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

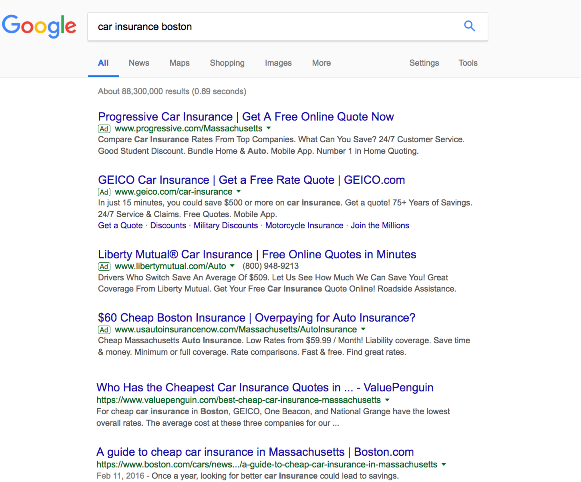
Francesco Decarolis Gabriele Rovigatti

(Bocconi University) (Bank of Italy)

BECCLE Conference
Bergen, 26/4/2019

Internet Advertising and Sponsored Search

- Internet advertising revenues in US: $88 billion dollars in 2017
- Sponsored search: main segment, 46% (next is banner 31%)



Sponsored Search and Marketing Agencies

Highly **concentrated supply**: Google's revenues range between 75% and 80% of total

Traditional view of the other players in sponsored search:

1) **Consumers:**
    - Search for products/services: known or new (learning)
    - Shop for product/services: ubiquitous online buy options
2) **Advertisers:**
    - Seek attention of relevant consumers: targeting
    - Have complex, sometimes conflictual interactions with search engines

Sponsored Search and Marketing Agencies

Highly **concentrated supply**: Google's revenues range between 75% and 80% of total

Traditional view of the other players in sponsored search:

1) **Consumers:**
   - Search for products/services: known or new (learning)
   - Shop for product/services: ubiquitous online buy options
2) **Advertisers:**
   - Seek attention of relevant consumers: targeting
   - Have complex, sometimes conflictual interactions with search engines
3) **Intermediaries - Digital Marketing Agencies (DMAs):**
   - Modern version of the traditional "Madison Avenue" agencies
   - At least since 2011, delegation of bidding to DMAs, who further delegate to their agency network's centralized Agency Trading Desk (ATD)

Demand Side Dynamics

Motivation and Findings

Intermediaries can significantly impact the marketplaces with effects that are both positive (more bidders/keywords) and negative (coordinated bids) for search engines' revenues

We use new, extensive data on both keyword bidding (40 million keyword **auctions**) and links advertisers-DMAs-ATDs (all DMAs and ATDs of 6,000 large advertisers) to quantify how increases in intermediaries' concentration affect Google's sponsored search revenues

Using an IV strategy, we find a sizeable, negative relationship between Google's revenues and buyers' HHI (1 s.d. increase in HHI or a 0.31 HHI increase in a zero to 1 scale, leads to 2% decrease in Google's revenues)

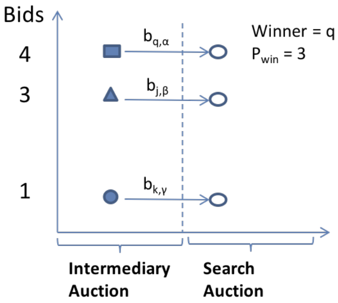
Implies that countervailing power can play a key role in disciplining market power in online platform markets and suggests that competition policy should monitor two aspects:

1. price pass-through to advertisers/consumers (algorithmic collusion, but beneficial?)
2. potential abuses in Google's response (increased reserve price; disintermediation; else?)

Theoretical Example

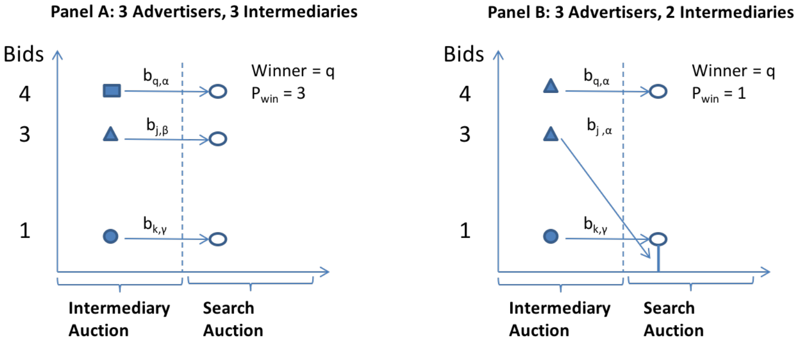
- Suppose there is a monopolist search engine selling 1 ad slot
- There are three advertisers (q, j, k) interested in the slot
- They have arbitrary bids: $b_q = 4$, $b_j = 3$ and $b_k = 1$
- They must bid through an intermediary (α, β or γ)
- 2-level Second Price Auction system

Panel A: 3 Advertisers, 3 Intermediaries



Theoretical Example

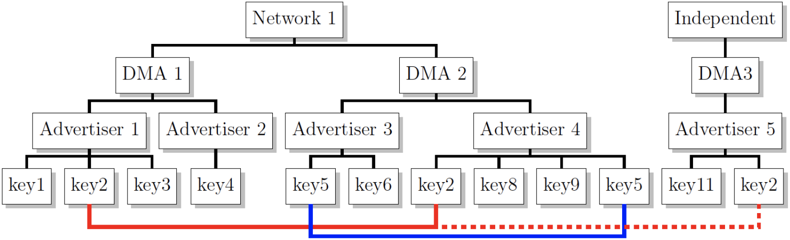
- Suppose there is a monopolist search engine selling 1 ad slot
- There are three advertisers (q, j, k) interested in the slot
- They have arbitrary bids: $b_q = 4$, $b_j = 3$ and $b_k = 1$
- They must bid through an intermediary (α, β or γ)
- 2-level Second Price Auction system



Data

- *Redbooks:*
  - Data on links advertisers-to-agencies
  - Yearly data 2011-2017 covering around 6,000 advertisers (i.e., web domains) per year active in all sectors ▸ advertisers
    - US: 4,400 publicly traded companies, plus largest private
    - Non US: top 2,000 global companies
  - For 2014-2017, link agencies to networks (ATD) ▸ networks
- *SEMrush:*
  - Data on links keywords-advertisers (URLs)
  - Google data on both paid and organic search
  - Up to the 50,000 most important keywords bid for each advertiser 2012 - 2017 (January), but with possibility to use higher frequency data (monthly/daily)
  - Keyword level: data on CPC, search volume, competition
  - Keyword/advertiser level: position, previous position, traffic

Data Structure



Data structure: keywords (SEMrush), advertisers (Redbooks/SEMrush), agencies and networks (Redbooks). Solid lines represent examples of coalitions: within DMA (blue) and network (red).

