Measuring Complementarities in Vertical Markets: Evidence from the Digital Advertising Industry

Andrea Chiantello, Francesco Decarolis, Maris Goldmanis and Antonio Penta*

February 23, 2021

Abstract

The digital advertising industry is characterized by a proliferation of specialized intermediaries helping advertisers in their purchases of online ad space. This study contributes to the analysis of the vertical complementarities in this market by posing and estimating a structural econometric model of how advertisers match to their intermediaries. Exploiting novel data and some recent methods in the estimation of many-to-many matching games for large markets, we quantify the value created by the matches, their driving forces and, counterfactually, evaluate the likely effects of the growing concentration among intermediaries. The estimates clearly show that competing advertisers benefit from dealing with a common intermediary and, moreover, offer a precise quantification of various forces including industry specialization, exclusive contracting and diversification.

JEL: C72, D44, L81.

Keywords: Online Advertising, Delegation, Matching, Sponsored Search Auctions.

*Chiantello, Bocconi University, *andrea.chiantello@studbocconi.it*. Decarolis, Bocconi University and IGIER, *francesco.decarolis@unibocconi.it*. Goldmanis, Department of Economics, Royal Holloway, University of London, *Maris.Goldmanis@rhul.ac.uk*. Penta, ICREA-UPF, Barcelona GSE and TSE Digital Center, *antonio.penta@upf.edu*.

1 Introduction

Digital advertising is a new, large and rapidly growing industry. In 2019, total digital ad spending is expected to be worth more than 300 billion US dollars globally. But taking advantage of the opportunities offered by digital advertising requires advertisers to adopt novel technologies and acquire specialized skills. Most firms, and especially those which are not in the tech sector, rarely invest to develop these capabilities in-house, but rely instead on outsourcing these functions to specialized intermediaries. These intermediaries, commonly referred to as Digital Marketing Agencies (DMAs), offer a disparate variety of services to their clients, ranging from managing their bidding strategies on online ad auctions platforms, to the design and management of their ad campaigns more broadly. Despite the relevance of this vertical relationship between advertisers and intermediaries, still only a few studies have analyzed its drivers and implications. In this study, we contribute to the understanding of such vertical relations by posing and estimating a structural econometric model of how advertisers match to their intermediaries. The main objective is to quantify the value created by these matches as well as their driving forces and, counterfactually, evaluate the likely effects of the growing concentration among intermediaries which this market is witnessing.

We build on earlier research about intermediaries in ad auctions to formulate a model of how intermediaries create value for their clients. Besides the diversity of services DMAs offer to their clients, the complexity of the vertical relationship we aim to understand is further increased by the peculiar industrial relationships which characterize the organization of this market. In particular, while at the level of single agencies this market appears to be relatively fragmented,[1] most of these agencies actually belong to a few *agency networks* (seven in the U.S.) that implement and sometime coordinate some of the DMAs' activities. Such a market structure thus creates a very rich set of potential channels through which agencies can create value for their clients. We consider various such channels, but we focus especially on the role of market specialization, exclusive contracting and diversification.

The aspects of the market discussed above open the door to a rich set of potential patterns for advertisers' preferences, who might in principle display both complementarities or substitutabilities over the matchings between other advertisers and marketing agencies. An advertiser, for instance, might evaluate either positively or negatively the possibility of hiring an intermediary who also manages the ad campaigns of its rival advertisers: Although sharing a marketing agency might create conflicts – as suggested, for instance, by both theoretical (Villas-Boas, 1994) and empirical (Silk and King III, 2013) studies on the advertising industry before the digital revolution – the benefits in the realm of digital advertising can be disparate, including access to better data and the possibility of bid coordination in the auctions where online ad space is sold.[2]

Our empirical analysis employs a large dataset that links, among the top 6,000 US advertisers, every firm active in the digital advertising market to one (or more) specific advertising agency for all major US industries between 2015 and 2017. Deals between firms and intermediaries are modeled

[1] According to the US Census Bureau's 2016 annual report, there are 12,000 marketing agencies in the US, with 8,000 of them having less than four employees and 10,800 less than 20. The advertising agencies with more than 100 employees are 285 and account for 60% of the employment in the sector.

[2] As shown by Decarolis, Goldmanis and Penta (2020), in the main existing auction formats, bid shading (and, hence, price reduction) is in fact feasible for an intermediary which handles ad purchases on behalf of multiple advertisers that are interested in the same type of ad space.

as the result of a (many-to-many) matching game. Although matching games are commonly used for modeling specific applications such as marriage, school applications, kidney exchanges and employment choices, they are relatively less used for vertical markets where matching between firms takes place along the supply chain. This is likely due to two challenges: first, the lack of systematic data covering the vertical matches in an industry and, second, the inherent complexity associated to analyzing matching problems in settings, such as the present one, in which parties are not constrained to a single match (hence the many-to-many matching framework). In this work, we overcome both difficulties. First, we rely on a novel firm-level dataset we specifically developed for this study, by combining several data sources on the digital ad industry. Second, we exploit the Maximum score estimator proposed by Fox (2018), which is ideal to deal with large markets such as the one we consider. As discussed below, this estimation approach exploits the features of the stability notion (Jackson and Wolinsky, 1996) used to solve the game to avoid the severe curse of dimensionality which would plague other approaches, if applied to large markets. The estimation procedure makes the problem computationally tractable and allows us to structurally estimate the preference parameters for advertisers and intermediaries which determine the matching between them, as well as to evaluate counterfactual changes in the market structure.

