-2019 period for owners that, for at least 2 of the 5 networks, are among the 10 largest shareholders. Black Rock and Vanguard are among the top 10 shareholders for each of the 5 networks. The other investors are among the top 10 shareholders of 2 or 3 networks. It is important, however, not to overstate the significance of this evidence on common ownership.

Overall, as stressed in the main text, there are conflicting views on the extent of competition in the US advertising and marketing services agency industry. Indeed, while we offered above five reasons why competition might be limited, the academic consensus is, however, that the industry is reasonably competitive. In the text, we referenced Silk and King [2013], which is a landmark study on concentration in this industry. It reports a set of concentration measures for the various sectors of the advertising and marketing services industry (Tables 2, 4 and 5) along with additional measures that apply to the holding companies/networks, whose dominance has long been overstated (Table 6). In an earlier study (Silk and Berndt, 1994), evidence is presented that the industry's diversity and low level of concentration were consistent with the MacDonald and Slivinski (1987) theory of the equilibrium structure of a competitive industry with multiproduct firms. King, Silk and Ketttelhohn (2003) investigated knowledge spillovers and externalities in the deagglomeration and growth of the advertising agency business. They found that a simple model of high demand, low wages, and externalities associated with clusters of related industries explained the dispersion of agency employment across states. Arzaghi, Berndt, Davis and Silk (2012) summarize a considerable body of stylized facts consistent with the market for advertising campaigns being contestable in the sense of Baumol et al. (1988).

Figure K.1: Common Ownership



Notes: for the 5 networks that are publicly traded, the figure reports the average ownership share in the 2010-2019 period for owners that, for at least 2 of the 5 networks, are among the 10 largest shareholders. Black Rock and Vanguard are among the top 10 shareholders for each of the 5 networks. The other investors are among the top 10 shareholders for 2 or 3 networks. The data source is the Eikon dataset, `https://www.refinitiv.com/en/products/eikon-trading-software`.

References

ANA. 2016. "Media Transparency Initiative: K2 Report." *K2 report for the Association of National Advertisers*, June.

Angrist, Joshua D, and Guido W Imbens. 1995. "Two-stage least squares estimation of average causal effects in models with variable treatment intensity." *Journal of the American statistical Association*, 90(430): 431–442.

Azar, Josa, Martin C Schmalz, and Isabel Tecu. 2018. "Anticompetitive Effects of Common Ownership." *The Journal of Finance*, 73(4): 1513–1565.

Bird, Steven, Ewan Klein, and Edward Loper. 2009. *Natural language processing with Python: analyzing text with the natural language toolkit.* " O'Reilly Media, Inc.".

Blake, Thomas, Chris Nosko, and Steven Tadelis. 2015. "Consumer Heterogeneity and Paid Search Effectiveness: A Large-Scale Field Experiment." *Econometrica*, 83(1): 155–174.

Dai, Weija. 2014. "Matching with Conflicts: An Application to the Advertising Industry." mimeo.

Decarolis, Francesco, Maris Goldmanis, and Antonio Penta. 2020. "Marketing Agencies and Collusive Bidding in Online Ad Auctions." *Management Science*, 66(10): 4359–4919.

Dobkin, Carlos, Amy Finkelstein, Raymond Kluender, and Matthew J Notowidigdo. 2018. "The economic consequences of hospital admissions." *American Economic Review*, 108(2): 308–52.

Edelman, Benjamin, Michael Ostrovsky, and Michael Schwarz. 2007. "Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords." *American Economic Review*, 97(1): 242–259.

ISBA. 2020. "Programmatic Supply Chain Transparency Study." *PwC Report for the Incorporated Society of British Advertisers.*

Kelley, Lawrence A, Stephen P Gardner, and Michael J Sutcliffe. 1996. "An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies." *Protein Engineering, Design and Selection*, 9(11): 1063–1065.

Lewis, R. A., and J. M. Rao. 2015. "The Unfavorable Economics of Measuring the Returns to Advertising." *Quarterly Journal of Economics.*

Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. 2011. "Scikit-learn: Machine learning in Python." *Journal of machine learning research*, 12(Oct): 2825–2830.

Pennington, Jeffrey, Richard Socher, and Christopher D. Manning. 2014. "GloVe: Global Vectors for Word Representation." 1532–1543.

Silk, Alvin J., and Charles III King. 2013. "How Concentrated Is the U.S. Advertising and Marketing Services Industry? Myth vs. Reality." *Journal of Current Issues & Research in Advertising*, 34.

Varian, Hal. 2007. "Position auctions." *International Journal of Industrial Organization*, 25(6): 1163–1178.