The relevant intermediary level is the **agency network** (in the example, Advertisers 1, 2, 3 and 4 are together under Network 1) ▸ descriptives

▸ Coalition Example ▸ DMA strategies ▸ Network strategies

Question and Strategy

- How do changes in intermediaries' concentration affect Google's revenues?
- A baseline regression model would be:

$$In(R^G_{mt}) = \beta DemandConcentration_{mt} + \phi X_{mt} + \tau_t + \gamma_z + \epsilon_{mt}$$

  - ${R^G}_{mt}$ = Search engine revenues in market m at time t
  - $DemandConcentration_{mt}$ = Measure of demand concentration
  - X_{mt} = Controls; time (τ_t), cluster (γ_z) FE

- But three main challenges:
  1. Definition of the relevant markets
  2. Measurement of relevant quantities
  3. Causal identification of β

1) Market Definition: two-step clustering

Advertisers' industries are *too broad*, but keywords are *too narrow*

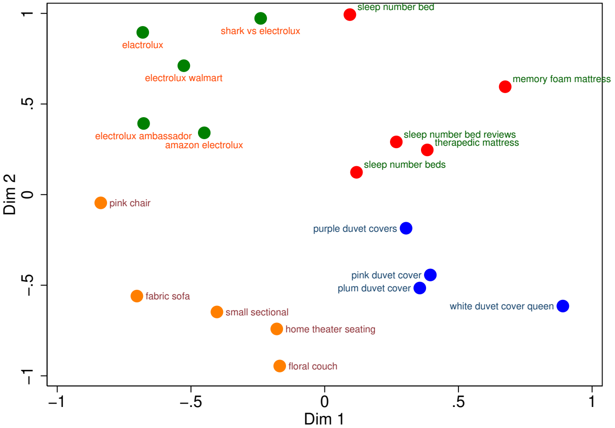
Our solution entails a two-layer clustering:

- Step 1: pool keywords together, but we have millions and many are related but not sharing any term. Solution: *GloVe*, unsupervised learning, pre-trained on 840B documents with 2.2M unique terms, from Common Crawl in English, featuring 300 dimensions ▸ details

Step 1: from Keywords to Thematic Clusters

Keyword	Industry
sleep number bed	Houseware
white duvet cover queen	Houseware
sleep number beds	Houseware
therapedic mattress	Houseware
memory foam mattress	Houseware
electrolux walmart	Houseware
elactrolux	Houseware
home theater seating	Houseware
amazon electrolux	Houseware
plum duvet cover	Houseware
shark vs electrolux	Houseware
pink duvet cover	Houseware
sleep number bed reviews	Houseware
purple duvet covers	Houseware
fabric sofa	Houseware
floral couch	Houseware
pink chair	Houseware
small sectional	Houseware
electrolux ambassador	Houseware

Step 1: from Keywords to Thematic Clusters



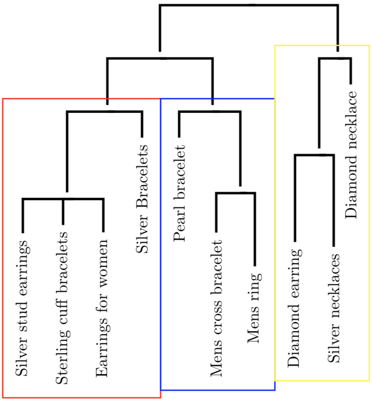
1) Market Definition: two-step clustering

Advertisers' industries are *too broad*, but keywords are *too narrow*

Our solution entails a two-layer clustering:

- Step 1: pool keywords together, but we have millions and many are related but not sharing any term. Solution: *GloVe*, unsupervised learning, pre-trained on 840B documents with 2.2M unique terms, from Common Crawl in English, featuring 300 dimensions details
- Step 2: Hierarchical clustering *within the thematic clusters of step 1* to account for competition (for any pair of keywords in a cluster, dissimilarity matrix built on co-occrrences of same advertisers)

Step 2: Hierarchical Clustering



Notes: Structure of competitive clusters: the three clusters - red, blue and yellow boxes - are identified through the Kelley, Gardner and Sutcliffe [1996] penalty parameter.

descriptives

2) Measurement of the Main Variables

We compute a proxy for R^G using data on the $i = 1, ..., N_r$ keywords bid by the sample of Redbooks' advertisers:

$$R_{mt} = \sum_{k \in K_m} CPC_{kmt} * Volume_{kmt} * CTR_{kmt}$$ distribution

- CPC_{kmt}: average Cost-per-Click of keyword k in market m at time t
- $Volume_{kmt}$ is the overall number of searches of k over an year
- CTR_{kmt} is the cumulative Click-through-Rate of all the sponsored ad slots shown for keyword k

And a proxy for demand concentration: $HHI_{mt} = \sum_{i=1}^{I} (s^i{}_{mt})^2$

- Market size (S_{mt}): sum of all the clicks of all the ad slots allocated in all the keywords in m: $S_{mt} = \sum_{k \in K_m} Volume_{kmt} * CTR_{kmt}$
- For intermediary i, representing the set of advertisers A_i, the market share in market m at time t is: $s^i_{mt} = \frac{1}{S_{mt}} \sum_{a \in A_i} \sum_{k \in K_m} \sum_{j \in J_k} CTR_{jkmt} * Volume_{kt} * 1\{a \text{ occupies } j \in J_k\}$

Alternatives: no CPC, agencies instead of networks, etc.

3) Causal Identification: IV Approach

- OLS unlikely to deliver causal effect due to OVB. Example: media attention to a phenomenon changes keyword entry/bid
- We adapt ideas from Dafny et al. (2012) of using M&A events as shocks to "local" market concentration ▸ mergers
- Hence, if in year t intermediary α merges with intermediary β, the merger-induced change in HHI is: details ▸ HHI(2017-2014)

$$sim\Delta HHI_{mt} = \underbrace{(s^{\alpha}_{m,t} + s^{\beta}_{m,t})^2}_{\text{Share of merged firm } \alpha+\beta} - \underbrace{((s^{\alpha}_{m,t})^2 + (s^{\beta}_{m,t})^2)}_{\text{Shares of single firms } \alpha \text{ and } \beta}$$

- Alternatives: we might want to exclude mergers too likely to be driven by specific keywords (too "local"); few overlapping markets; mergers with insufficient pre or post periods pre/post

Results: Baseline Estimates

	(1)		(2)		(3)		(4)		(5)	
	RF	FS	RF	FS	RF	FS	RF	FS	RF	FS
sim$\Delta\hat{HHI}$	-7.454***	0.605***	-4.070***	0.957***	-3.842***	0.830***	-3.831***	0.829***	-3.723***	0.831***
	(0.929)	(0.141)	(0.973)	(0.0765)	(0.993)	(0.0855)	(0.993)	(0.0855)	(0.988)	(0.0853)
Weak Id. F-Test	18.42	18.42	156.75	156.75	94.12	94.12	94.02	94.02	94.9	94.9
Underid. F-test	6.43	6.43	23.97	23.97	19.21	19.21	19.21	19.21	19.25	19.25
Observations	54,661	54,661	52,476	52,476	52,476	52,476	52,476	52,476	52,476	52,476
Cluster FE			✓		✓		✓		✓	
Year FE					✓		✓		✓	
Organic Results							✓		✓	
Keyword Characteristics									✓	

	OLS					IV				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\hat{HHI}$	-2.110***	-2.120***	-2.129***	-2.122***	-2.130***	-12.31***	-4.252***	-4.630***	-4.620***	-4.479***
	(0.0417)	(0.0457)	(0.0459)	(0.0459)	(0.0458)	(3.027)	(0.938)	(1.070)	(1.072)	(1.061)
Organic Results (billion)				0.252***	0.263***				0.206***	0.225***
				(0.0458)	(0.0484)				(0.0454)	(0.0478)
Keywords Characteristics										
Branded Keyword					0.396***					0.458***
					(0.0430)					(0.0532)
Long-tail Keywords					-0.0908***					-0.0491
					(0.0294)					(0.0356)
Observations	54,661	52,476	52,476	52,476	52,476	54,661	52,476	52,476	52,476	52,476
Cluster FE		✓	✓	✓	✓		✓	✓	✓	✓
Year FE			✓	✓	✓			✓	✓	✓

Robustness and Extensions

- **Validation and Channels**
  - Cluster validation → Amazon Mechanical Turk;
  - Heterogeneous effects at industry level → β_{IV} industry-level
  - Different Channels → Channels
- **Robustness**
  - Different definition of clusters → Table
  - Alternative proxies for R_{mt} and HHI_{mt} → Robustness
  - Individual Mergers → β_{IV} merger-level
- **Alternative Identification Strategies**
  - “Merged” markets only → β_{OLS} and β_{IV}

Conclusions

Main findings:

- First evidence that intermediaries' concentration reduces Google's revenue
- Novel approach for market definition in sponsored search

Considerations for competition policy:

1. Risk of abuses in Google's response to intermediaries:
   - Higher reserve prices: Google started increasing its reserve price in May 2017. AdRank made them "context specific" and more heavily based on max CPC. Who are the real losers?
   - Disintermediation: pay attention where Google's seeks to replace agencies, like with DoubleClick Search
2. When is growing buyers' power desirable:
   - Pass-through to advertisers (consumers) of lower prices or algorithmic collusion for the benefit of intermediaries?
   - Heterogenous impacts on smaller platforms (Bing, etc.)?