The results, albeit still preliminary, clearly show that competing advertisers benefit from dealing with a common intermediary. In particular, the estimates offer a precise quantification of the industry specialization effect for intermediaries, exclusive contracting, diversification and other non evident network phenomena that are usually of interest in the study of industrial organization. Perhaps the most interesting finding is that advertisers seem to prefer not to share *marketing agencies* with their direct competitors, but at the same time they prefer to match to agencies belonging to the same *agency network* as the agencies that manage their competitors. A closer look at the way DMAs operate suggests a simple economic explanation for this result. In particular, while the creative activities and strategic decisions associated to the design and management of the advertising campaigns take place at the level of the marketing agencies, the functions of (algorithmic) bidding and data analysis happen at the level of the agency networks (within specialized units known as "agency trading desks", or ATDs – see Section 3 for a more detailed description of the market's institutional details). Hence, on the one hand, our analysis confirms earlier findings in the literature on common agency phenomena (Silk and King III, 2013): as long as traditional marketing activities are concerned, advertisers still prefer to avoid joining agencies which also manage the campaigns of their direct competitors; on the other hand, our results show that agencies do value the positive spill overs made possible by the modern activities specific to online advertising, such as the data analysis and bidding functions.

It thus appears that the two-layers structure that has emerged in this market – with more traditional functions operated at the agency level, and the more modern functions of algorithmic bidding and data analysis at the higher agency network level – provides a way of balancing the pros and cons from "sleeping with the enemy" (that is, of joining an intermediary also hired by a direct competitor, Villas-Boas (1994)), by closely mirroring the structure of advertisers' preferences as emerged by our estimation: all traditional activities, which the earlier literature had identified as reasons to avoid a common intermediary, are still operated at the agency level; the more modern functions of algorithmic bidding and data analysis, which were expected to generate positive spill overs from sleeping with the enemy, are operated at the higher agency network level.

We conclude with an assessment of the potential effects of mergers between intermediaries. This analysis is motivated by the rapidly increasing concentration among intermediaries observed in recent years, and by the importance of understanding its impact on the advertisers' welfare of advertisers. At this stage, our approach is not a full counterfactual analysis, but more an evaluation of the advertisers' change of surplus under different ownership structures for the agencies. Although our results are particularly preliminary in this part of the study, we believe that this analysis might opens up a new angle and methodology to quantitatively analyze the changing competition landscape in digital markets, taking into account the rich structure due to both the diversity of services offered by the agencies, as well as the specific organizational structure emerged in the industry.

The rest of the paper is organized as follows: Section 2 reviews the relevant literature concerning vertical markets, marketing agencies and matching games. Section 3 describes the institutions and data. Section 4 introduces the theoretical framework. Sections 5 discusses identification. Section 6 presents the estimation approach and the results. Section 7 reports the counterfactual merger evaluation. Section 8 concludes.

2 Related Literature

This work contributes primarily to three branches of the literature. First and most directly, it studies the behavior of advertisers in digital advertising. The economics literature on the digital advertising market is becoming more mature, but it is still primarily focused on the publishers' perspective and on what is the optimal mechanism to sell ad space. This is perhaps the reason why this literature has focused more on search than on display ads. As discussed below, the former entails ad space sold on the result page of search engines, while the latter regards the ad space sold on any other web site (from those of established newspapers to low traffic web pages, all collectively indicated as publishers). The studies on the search ad auctions are closely related to the classical auction theory and mechanism design literature. They usually analyze how allocative efficiency and publishers' revenues change under different mechanism designs or model assumptions. Early papers treat the auctions for ads as classical multi-item auctions (Varian, 2007; Gomes and Kane, 2014), while later works aim to account for peculiar aspects of the search auctions such as the fact that the bids are not updated every time an auction is run (Athey and Nekipelov, 2010) or that ranking positions yield heterogeneous clicks amounts for different keywords (Goldman and Rao, 2014). Another stream of literature takes the perspective of the advertiser and studies how the competition for ad slots affects their revenues (Agarwal and Mukhopadhyay, 2016) or what is the best strategy to maximize revenues while choosing the keyword portfolio under budget constraints (Baardman et al., 2019).

If display advertising is somewhat less studied in the economics literature, it is more popular in the engineering and data science ones. An exhaustive literature review is conducted by Choi et al. (2017). The economics literature on the subject focuses again on the optimal mechanism design to allocate ads by exchanges (McAfee, 2011*a*). Some works enquire the relation between advertisers and other actors by theoretically investigating the optimal types of contracts (Balseiro and Candogan, 2017), but without studying how the players are chosen.

The second branch of literature is the one on vertical markets with intermediaries. Although

middlemen are obviously fundamental in many markets, their presence and influence is often understudied. The most important actions carried out by economic intermediaries are, among others: setting prices and clearing markets, providing liquidity and immediacy, coordinating buyers and sellers, guaranteeing quality and monitoring performance (Spulber, 1996). Recent works in empirical industrial organization try instead to study which role intermediaries have, how they benefit the market, and try to quantify such effects. Since vertical markets often witness upstream actors interacting with many downstream counterparts and vice versa, network models are often employed (Condorelli, Galeotti and Renou, 2016) to capture the nature of the relation. In many cases, the whole network is described according to various measures (such as connectivity, density or nestedness) to shed light on market functioning. There are numerous examples, from macro level, on how value chains of countries predict growth (Hausmann and Hidalgo, 2011), to micro, on how the "nestedness" measure of a value chain in retail can predict failure of firms (Uzzi, 1996). Works of this kind are close to the network science literature and focus on networks as a whole. The current work, on the other hand, explicitly models the economic agents' profit functions, which paired with an equilibrium concept leads to a network formation game (or matching game).[3]

The third and final strand of related literature is that on the estimation of matching games. To our knowledge, Day (2014) was the first to propose matching estimation as a way to analyze the relationship between advertising agencies and their clients, in her case in the context of traditional advertising for pharmaceutical products. Since on both ends of the market there are firms making choices, matching models offer a reasonable modleing framework to estimate agents' preferences over the characteristics of their potential partners.[4] The revealed preference approach for the study of matching games encompasses several different estimation methods deriving from the discrete choice model literature. It is not a surprise then, that maximum likelihood and method of moments estimators, extensively used in the demand estimation literature (McFadden et al., 1973; Train and Weeks, 2005; Berry, Levinsohn and Pakes, 1995), are also the most common choice in matching games by large (Choo and Siow, 2006; Abdulkadiroğlu, Agarwal and Pathak, 2017). These two classes of estimators usually assume the idiosyncratic error term distribution to be known, while they let the deterministic part of the utility function to be flexible.[5] The maximum score estimator we adopt takes the opposite direction: it specifies the deterministic part of the surplus function, and lets the random term be unrestricted. The maximum score estimator used in this works, which was developed by Fox (2018), builds on the seminal work of Manski (1975). It has been applied to many markets involving one-to-many matchings (Chen and Song, 2013; Fox and Bajari, 2013), and to a few which involve many-to-many matchings (Nosal, 2016; Fox, 2018) or even to entry dynamics (Ellickson, Houghton and Timmins, 2013). A fourth estimator is the inequality moment estimator (Pakes et al., 2015) which is similar to the one used here since they both rely on inequality conditions, but differs substantially due to the form taken by the profit function.

The choice of a maximum score estimator in this study was motivated by two reasons. First,

[3]Note that we do not model intermediaries as both seller and buyers. We focus on the downstream section that sees only the advertiser and agencies. Multistage matching models that treat agents as both sellers and buyers are possible in theory but are substantially more complicated in practice (Azevedo and Hatfield, 2018). Translated to the digital advertising industry this means that deals between agencies and other intermediaries along the value chain such as DSPs, Ad exchanges or Search engines are not directly addressed.

[4]Although an important strand of the empirical matching literature focuses on non-transferable utility models – see Agarwal (2017) for a recent overview, for applications related to the vertical links between firms a transferable utility framework is evidently the right approach. We thus follow Fox (2018), as discussed next.

[5]Most common choice is GEV type I, for the same reason provided by demand estimation literature.

the unrestricted distribution of the error term eases the computational problem, which is relevant in large, many-to-many matching markets such as this one. Importantly, many-to-many matching is the core of the work since its objective is to measure complementarities. Second, the advertisers and the agencies bargain for a payment in exchange of a service after choosing each other, but the data on payments are not available. Our model presumes an environment with Transferable Utility (TU), in which transfers are known to the agents but unobserved by the econometrician. Some alternative estimators circumvent the second problem by considering a joint surplus function, but do not provide any solution to the first one (Chiappori and Salanié, 2016).

Papers which study how intermediaries affect the market and the corresponding network formation cover different industries, including healthcare (Ho and Lee, 2019), food and beverage (Berto Villas-Boas, 2007), physical advertising (Donna et al., 2018), etc. Many papers study the vertical relationship through a Nash bargaining model that exploits data on wholesale prices (Donna et al., 2018), intermediaries' marginal costs (Berto Villas-Boas, 2007) or both (Ho and Lee, 2019). The Nash bargaining model is a popular choice because, if such data are available, they can be used to estimate bargaining power parameters across the vertical relation and more importantly to split the generated surplus. Nevertheless, without data on wholesale prices or costs, it is to our knowledge not possible to split the surplus generated by the intermediaries between them and their clients, thereby making the Nash bargaining model less useful in our context. As a matter of fact, difficulties in analyzing vertical markets often arise from the scarcity of these data, which are enterprises' private information. Even though this work does not introduce any econometric methodological innovation, it provides an interesting application of this viable method to analyze externalities without data on transfers.

3 Institutions and Data

3.1 Digital Advertising

The market of digital advertising is rapidly growing. In 2019, according to industry reports by eMarketer, expenses on online ad surpassed for the first time those on print and television combined. Digital advertising can be divided into four main categories: display (or banners), search, video, and social media. The first two segments are older and account for most of the expenditures, while the latter two are a decade younger but rapidly growing in relevance.

Different processes are used to buy and sell ad space depending on the category of digital ad. For all four categories, however, auctions are the key mechanisms used to allocate ad space and determine payments. The details behind these auction systems are complex and evolving over time. Importantly, many kinds of actors are involved beside the advertisers and the publishers (i.e., the owners of the internet page on which the ad space is shown): digital marketing agencies (DMA), demand side platform (DSP), Ad Exchanges, supply side platforms (SSP), ad networks. Market dynamics are further complicated by the fact that boundaries between actors are sometimes blurred and players are vertically integrated.[6]

[6] DMAs are specialized companies that work as the interface between firms marketing departments and publishers, using their expertise on the market. They help make decisions such as the creative content of the ad campaign as well as how to reach the target audience, the best channels to use for the marketing campaign and the other intermediaries to contact. It is fair to say that DMA are so different among them, that a stringent definition is impossible. In order to buy slots and place the ads the DMAs interact with DSPs: these are firms that manage large inventories of ads and own the technology to perform programmatic and real-time purchases of ad space

For Search advertising, things are somewhat simpler than they are for display ad. Search ads show up on the results page of a search engine like Google or Bing after a user queries a keyword. Thus, the only seller for such ad spaces is the search engine itself, which literally runs an auction to allocate the space to advertisers every time a consumer queries a keyword.[7] The main source of complexity is the number of keywords for which any advertiser is allowed to bid. Considering that 63,000 searches every second are submitted to Google, potential advertisers need some DMAs and DSPs to manage their buying process effectively. Sometimes, however, the DSP service is provided by the search engine itself, in a process of disintermediation. Google Ads (previously AdWords) is technically a DSP for ads on Google's result pages, while Bing Ads is the competitor's equivalent serving both Bing (Microsoft) and Yahoo! Search (Verizon).