Intermediated Bidding and Demand Concentration

back

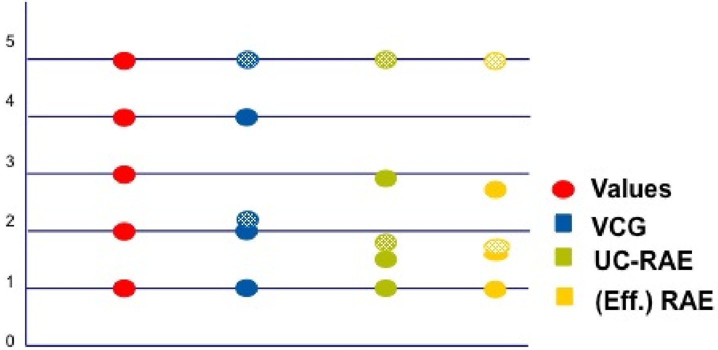
The demand side has vastly changed thanks to intermediaries:

- **Technological innovations**: automated bidding systems to address the need for more speed (high frequency or even real-time) and better data usage
- **Growing concentration**: 7 large ATDs, active at the agency network level ATD list

	Search Volume Share				Presence across Keywords			
	2014	2015	2016	2017	2014	2015	2016	2017
IPG	0.21	0.19	0.21	0.19	0.26	0.32	0.33	0.38
WPP	0.17	0.20	0.16	0.23	0.29	0.29	0.33	0.43
Omnicom	0.17	0.16	0.17	0.14	0.39	0.38	0.37	0.38
Publicis	0.14	0.13	0.13	0.18	0.30	0.30	0.29	0.30
MDC	0.09	0.09	0.08	0.09	0.17	0.17	0.17	0.24
Havas	0.05	0.07	0.06	0.02	0.12	0.14	0.12	0.06
Dentsu-Aegis	0.05	0.08	0.10	0.09	0.14	0.17	0.19	0.25
Indep Age	0.13	0.09	0.08	0.06	0.42	0.38	0.35	0.22

Review of Decarolis-Goldmanis-Penta (2017): Theory

back



Comparison: VCG, UC-RAE, E-RAE, RAE

Example of Data and Coalition

Case Study - DD back

Merkle: large DMA with multiple clients (Redbooks data) active on the same keywords (SEM Rush data)

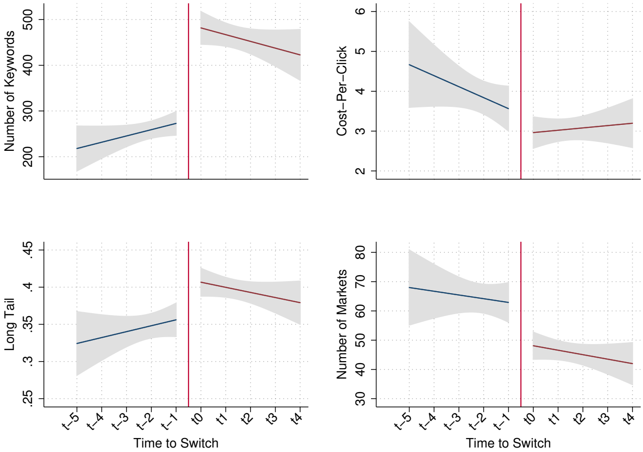
Example from charity sector: *Habitat for Humanitas* and *Salvation Army*

Keyword	CPC ($)	Volume (mil)	Position	
			Habitat	*Salv.Army*
habitat for humanity donations pick up	4.01	40	1	4
charities to donate furniture	1.08	20	3	9
donate online charity	0.93	20	11	10
website for charity donations	0.90	19	11	6
salvation army disaster relief fund	0.03	20	2	1

In July 2016, Merkle acquired by Dentsu-Aegis for $1.5 billion dollars. Change in concentration in many markets with Merkle/Dentsu-Aegis advertisers

DMA strategies: effects of affiliation

back

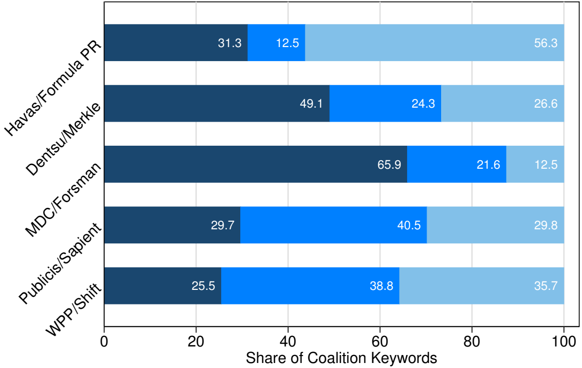


Network Strategies: Coalitions and Market Split

Case Study - DD

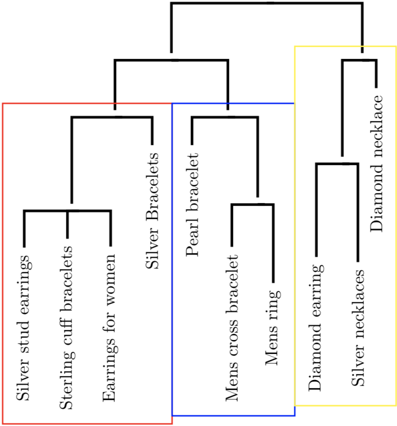
back

- One illustrative M&A per network DMA strategies
- Sample of common keywords (pre, post, or both) in a 2-years window around the acquisition



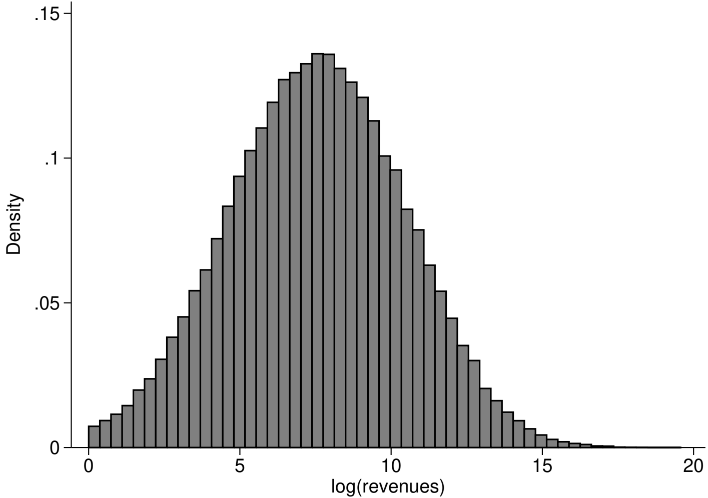
Example of Hierarchical Clustering

back



Notes: Structure of competitive clusters: the three clusters - red, blue and yellow boxes - are identified through the Kelley, Gardner and Sutcliffe [1996] penalty parameter.