The reason behind such a heavy intermediation can be found in three characteristics regarding ads and their allocation: (i) the speed at which the auctions must be accomplished; (ii) the minuscule value and high volume of the items that are traded; and (iii) the need to use automated systems for bidding, evaluation, and execution of the trades (McAfee, 2011*b*). In fact the third characteristic follows from the first two. Ads are allocated extremly quickly: Every time visitors click on a web page link, or press a search engine's "search" button, multiple auctions are performed to select the ads that will be loaded on the website or at the top of the results page. Everything is completed in a fraction of seconds. It is clearly impossible for a human economic agent to optimize the bids for such a volume of auctions on items of little value, and therefore the entire process is automated. To adopt the industry specific jargon, the term "Programmatic bidding" refers to the fact that ads are bought and sold by computers without the participation of humans. Programmatic (or automated, or algorithmic) bidding is also performed because the items on sale are difficult to evaluate. In fact, the value of a click on a given ad varies extensively according to who clicks on it. Advertisers therefore, collect meaningful data on potential customers on the internet through proprietary cookies, the IP address of visitors, or simply by acquiring the data from third parties. Once all this information on the target audience is acquired, it is used to optimize the bidding process, selecting the right timing and location of the ads to show. In order to perform real-time bidding, automation is necessary.

3.2 Digital Marketing Agencies

Marketing agencies are the core of the advertising and marketing sector. They have been around since long before the coming of digital advertising, and consist of firms primarily engaged in

for DMAs or advertisers. The major players among DSPs are Mediamath and Simpli.fi (Datanyze). Sometimes DSPs license their technology directly to the DMAs. For those DMAs belonging to an agency network, however, the programmatic technology is typically accessed by delegating bidding to specialized units within the agency networks, known as agency trading desks (ATDs). A concise explanation of programmatic and real time purchasing is given in the next paragraph. Moving on we find Ad Exchanges: these are simply firms that manage a clearinghouse for ads: a market place where all the different types of actors can buy and sell through auction mechanism ad slots. Major ad exchanges are Google DoubleClick, AT&T Appnexus, Rubicon Project. Given the high level of automation, everything takes place at such a speed that it is impossible for buyers to buy without going through a DSP, or for sellers to sell without the help of an SSP. SSPs are the symmetric supply side counterparts of DSP and help publishers to sell the space on their websites in an optimal and automated fashion. The major player in this area is Pubmatic. Finally Ad networks are large wholesale buyers of ads which buy or sell either on Ad exchanges or through direct deals with DSPs, SSPs, publishers and advertisers. The most important player in this group is Google Adsense. Sometimes firms like Google Doubleclick are vertically integrated and act as DSP, Ad exchange and SSP.

[7] Social Media advertisement works in a similar fashion, with the social media platform being the only supplier of space and allocating it through auctions to the publishers, with the only difference being that ads are shown on the personal feed and not among query results.

creating advertising campaigns and placing such advertising in periodicals, newspapers, radio and television, or other media. They traditionally provide multiple services including account management, production of advertising material, media planning, buying (i.e., placing ads), as well as cretive services. According to the US Census Bureau's 2016 annual report, there are 12,000 marketing agencies in the US, with 8,000 of them having less than four employees and 10,800 less than 20. The advertising agencies with more than 100 employees are 285 and account for 60% of the employment in the sector. While this data might suggest a rather fragmented industry, the reality is more nuanced as most agencies belong to just one out of seven large agency networks. These are holdings, sometimes referred to as "DMA Networks": WPP, Publicis Group, Dentsu Inc, Interpublic, Omnicom Group, Havas and MDC. DMA Networks' revenues range from less than 2 billion US dollars for MDC, up to more than 15 billion dollars for WPP, the market leader.

These parent networks own multiple agencies which act as independent companies, offering a wide range of services but remaining ,similar in their nature. While the very concept of a network of agencies was first introduced by Interpublic Group as a way of circumventing the industry's non-spoken rule that forbids a DMA to serve direct competitors, according to Silk and King III (2013) the main advantages deriving from a common affiliation are support programs, technologies, and access to capital. Indeed, it is often the case that agencies belonging to the same network even compete for clients, and it does seem to be crucial for agencies to retain as much independence as they can, in order to appear credible to their clients.

With the advent of digital advertising, both marketing agencies and their networks have evolved. New technologies were developed to aid advertisers buying ad space in the complex enviornments described above. Clear economies of scale and scope could be achieved by centralizing all those activities involving data storage, analysis and bid optimization. In this way, agency networks developed (either internally or through acquisitions) their agency trading desks, and entrusted them with the bidding activities of all the clients under the DMAs in the network. Similarly, independent agencies (i.e., those outside the networks) created their own programmatic systems, with some companies being purely digital in the sense that they only help their clients to bid in the online ad auctions. Overall, DMAs de facto purchase ads on behalf of advertisers.

DMA dimensions vary widely with the type of service they offer. As the market is young, it is also extremely dynamic, with many acquisitions, mergers, new entries and exits every year. Even if we do not have precise statistics on the digital segment, it is likely to mirror the distribution on the whole advertising industry outlined above. Similarly, DMAs clients range from big corporations to very small family-owned enterprises, even if for the former there is a strong case for managing marketing in-house because of the cost. The billing methods used by DMAs are multiple, ranging from fixed fees per ad campaign (the most widely used method), to hourly rates, to a mixture of both. Performance-based billing is not the standard in the industry, and for this reason firms are trying to internalize the service offered by the DMAs, which are sometimes considered to be rather opaque. Another trend occurring in this industry is project-based billing, a phenomenon that could explain why advertisers deal with so many different DMAs during a single year.