Distribution of $log(\hat{R})$ back



Merger Events

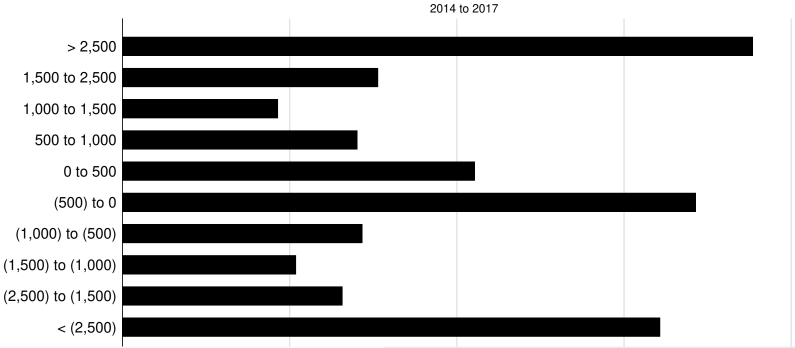
back

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

Change in local concentration - 2014 to 2017 descriptives-mkt

back

- we observe 21 M&A and 2 divestures
- $HHI_{m,2017} - HHI_{m,2014}$
- $HHI \in [0 - 10,000]$



Largest Individual Mergers of Four Different Agency Networks back

Panel a): Individual Mergers								
	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***
	(2.160)	(0.363)	(1.126)	(0.0363)	(3.236)	(0.192)	(5.345)	(0.165)
Observations	4,776	4,776	3,047	3,047	3,013	3,013	981	981
Panel b): Individual Mergers: no Top 10% markets								
	Sapient		Merkle		Shift		Forsman & Bodenfors	
	RF	FS	RF	FS	RF	FS	RF	FS
sim$\Delta\hat{HHI}$	-2.757	1.033***	-5.216***	1.374***	4.655	0.789***	-8.862*	6.450***
	(2.295)	(0.354)	(1.126)	(0.0412)	(2.952)	(0.212)	(4.733)	(0.159)
Observations	4,330	4,330	2,736	2,736	2,719	2,719	909	909
Industry FE	✓		✓		✓		✓	
Year FE	✓		✓		✓		✓	
Organic Results	✓		✓		✓		✓	
Keyword Characteristics	✓		✓		✓		✓	

Results: Different Channels (IV estimates) back

	$log(\hat{R})$	$log(cpc)$	$log(volume)$	$log(\#keywords)$
	(1)	(2)	(3)	(4)
$\hat{HHI}$	-3.024***	-2.473***	-0.734	2.681***
	(1.143)	(0.507)	(0.797)	(0.941)
Organic Results (billion)	0.604***	0.0955***	0.502***	-0.0626*
	(0.131)	(0.0369)	(0.104)	(0.0325)
Observations	21,917	21,917	21,917	21,917
Industry FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓

Main AD Networks and their Agency Trading Desks

back

Network	ATD	Year
Dentsu-Aegis	Amnet Accordant Media	2011 2016
Publicis Groupe	Vivaki (Audience on Demand, AOD) RUN Precision	2008-2014 2014-2014 2017
IPG	Cadreon (Mediabrands Audience Platform)	2009
Omnicom Group	Accuen	2009
WPP/Group M	Xaxis	2011
Havas	Adnetik (spun off as an independent company in 2010) Affiperf	-2010
MDC	Varick Media	2008

Agency Networks and Trading Desks

An advertiser bids by itself or via DMA, possibly part of an agency network, typically paying it a negotiated lump sum amount per campaign

Programmatic buying: the algorithmic purchase of ads space in real time. Software automatizes the buying, placement, and optimisation of media inventory via bidding.

Agency Trading Desk: the unit within a media buying agency that centralizes programmatic buying for "biddable" media like Google, Bing, Twitter, iAd, and Facebook. ATDs

Agency Network	Agency Trading Desk	Number of Advertisers	Number of Agencies
IPG	Cadreon	742	175
WPP	Xaxis	858	294
Omnicom Group	Accuen	951	248
Publicis Groupe	Vivaki	685	172
MDC	Varick Media	225	35
Havas	Affiperf	169	46
Aegis-Dentsu	Amnet	185	47
Other	ITD (∼5-50)	5,859	2,565

GSP with Quality Scores

- Google and Bing-Yahoo! form of the GSP uses advertiser specific *'quality scores'* (e_i)
- Suppose CTR are: $CTR(i) = e_i \cdot x^{\rho(i)}$
- Ranking of advertisers is now by $e_i \cdot b_i$
- Price-per-click for position $\rho(i)$ is $p_i = e^{\rho(i+1)} b^{\rho(i+1)} / e^{\rho(i)}$
- Necessary and sufficient condition for EOS is: MEF

$$v_i = \frac{b_i x^{i-1} - b_{i+1} x^i}{x^{i-1} - x^i} > \frac{b_{i+1} x^i - b_{i+2} x^{i+1}}{x^i - x^{i+1}} = v_{i+1}$$

- Relabeling advertisers so that $e_i v_i > e_{i+1} v_{i+1}$, EOS condition becomes:

$$e_i v_i = \frac{e_i b_i x^{i-1} - e_{i+1} b_{i+1} x^i}{x^{i-1} - x^i} > \frac{e_{i+1} b_{i+1} x^i - e_{i+2} b_{i+2} x^{i+1}}{x^i - x^{i+1}} = e_{i+1} v_{i+1}$$

Detecting Coordination

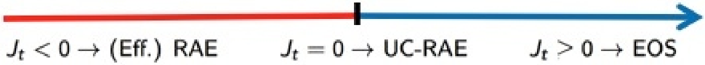
A simple criterion to detect collusion based on the *only* observable difference between collusion and (EOS) competition:

- for both competitive and collusive equilibria, the inequality below holds for all independent bidders:

$$e_i v_i = \frac{e_i b_i x^{i-1} - e_{i+1} b_{i+1} x^i}{x^{i-1} - x^i} > \frac{e_{i+1} b_{i+1} x^i - e_{i+2} b_{i+2} x^{i+1}}{x^i - x^{i+1}} = e_{i+1} v_{i+1}$$

- but, in the collusive equilibria, it is violated for all colluders that are not the highest-valuation bidder
- consider agency bidder j, $j \notin \{min(\mathcal{C})\}$, then:

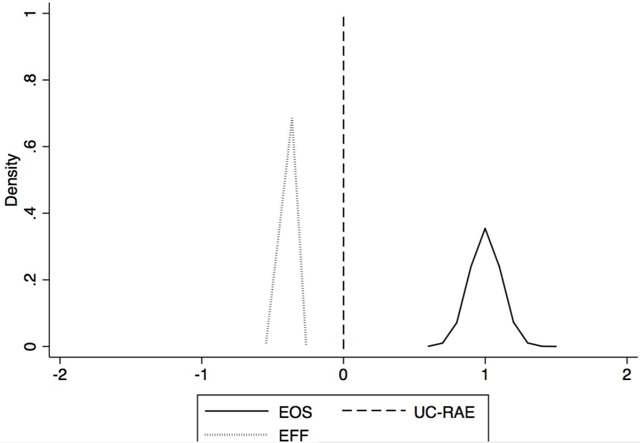
$$J_t = \frac{e_j b_j x^{j-1} - e_{j+1} b_{j+1} x^j}{x^{j-1} - x^j} - \frac{e_{j+1} b_{j+1} x^j - e_{j+2} b_{j+2} x^{j+1}}{x^j - x^{j+1}}$$



Simulation: Baseline Case

Fix the valuations, CTRs and coalition structure as in the example.
Simulate 100,000 auctions by iid draws of $e_{it} \sim N(\mu = 1, \sigma = .03)$

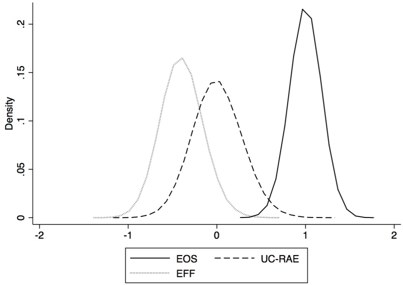
Case 1: No Noise



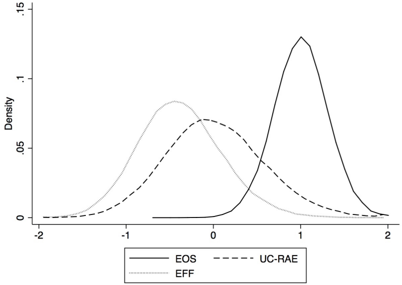
Simulation with Belief Errors on Quality Scores

True quality score is e_{it}, but bidders believe score to be $\tilde{e}_{it} = d_{it} \cdot e_{it}$

Case 2: Small noise
$d_{it} \sim \mathcal{N}(1, 0.05^2)$

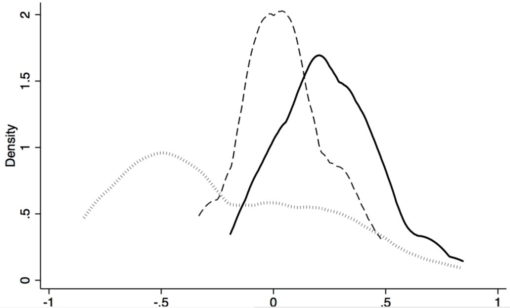


Case 3: Big Noise
$d_{it} \sim \mathcal{N}(1, 0.1^2)$



Detecting Coordination in Real Data: 3 Example Keywords

- **Data**: 8-slot auctions held in 2011; one SEMA handling 2 bidders; 71 keywords (with different SEMA/bidders)
- **Criterion**: 95% C.I. for median of J_t
- **Results**: 3 keywords $\rightarrow$ (E-)RAE; 36 keywords $\rightarrow$ UC-RAE; 32 keywords $\rightarrow$ EOS



Revenue Quantification

Details

1. Use J_t to detect cases of likely coordination
2. Recover independents' values by inverting equilibrium bids
3. Use values for independents to bound agency bidders' values
4. Compute counterfactual upper/lower bound competitive bids
5. Example: revenue effects for 36 keyword detected as UC-RAE

	Observed	Lower Bound	Upper Bound	Δ *UpB.-Obs.*
Agency	33.2	32.1	35.3	2.1
				[1.5; 2.7]
Others	66.8	64.2	72.6	5.8
				[3.7; 7.9]
Total	100	96.3	107.9	7.9
				[5.3; 10.4]

Revenue Quantification

Back

1. Use J_t to detect cases of likely coordination
2. Recover independents' values from data, inverting equilibrium bids
   - Obviously, equilibrium restrictions are not exactly satisfied by the data
   - *Varian's method:* assume data are generated by a compl.info. model in which quality scores are $e_i' = d_i \cdot e_i$ (distance $|d_i - 1|$ identifies **belief error on quality scores**)
   - *Small belief-errors are needed to reconcile data with compl.info. model*
   - Separately for each auction, recover the *smallest belief errors d* needed to rationalize data under the chosen equilibrium: $min_d \sum_{i>1}(d_i - 1)^2$ s.t. eq. restrictions with d
3. Use the inferred values for the independents to bound agency bidders' values:
   - If j is the lowest valued agency member, v_j bounded from below by the value of the bidder in position $\rho(j+1)$ and bounded above by the bidder in position $\rho(j-1)$
4. Compute counterfactual upper/lower bound **competitive bids and revenues**

Entrant agencies *per network* in 2017

Panel a): Previously Independent						
Dentsu-Aegis	Publicis	IPG	Omnicom	WPP	Havas	MDC
Grip	North Strategic	BPN Worldwide		Cavalry Agency		Forsman & Bodenfors
Gyro		ReviveHealth		Deeplocal		Laird+Partners
Happy Creative Services		StickyEyes		Essence Digital		
Merkle				Mirum Global		
				Muh-Tay-Zik Hof-fer		
				SHIFT Communications		
				Zubi Advertising Services		
				iStrategyLabs		
Panel b): Brand New						
Dentsu-Aegis	Publicis	IPG	Omnicom	WPP	Havas	MDC
Band Pte		Flipside Group	Hearts & Science	Code Computerlove	Ignition Holdings	
Barnes Catmur & Friends		Healix	United State of Fans	Conrad Caine Gmbh		
C2C Outdoor		Rapport Worldwide		Famous nv/sa		
IMPAQT		SociedAD		Quirk		
Perfect Relations		Trilia Media		m/SIX		
				Tank		

Notes: Previously Independent (panel a) and brand new (panel b) agencies merged and acquired by the 7 networks during 2016.

Summary Statistics by Keyword - Advertisers

back

	Keywords with at Least 1 Network Years 2014/2017			
	Mean	Median	SD	Obs
Cost-per-click	2.34	0.90	5.79	15,383,769
Volume (000)	498	40	34,916	15,383,769
Traffic	0.01	0.00	0.53	15,383,769
Competition	0.58	0.69	0.39	15,383,769
Num of Advertisers	1.30	1.00	0.68	15,383,769
Organic Results	47	1.8	260	15,383,769
# Characters	22.79	22.00	7.74	15,383,769
# Words	3.71	4.00	1.35	15,383,769
Long Tail	0.50	1.00	0.50	15,383,769
Branded	0.10	0.00	0.29	15,383,769
Coalition	0.15	0.00	0.36	15,383,769
Coalition Size	2.38	2.00	0.69	332,017

Summary Statistics by Keyword - Advertisers / 2

back

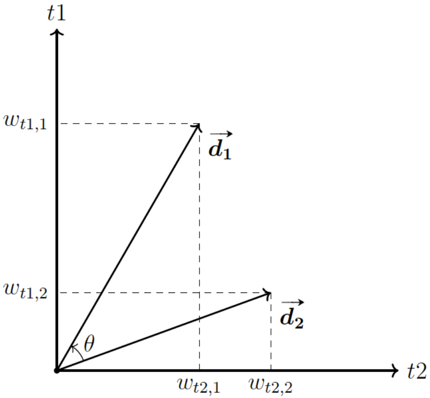
	Keywords with at Least 1 Independent Years 2014/2017			
	Mean	Median	SD	Obs
Cost-per-click	2.39	0.89	6.11	21,525,056
Volume (000)	362	40	99,845	21,525,056
Traffic	0.06	0.00	1.27	21,525,054
Competition	0.59	0.73	0.39	21,525,056
Num of Advertisers	1.21	1.00	0.52	21,525,056
Organic Results	3.8	0.16	19	21,525,056
# Characters	22.86	22.00	7.59	21,525,056
# Words	3.66	3.00	1.30	21,525,056
Long Tail	0.48	0.00	0.50	21,525,056
Branded	0.07	0.00	0.25	21,525,056

GloVe algorithm ▸ Back

- the GloVe approach starts by building a matrix of word co-occurences within documents in a corpus. In our case, 840 billions+ documents gathered within the Common Crawl, all in English; these correspond to $\approx$ 2.2 million unique terms (g)
- through a log-bilinear regression model - i.e., a weighted version of the global factorization methods like latent semantic analysis - the model yields a matrix of dimension $g \times d$ (in our case, $d = 300$)
- we merge the keywords *term by term* ($\approx$ 1 million) with the GloVe pre-trained set - with around 85% matches
- aggregate the resulting vectors taking the sum of GloVe vectors (baseline) or the mean (robustness)

K-means algorithm on cosine distance

Back



- we take the cosine of the distance
- run a spherical K-means on the cosine distances between vectorized keywords (K = 1,000)

Example Results - *Pharmaceutical & Health* ▸ Back

Cluster	Keyword
85	aarp pharmacy prices
85	insurance with medicare
85	medical supplies medicare
85	medicare approved drug list
85	medicare approved pharmacies
85	medicare health providers
85	medication coverage
85	medication insurance coverage
65	best caterers in boston
65	catering denver colorado
65	catering in hamilton nj
65	food catering denver
65	italian catering denver
65	metro detroit catering
65	omaha catering restaurants
65	sushi catering boston

▸ Check by Amazon Mechanical Turk

Dentsu-Aegis acquisition of Merkle

Back

What are the actual effects of DMA concentration? The idea is to analyze a major M&A case

- **Strategy**: diff-in-diff analysis exploiting Dentsu-Aegis acquisition of Merkle in July 2016 creating at least 7 cases:
  - Electronics: Dell, Samsung $\rightarrow$ Apple, HP, IBM/Lenovo, Intel (also: eBay, HomeDepot, Target, Walmart)
  - Financial: LendingTree, MetLife $\rightarrow$ Capitalone, Discover, Fidelity, Equifax, JP Morgan-Chase
  - Car manufacturers: Fiat Chrysler Automotive, Mercedes-Benz USA $\rightarrow$ Toyota, Volkswagen, Subaru (also: Autotrader, KBB, eBay)
  - Phone services: Vonage $\rightarrow$ Tmobile
- Model specification, run separately for each Merkle advertiser:

$$CPC_{kt} = a_k + b_t + \beta(PostMerger_t * SharedKeyword_k) + \varepsilon_{kt}$$

where k = keyword and t = month and year pair.