The market maintains its dynamism mainly due to two reasons. First, there are very low barriers to entry that prevent to set up a digital marketing firm. The demand for specialists ensures that top-tier talent will always have the option of setting out on their own. This option, coupled with the challenging work environment at many DMAs, makes "working for yourself" an

attractive option to many. Secondly, business reports state that benefits to scale for the digital marketer remains limited. Working with a large monolith may actually be seen as a negative by advertisers. Mergers and acquisitions in this industry have remained steadily high over the last few years, but not due to cost reduction opportunities, which are reported as negligible, but rather driven by the objective of acquiring strategic capability.

3.3 Data

The dataset used in this study has been assembled by Decarolis and Rovigatti (2019) in their study of how intermediary concentration affects Google's revenues from search auctions. Their data combine two sources: a dataset linking advertisers to DMAs (and DMAs to their agency networks) and a second one linking advertisers to sponsored keywords. In this study, we are chiefly interested in the former dataset, Redbooks.

Table 1: Number of Advertisers and DMAs Active Each Year in Each Industry

Year	2014		2015		2016	
Industries	Adv	DMA	Adv	DMA	Adv	DMA
Beauty & Personal care	7	7	6	6	8	9
Apparel & Shoes	17	20	14	15	20	18
Institutions	8	7	11	12	12	13
Automotive & Transport	19	20	23	28	19	24
Jobs & Education	5	5	6	7	8	9
Financial Service	46	33	51	40	61	45
Food & Agriculture	20	22	17	21	28	29
Food & Beverage	7	11	8	7	11	15
Government	5	5	6	6	8	8
Media	15	18	19	22	22	23
Housewares	10	13	13	13	13	12
Industrial	27	25	28	25	28	27
News and Media	51	44	63	54	66	56
Pharma & Health b	14	16	21	23	16	19
Recreation	22	22	21	22	33	29
Restaurants	19	20	23	21	25	24
Retailers & Merchandise	10	12	11	14	11	14
Technology	34	31	39	36	39	34
Telecom	9	13	9	14	9	16
Real Estate	36	31	34	28	37	31
TOT	383	147	431	158	484	166

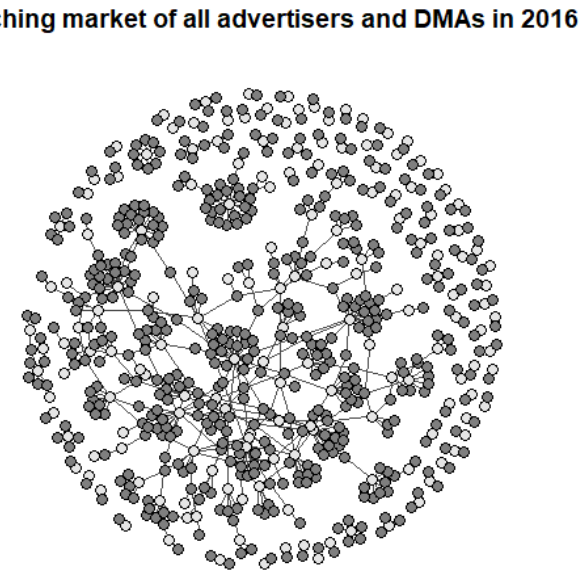
Notes: *Adv* column reports the number of advertisers clients of DMAs in a given market and year. *DMA* column reports the number of DMAs. The number of DMAs is almost always higher than advertisers.

Redbooks is a comprehensive database on marketing agencies: it links a list of advertisers, which represent the largest US corporations active in online marketing, to the full list of marketing agencies that deal with them. Moreover, for each agency it reports whether the DMA is independent or it belongs to one of the seven agency networks, and which one. Redbooks also provides data about the nature of the agencies, indicating whether or not they are involved in digital advertising. For computational tractability, and to focus on the digital advertising sector, we only keep in the sample the DMAs flagged as digital (around 150 per year) and their associated advertisers

(between 400 and 500 per year).[8] Table 1 reports the exact number of advertisers and DMAs, separately for each year (from 2014 to 2016) and advertisers' industry (20 industries), as classified by Redbooks. Every industry in a given year defines a matching market: the pool of potential matches is thus assumed to be that of advertisers and DMAs observed within each year-industry pair. A match is realized if the agents are linked according to the Reedbooks data. In interpreting the numbers in Table 1, note that the same DMAs can appear in multiple industries (in contrast, by construction, advertisers only appear in one industry).

The network structures in Figures 1 and 2 help visualize the data. Figure 1 pools together all the matchings observed across all industries in 2016. We do not use this type of data structure and, instead, further partition the market into year-industry combinations: Figure 2 pools exclusively advertisers and DMAs active in the pair (2016, Real Estate). It describes the type of data structure that we use. Both graphs indicate the presence of two types of DMAs: on the edges of the graph, we notice a universe of independent DMAs dealing with one or a few clients, while, at the center, the biggest component of the network reunites more interconnected DMAs dealing with multiple clients. There are plenty of advertisers who match to more than one DMA and plenty of DMAs matching with more than one advertiser. Although the network in Figure 2 is simpler than that in Figure 1, it still presents a many-to-many matching structure.