The Case of Dentsu-Aegis/Merkle: Diff-in-diff

Back

- Define treatment/control: ever shared vs never shared
- Select keywords: top 30, 50, 100, 500 by traffic volume

Industry	Advertiser	30 key	50 key	100 key	500 key
Electronics	Dell	-2.84***	-1.82***	-1.33***	-0.22
		(0.16)	(0.16)	(0.08)	(0.59)
	Samsung	-0.04	0.22	0.14	-0.10
		(0.88)	(0.52)	(0.67)	(0.38)
Financial	LendingTree	-0.75**	-0.82***	-0.25***	-0.36***
		(0.46)	(0.31)	(0.62)	(0.07)
	MetLife	-1.57***	-1.27***	-0.36	0.74*
		(0.22)	(0.38)	(0.71)	(0.39)
Automotive	FCA	-2.05***	-1.28***	-0.99***	-0.54***
		(0.02)	(0.07)	(0.06)	(0.14)
	MBauto	0.22	0.08	0.47***	0.65***
		(0.45)	(0.73)	(0.00)	(0.00)
Telecommunications	Vonage	3.37***	3.13***	2.97***	2.03***
		(0.17)	(0.22)	(0.17)	(0.11)
Observations		120	200	400	2,000

Dependent Variable and Market Definition

Main outcome variable obtained by aggregating at market level:

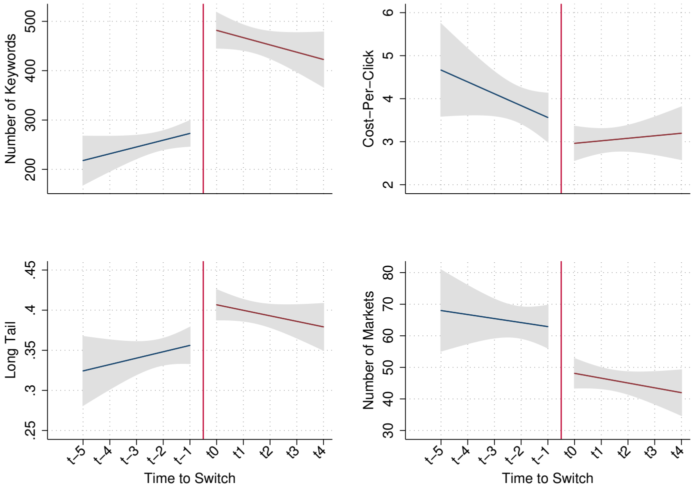
- $\hat{R}_{jt} = \sum_{i=1}^{N_R} CPC_{it} * Volume_{it} * \hat{CTR}(B_{it}) * 1(market_i == j)$, where $j \in [1, .., J]$ stands for advertisers' market

Definition of the market is thus crucial step. Various options:

- Ideal: as in antitrust/merger, but lack data on demand;
- Redbooks: use the *industry* definition provided on a subset of advertisers, and use SEMrush data to impute the rest;
- SEMrush: text clustering. We use a k-means algorithm (J = 3,000/5,000/10,000) on vectorized keywords - more on that later.

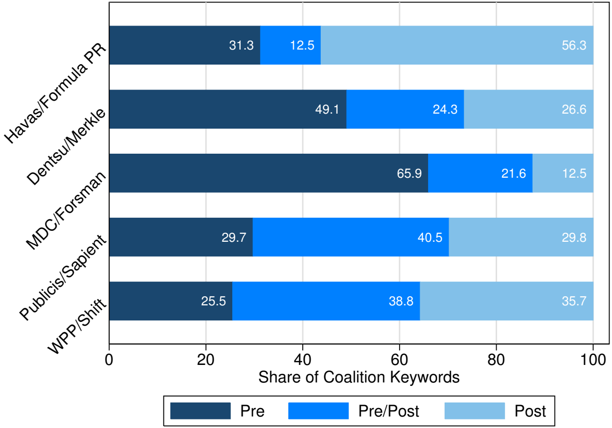
Effects of DMA affiliation - key metrics

back



Pre/post coalition keywords in mergers

back



Results: Robustness without Publicis

back

Panel a) OLS and IV Estimates

	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\hat{HHI}$	-0.943***	-0.934***	-0.787***	-0.777***	-4.226***	-4.250***	-2.679**	-2.675**
	(0.0424)	(0.0427)	(0.0460)	(0.0460)	(1.296)	(1.309)	(1.224)	(1.227)
Organic Results (billion)		0.185***	-0.0943	-0.119*		-0.495*	-0.306**	-0.319**
		(0.0703)	(0.0608)	(0.0619)		(0.290)	(0.156)	(0.150)
Keywords Characteristics								
Branded Keywords				0.0116				-0.0226
				(0.0326)				(0.0435)
Long-tail Keywords				-0.115***				-0.0394
				(0.0268)				(0.0565)
Observations	39,179	39,179	39,179	39,179	39,179	39,179	39,179	39,179
Industry FE			✓	✓			✓	✓
Merger FE			✓	✓			✓	✓
Year FE			✓	✓			✓	✓

Panel b) Reduced Form and First-Stage Results

	(1)		(2)		(3)		(4)	
	RF	FS	RF	FS	RF	FS	RF	FS
sim$\Delta\hat{HHI}$	-5.570***	1.318***	-5.546***	1.305***	-4.001**	1.493***	-3.979**	1.488***
	(1.730)	(0.106)	(1.731)	(0.105)	(1.862)	(0.0833)	(1.854)	(0.0876)
Weak Id. F-Test	153.59	153.59	153.35	153.35	321.56	321.56	288.13	288.13
Underid. F-test	6.72	6.72	6.72	6.72	7.73	7.73	7.60	7.60
Observations	39,179	39,179	39,179	39,179	39,179	39,179	39,179	39,179
Organic Results			✓		✓		✓	
Industry FE					✓		✓	
Merger FE					✓		✓	
Year FE					✓		✓	
Keyword Characteristics							✓	

IV Estimates: Different Outcomes

back

Table: Analytical Refinements: IV Estimates on Different Outcomes

	Industry Level		Clustering		Complete		No *Publicis*		No Outliers	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\hat{HHI}$	-22.80	0.258	-28.96	-6.377***	-22.08	-5.534***	-43.12*	-5.517***	-41.16	-6.256***
	(19.16)	(15.05)	(32.27)	(0.828)	(15.81)	(0.488)	(13.67)	(0.160)	(61.20)	(0.349)
Organic Results (billion)		0.519		0.656**		0.680**		0.737**		0.853**
		(1.020)		(0.164)		(0.149)		(0.168)		(0.247)
Observations	92	92	214,107	214,107	230,616	230,616	214,842	214,842	187,735	187,735
Industry FE		✓		✓		✓		✓		✓
Merger Dummies		✓		✓		✓		✓		✓
Year FE		✓		✓		✓		✓		✓

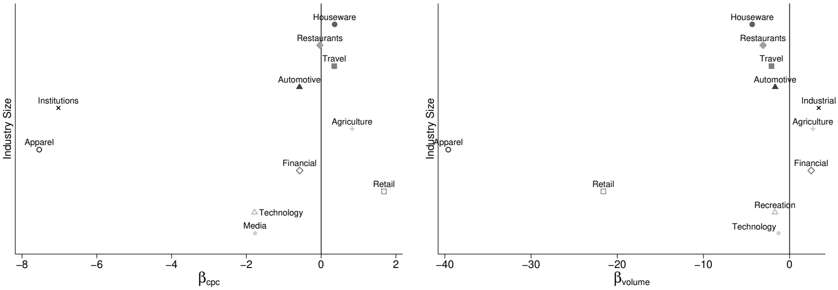
Results without *Media* and *Pharmaceutical* back

Table: IV Estimates without *Media* and *Pharmaceutical*

	$\Delta ln(\hat{R})$		$\Delta ln(\#keywords)$		$\Delta ln(volume)$		$\Delta ln(cpc)$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\hat{HHI}$	-4.093***	-3.295***	-0.252	-0.308	1.509	1.645	-0.958*	-0.437
	(1.274)	(1.232)	(0.744)	(0.686)	(1.028)	(1.041)	(0.577)	(0.466)
Organic Results (billion)		-0.364*		-0.247**		0.518***		0.0168
		(0.206)		(0.119)		(0.168)		(0.0809)
Observations	35,050	35,050	35,050	35,050	35,050	35,050	35,050	35,050
Industry FE		✓		✓		✓		✓
Merger Dummies		✓		✓		✓		✓
Year FE		✓		✓		✓		✓

Industry-level IV estimates distribution

back



The Case of Dell: Keyword Example

Top 10 shared keywords (by traffic volume)

		Pre Merger			Post Merger		
keyword	N. Obs.	CPC	Position	N. Bidders	CPC	Position	N. Bidders
build your computer	17	1.87	3.80	2.17	1.23	1.00	2.00
		(0.17)	(0.84)	(0.75)	(0.31)	(0.70)	(0.12)
cloud computing	16	71.90	3.10	2.10	35.64	2.00	1.67
		(30.75)	(2.08)	(0.74)	(0.00)	(1.00)	(0.82)
computer deals	17	1.73	2.63	2.33	1.71	1.75	2.20
		(0.11)	(1.60)	(0.89)	(0.00)	(0.50)	(1.10)
dell 2 in 1	17	0.87	1.09	2.17	1.04	1.50	2.00
		(0.28)	(0.30)	(0.83)	(0.00)	(1.00)	(0.00)
desktop computer	16	1.35	2.25	3.67	1.63	1.75	2.25
		(0.24)	(1.36)	(1.30)	(0.00)	(0.96)	(1.26)
desktop computers	18	2.52	2.08	3.67	1.99	2.00	2.33
		(0.47)	(1.00)	(1.30)	(0.00)	(1.41)	(0.52)
laptops	17	3.65	3.20	4.25	5.26	3.00	2.00
		(0.99)	(2.20)	(1.06)	(0.00)	(1.73)	(0.71)
laptops on sale	15	1.93	4.56	4.00	2.59	5.00	4.00
		(0.41)	(1.81)	(1.28)	(0.00)	(2.83)	(0.00)
small laptop	16	1.86	4.22	4.45	3.55	3.00	2.20
		(0.84)	(2.11)	(1.21)	(0.00)	(1.41)	(1.10)
windows laptops	13	2.60	2.27	3.33	3.01	3.00	3.00
		(0.43)	(1.35)	(1.15)	(0.00)	(0.00)	(0.00)

Amazon Mechanical Turk - the Task

back

- "Non-machine" learning test for the correctness of data-driven clusters
- Generally used for similar tasks - e.g. generate training sets for neural netowrks (patterns recognition, captcha, optical character identification)
- 23,000 clusters to be tested $\rightarrow$ impossible for individuals
- Simple task $\rightarrow$ given a reference keyword belonging to cluster k, link another term among two alternatives, one drawn from k, one from cluster j in the same industry
- Two versions:
  - Alternative keywords drawn from *all* other keywords
  - Alternative keywords drawn from the set of keywords with *no term* in common with the reference keyword

Amazon Mechanical Turk - the Task

back

Mergers 2014-2017: All networks

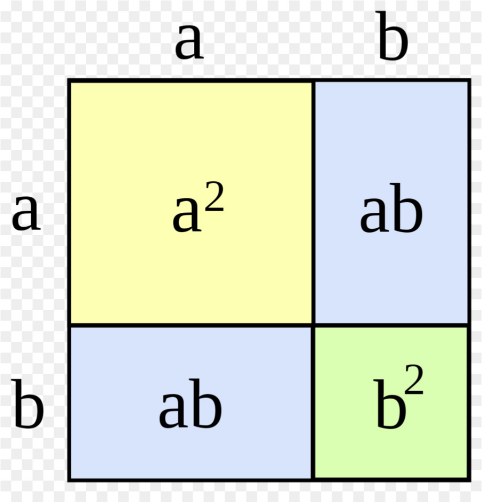
back

Agency	Acquiring Network	Acquisition year	Number of Advertisers	Number of Industries	Number of Markets
The Brooklyn Brothers	IPG	2016	6	2	19
Essence Digital Limited	WPP	2015	1	1	74
Quirk	WPP	2015	5	2	272
SHIFT Communications	WPP	2017	13	8	700
Deeplocal Inc.	WPP	2017	5	1	74
Maruri GREY	WPP	2017	1	1	133
Zubi Advertising Services, Inc.	WPP	2017	3	2	185
Campfire	Publicis	2015	3	1	21
La Comunidad	Publicis	2015	9	5	181
Sapient Corporation	Publicis	2015	17	6	514
Blue 449	Publicis	2016	4	2	76
Forsman & Bodenfors	MDC	2017	5	1	155
Formula PR	Havas	2015	6	4	189
FoxP2	Dentsu-Aegis	2015	1	2	31
Rockett Interactive	Dentsu-Aegis	2015	1	1	12
Covario, Inc.	Dentsu-Aegis	2015	3	1	54
Achtung	Dentsu-Aegis	2016	2	1	100
Gravity Media	Dentsu-Aegis	2016	5	3	249
Grip Ltd.	Dentsu-Aegis	2016	3	2	70
Merkle	Dentsu-Aegis	2017	18	7	567
Gyro	Dentsu-Aegis	2017	12	6	270

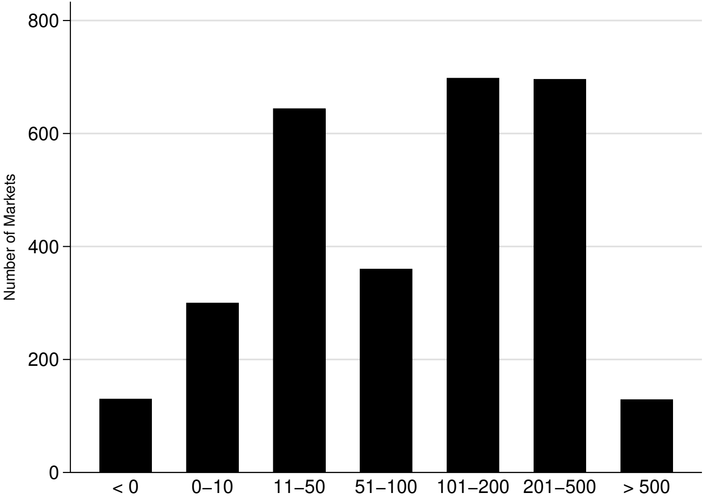
$sim\Delta HHI_{mt}$: instrument assessment back

- Instrument definition depends on the number, and the extent, of network M&A in our data list
- Main assumption: there is no reverse causality at the **local market** level, in the sense that the merger did not take place with the aim of increasing concentration in local markets relevance
- $sim\Delta HHI_{mt}$ takes different values, depending on the merger and the market distribution