Figure 1: Connections between all advertisers and DMAs active in 2016

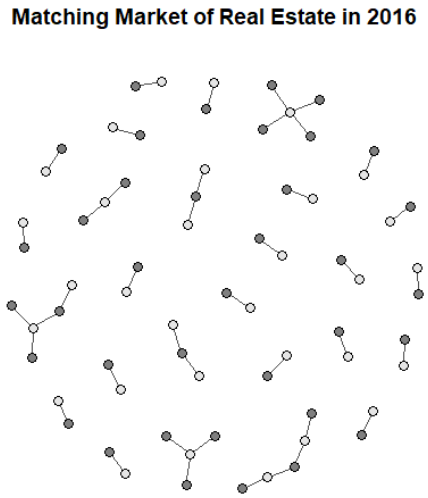


Notes: advertiser in grey, DMAs in white. The figure is a bipartite graph with advertisers only connected to DMAs and viceversa. A link between nodes represents a match.

Observing that advertisers use multiple DMAs is not surprising. In fact, the advertisers in

[8] To focus on a more homogeneous set of advertisers, in every year the sample includes only advertisers matched to at least one DMA.

Figure 2: Matching market of Real Estate in 2016



Notes: advertiser in grey, DMAs in white. The figure is a bipartite graph with advertisers only connected to DMAs and viceversa. A link between nodes represents a match.

the Redbooks data are large corporations, needing different types of marketing services and active across multiple markets. DMAs, in turn, are specialized both by the types of services they offer and by the markets in which they have expertise. Indeed, an industry (e.g., real estate) will typically contain multiple separate markets. Moreover, it is not uncommon that advertisers deal with DMAs for a single marketing campaign, like the launch of a specific product, and therefore deal with many DMAs throughout the same year.

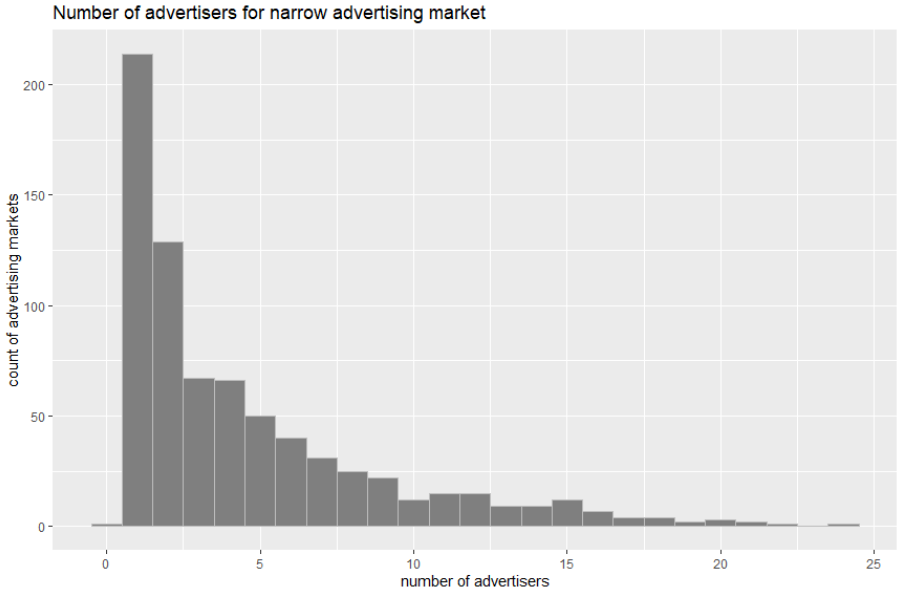
Decarolis and Rovigatti (2019) combine the Redbook data are combined with those from SEMrush covering on Google's sponsored search auctions held in the US market in the same time period. SEMrush is one of the main providers of online advertising data and its dataset allows observing multiple variables for each keyword, including the set of advertisers winning an ad slot, the search volume, and the average prices (i.e., cost-per-click). Decarolis and Rovigatti (2019) propose a methodology to cluster together keywords in a way that the clusters resemble a market: each cluster contains keywords that are related for consumers and for which advertisers are in competition for the scarce ad space.[9] These clusters represent the markets where competition takes place for the ad slots. For example, keywords like "camera lens,""professional camera" and "photography kit" belong to the same thematic cluster, because people searching for "camera lens" are potentially also interested in purchasing a "photography kit" and, as a consequence, it is likely that advertisers sponsor competing or similar products by bidding on the same keywords. Their proposed clustering measure is a crucial input for our own work, as it narrowly defines where each advertiser is interested in placing their ads and tries to define a relevant market for them. Two

[9]They use an unsupervised learning algorithm to represent keywords as numerical vectors (keyword vectorization using the GloVe algorithm) and group the vectorized keywords into clusters according to their semantic similarity (thematic clustering).

caveats are important. First, we shall stress that we are considering competition for ad spaces and not between the products object of the ad. Hence, as far as advertisers are in competition for the limited ad space, the fact that both substitutes and complements are likely to be pooled together is not a problem. That is, even firms that sell rather different products might be competing to grab the attention of the same consumer as he browses the web. Second, albeit these markets are constructed using Google's sponsored search auctions alone, the paramount importance of this form of online advertising ensures that we can use these markets to approximate more generally the degree of overlap between advertisers. Indeed, it is reasonable to assume that advertisers competing for search ad space will also be competing with display ad and social media ad spaces, since the target audience and the type of ad will be similar. Video and audio ads might, instead, represent a somewhat distinct market, but we have no data to address them separately.

Figure 3 shows a typical distribution of the number of advertisers that bid in each cluster (or, simply, market). In particular, the example shows the distribution for the real estate industry in 2016, but similar distributions characterize the other matching markets. Markets are quite specific: several of them are participated by 1 advertiser only, while very few are participated by more than 10. In turn, the typical advertiser active in the Real Estate industry in 2016 enters 100 markets, and 70% of advertisers are active on a number between 50 and 150 markets.

Figure 3: Number of clusters ("markets") by the number of active advertisers (2016, Real Estate)



Notes: the histogram shows how the markets in the (2016, Real Estate) matching market are populated by different numbers of advertisers. Two advertisers active in the same market are considered as competitors in their purchase of internet ad space.