$sim\Delta HHI_{mt}$: exogeneity back

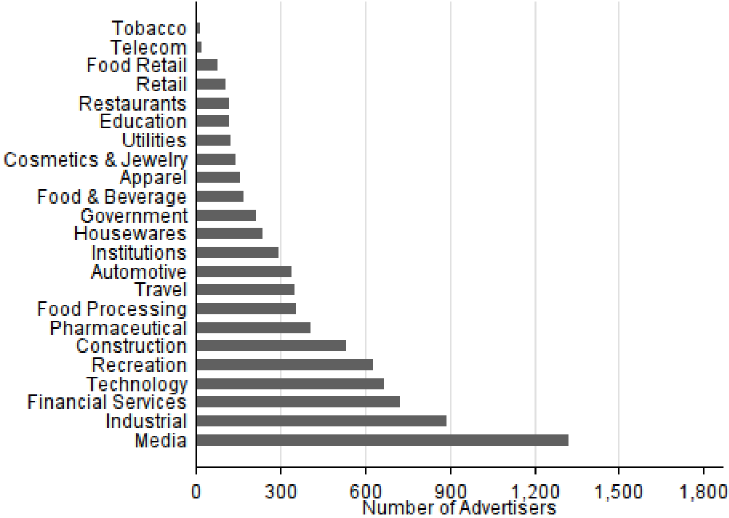


$sim\Delta HHI_{mt}$: distribution back



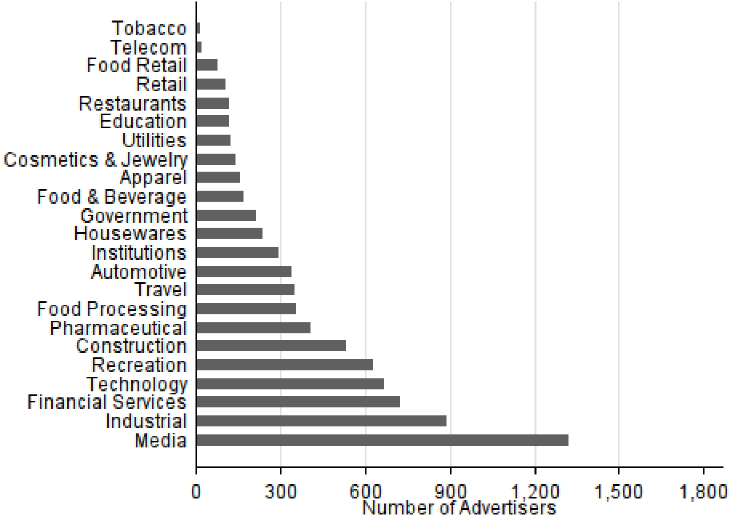
Distribution of advertisers per industry

back



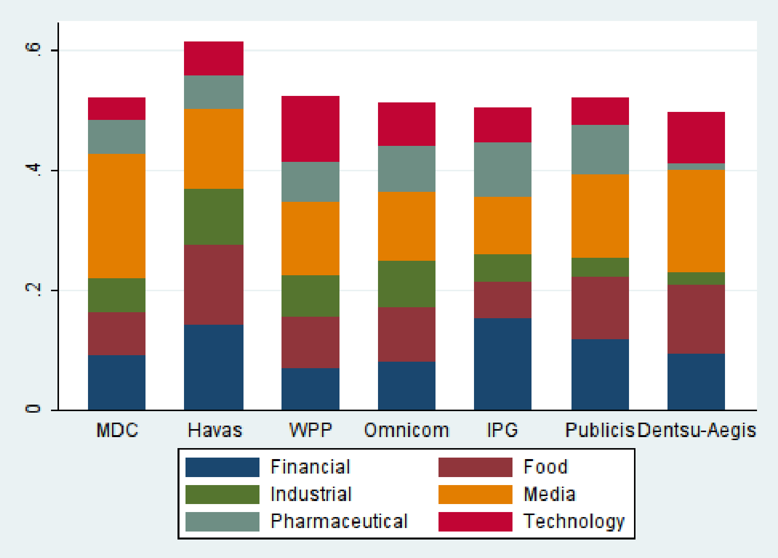
Distribution of advertisers per industry

back



Network Industry Specialization

back



Results: Robustness Checks back

- Control for Agency Trends, market by market

	Agency Trend		$\bar{R}$ on $\hat{HHI}$	$\bar{R}$ on $\bar{HHI}$
	(1)	(2)		
$\hat{HHI}$	-2.442	-3.187***		
	(1.543)	(1.208)		
$\bar{HHI}$				
Organic Results (billion)		-0.348**		
		(0.151)		
Observations	39,179	39,179		
DMA × Trend	✓			
Industry FE		✓		
Merger FE		✓		
Year FE		✓		

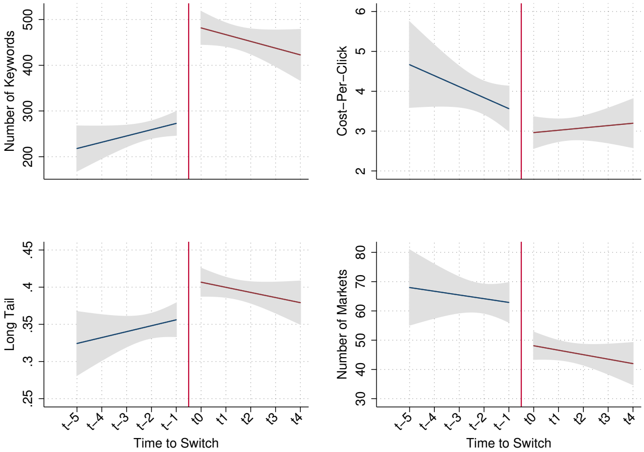
Results: Robustness Checks back

- Control for Agency Trends, market by market
- Alternative outcomes and concentration measures

	Agency Trend		$\bar{R}$ on $\hat{HHI}$		$\bar{R}$ on $\bar{HHI}$	
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{HHI}$	-2.442	-3.187***	-3.919**	-2.897*		
	(1.543)	(1.208)	(1.654)	(1.576)		
$\bar{HHI}$					-3.830**	-2.865**
					(1.525)	(1.442)
Organic Results (billion)		-0.348**		-0.254		-0.258
		(0.151)		(0.182)		(0.173)
Observations	39,179	39,179	39,179	39,179	39,179	39,179
DMA $\times$ Trend	✓	✓				
Industry FE		✓		✓		✓
Merger FE		✓		✓		✓
Year FE		✓		✓		✓

DMA strategies: effects of affiliation

back



Summary Statistics - Markets

back

	Mean	SD	Median	Obs		Mean	SD	Median	Obs
$log(Revenues)$	6.96	2.96	6.98	90,138	ΔR	-0.09	2.05	-0.04	60,336
HHI	4,741	2,922	3,970	95,516	ΔV	0.03	0.55	0.05	63,405
Long-tail	0.40	0.42	0.23	95,516	ΔK	-0.13	0.78	0.00	63,405

Summary Statistics by Keywords and Advertiser Type

	Keywords with at Least 1 Network Years 2014-2017				Keywords with at Least 1 Independent Years 2012-2017			
	Mean	Median	SD	Obs	Mean	Median	SD	Obs
Cost-per-click	2.34	0.90	5.79	15,383,769	2.39	0.89	6.11	21,525,056
Volume (000)	497	40	34,916	15,383,769	362	40	99,845	21,525,056
Traffic	0.01	0.00	0.53	15,383,769	0.06	0.00	1.27	21,525,054
Competition	0.58	0.69	0.39	15,383,769	0.59	0.73	0.39	21,525,056
Num of Advertisers	1.30	1.00	0.68	15,383,769	1.21	1.00	0.52	21,525,056
Organic Results	4.70	0.18	26	15,383,769	3.8	0.16	19	21,525,056
# Characters	22.79	22.00	7.74	15,383,769	22.86	22.00	7.59	21,525,056
# Words	3.71	4.00	1.35	15,383,769	3.66	3.00	1.30	21,525,056
Long Tail	0.50	1.00	0.50	15,383,769	0.48	0.00	0.50	21,525,056
Branded	0.10	0.00	0.29	15,383,769	0.07	0.00	0.25	21,525,056
Coalition	0.15	0.00	0.36	15,383,769	0.00	0.00	0.00	21,525,056
Coalition Size	2.38	2.00	0.69	332,017	-	-	-	-