Before moving on, it is important to remark once more the different notions of markets and matching markets. Both can be described as collections of advertisers ad DMAs. However, while a *matching market* is the set of all advertisers and DMAs which are active in an year-industry pair, *markets* are composed by those sets of advertisers and DMAs which are active on at least one keyword belonging to the clusters as identified by Decarolis and Rovigatti (2019). The usage of these two notions will be different: markets will be used to build measures that define, for

any advertiser, which are its competitors among other advertisers; while matching markets will be used to determine for each advertiser and agency pair, which are their potential matches.[10] This assumption implies that if an advertiser wishes to choose a different DMA, he is only allowed to choose among DMAs matched with advertisers in the same industry and active in that year. Given the narrow specialization of DMAs relative to the broad notion of an advertiser's industry, this is a rather innocuous assumption, but one which greatly reduces the computational burden.

4 Model

4.1 Model

The theoretic model presented in this section is a substantial simplification of Azevedo and Hatfield (2018). Most importantly, for tractability and ease of exposition we switch from their continuum-of-agents setup to one with a discrete agent set, borrowing some of the notation and definitions from the networks model in Jackson and Wolinsky (1996). This simplification raises important questions regarding the existence and uniqueness of equilibrium (the general existence and uniqueness results of Azevedo and Hatfield (2018) no longer apply in the present setup). For now, we leave these questions aside and focus on estimation of preferences under the assumption that the observed market is in equilibrium.

Consider a matching market with given sets of advertisers and DMAs.[11] Let $[I]$ be the set of advertisers in the market, indexed by i, and let $[J]$ be the set of DMAs, indexed by j.[12] Each possible matching is then a bipartite graph on $[I]$ and $[J]$. Because the node set remains fixed, we identify each matching with its edge set, which is a subset of $[I] \times [J]$. We refer to members of matchings as matches. Given a matching Φ, for each $i \in [I]$ we let Φ_i be the set of DMAs matched with advertiser i ($\Phi_i = \{j \in J \mid (i,j) \in \Phi\}$). Similarly, for each $j \in [J]$, we let Φ^j be the set of advertisers matched to DMA j ($\Phi^j = \{i \in I \mid (i,j) \in \Phi\}$). We allow many-to-many matchings (each advertiser can be matched to multiple DMAs and vice versa), but prohibit unmatched agents.[13] Thus, the set of all possible matches is $\Omega = \left\{\Phi \in \mathcal{P}([I] \times [J]) \,\middle|\, \Phi_i \neq \emptyset \,\forall i \in [I] \text{ and } \Phi^j \neq \emptyset \,\forall j \in [I]\right\}$. Following Jackson and Wolinsky (1996), we will use "ij" as a shorthand for (i,j), "$\Phi + ij$" as a shorthand for $\Phi \cup \{(i,j)\}$ and "$\Phi - ij$" as a shorthand for $\Phi \setminus \{(i,j)\}$.

Each advertiser and each DMA derives some inherent value from each possible matching. Let advertiser i's matching valuation function be $\nu_i : \Omega \to \mathbb{R}$ and let DMA j's matching valuation function be $\nu^j : \Omega \to \mathbb{R}$. We allow these valuation functions to depend in a nonlinear way on the whole matching; it is not assumed that they are additively separable in valuations of single matches. In particular, the valuation $\nu_i(\Phi)$ may depend not only on the set of DMAs that advertiser i is matched to, but also on the identities of other advertisers matched to these DMAs (and similarly for the other side of the market).

For example, let us assume that there are three advertisers named a, b, c and three DMAs d, e, f. a is matched with d and e while both b and c are matched with f. Let us call this matching

[10] Dynamic behaviour such as entry and exit from the market is not taken into consideration, therefore the time dimension is left unexploited.

[11] In the full model to be estimated, there will be M different matching markets, indexed by m. However, we will analyze each market separately and hence present the model for a single market.

[12] Here and throughout the paper we use $[n]$ to denote the set of natural numbers from 1 to n inclusive.

[13] This assumption derives from the data itself. Alternative models allow independent advertisers to be matched with empty sets (Nosal, 2016).

A: $A = \{ad, ae, bf, cf\}$. Now consider an alternative matching B, which is exactly the same as A apart from the fact that b also deals with e, i.e., $B = \{ad, ae, be, bf, cf\}$. The valuation function of a may take different values under matchings A and B even though the matches involving a are the same in both. This is because in matching B an additional advertiser, b, is matched to e, an agency associated with a. It is easy to see that these valuation function features are necessary to capture complementarities.

In addition, for each matching $\Phi \in \Omega$ we define an $I \times J$ "transfers matrix" T^Φ, whose (i,j)-th entry, t^Φ_{ij}, is the payment that advertiser i makes to DMA j when the matching is Φ. We assume that advertisers make payments only to agencies to which they are matched, so that $t^\Phi_{ij} = 0$ whenever $ij \notin \Phi$. Thus the set of valid transfers matrices for the matching Φ is $\mathcal{T}^\Phi = \left\{T^\Phi \in \mathbb{R}^{I\times J} \,\middle|\, j \notin \Phi_i \implies t^\Phi_{ij} = 0\right\}$.

The overall profits are assumed to be quasilinear. In particular, advertisers' and DMAs' profit functions are $\pi_i, \pi^j : \left\{(\Phi, T^\Phi) \in \Omega \times \mathbb{R}^{I\times J} \,\middle|\, T^\Phi \in \mathcal{T}^\Phi\right\} \to \mathbb{R}$, given by:

$$\pi_i(\Phi, T^\Phi) = \nu_i(\Phi) - \sum_{j\in\Phi_i} t^\Phi_{ij}; \tag{1}$$

$$\pi^j(\Phi, T^\Phi) = \nu^j(\Phi) + \sum_{i\in\Phi^j} t^\Phi_{ij}. \tag{2}$$

Since payments flow from one side to the other, the joint profit of advertisers and DMAs is not affected by the transfers:

$$\sum_{i\in I} \pi_i(\Phi, T^\Phi) + \sum_{j\in J} \pi^j(\Phi, T^\Phi) = \sum_{i\in I} \nu_i(\Phi) + \sum_{j\in J} \nu^j(\Phi)$$

Similarly, for advertiser i matched to DMA j,

$$\pi_i(\Phi, T^\Phi) + \pi^j(\Phi, T^\Phi) = \nu_i(\Phi) + \nu^j(\Phi) - \sum_{k\in\Phi_i\setminus\{j\}} t^\Phi_{ij} + \sum_{k\in\Phi^j\setminus\{i\}} t^\Phi_{ij} \tag{3}$$

because the transfer from i to j cancels out.

4.2 Equilibrium Concept

Our equilibrium concept is pairwise stability. A matching is pairwise stable if it cannot be blocked by any pair of agents $(i,j) \in [I] \times [J]$ or any individual agent $i \in [I]$ or $j \in [J]$. Assuming that forming a new match requires both parties to agree, whereas an existing match can be severed unilaterally by either party, a matching can be blocked either by (1) an agent profitably breaking one or more of its existing matches, or (2) a pair of agents not matched to each other profitably forming a match between them (and possibly dissolving one or more of their existing matches with others). If neither of these is possible, the matching is pairwise stable. To complete the definition in our transferable utility setting, we must specify how the transfers will change in a blocking deviation. We postulate that the formation or dissolution of match ij will not affect any transfers other than the one between i and j, i.e., for all $kl \neq ij$, $t^{\Phi-ij}_{kl} = t^\Phi_{kl}$ and $t^{\Phi+ij}_{kl} = t^\Phi_{kl}$. Note that this assumption uniquely pins down $T^{\Phi-ij}$ given T^Φ (because $t^{\Phi-ij}_{ij} = 0$, as $ij \notin \Phi - ij$).

Given valuation functions $(\nu_i)_{i\in[I]}$ and $(\nu^j)_{j\in[J]}$, a matching $\Phi \in \Omega$ with transfers matrix T^Φ

is stable if and only if the following two conditions hold:

No agent blocks by dissolution: $\forall i \in I, \forall K \subset \{ij \in \Phi\}$:
$\pi_i(\Phi, T^{\Phi}) \geq \pi_i(\Phi \setminus K, T^{\Phi \setminus K})$ and $\forall j \in J, \forall K \subset \{ij \in \Phi\}, \pi^j(\Phi, T^{\Phi}) \geq \pi^j(\Phi \setminus K, T^{\Phi \setminus K})$

No pair of agents blocks by rematching: $\forall ij \notin \Phi, \forall K \subseteq \{ik \in \Phi\} \cup \{kj \in \Phi\}, \forall T^{\Phi+ij \setminus K}$ s.t. $\forall kl \neq ij, t_{kl}^{\Phi+ij \setminus K} = t_{kl}^{\Phi \setminus K}$:
$\pi_i(\Phi + ij \setminus K, T^{\Phi+ij \setminus K}) > \pi_i(\Phi, T) \implies \pi^j(\Phi + ij \setminus K, T^{\Phi+ij \setminus K}) < \pi^j(\Phi, T)$ and
$\pi^j(\Phi + ij \setminus K, T^{\Phi+ij \setminus K}) > \pi^j(\Phi, T) \implies \pi_i(\Phi + ij \setminus K, T^{\Phi+ij \setminus K}) < \pi_i(\Phi, T)$

As transfers are not observable to the researcher, some manipulation on the definition of pairwise stability is needed to obtain a characterization that does not contain them. This characterization will be employed for estimation. With this aim in mind, let's consider a matching Φ with $ij, i'j' \in \Phi$ and at the same time $ij', i'j \notin \Phi$. Moreover, let's define $\bar{t}_{ij'}$ as the transfer that would make j' indifferent between the current matching and one where the matches $i'j'$ and ij are replaced with ij'. Isolating the transfer from the profit equivalence:

$$\bar{t}_{ij'} = \nu^{j'}(\Phi) + t_{i'j'}^{\Phi} - \nu^{j'}(\Phi - ij - i'j' + ij').$$

Now, let $T^{\Phi - ij - i'j' + ij'}$ be obtained from T^{Φ} by setting the ij and $i'j'$ terms to zero and setting the ij' term to $\bar{t}_{ij'}$ (this clearly satisfies the condition for a valid transfer matrix in the definition of stability). Because $\pi^{j'}(\Phi - ij - i'j' + ij', T^{\Phi - ij - i'j' + ij'}) = \pi^{j'}(\Phi, T^{\Phi})$, the matching Φ will be jointly blocked by j' and i unless $\pi_i(\Phi - ij - i'j' + ij', T^{\Phi - ij - i'j' + ij'}) \geq \pi_i(\Phi, T^{\Phi})$, i.e.,

$$\nu_i(\Phi) - t_{i,j}^{\Phi} \geq \nu_i(\Phi - ij - i'j' + ij') - \bar{t}_{ij'}$$

Let's now plug the definition of $\bar{t}_{i,j'}$ into the latter inequality:

$$\nu_i(\Phi) + \nu^{j'}(\Phi) - t_{ij} + t_{i'j'} \geq \nu_i(\Phi - ij - i'j' + ij') + \nu^{j'}(\Phi - ij - i'j' + ij').$$

One way to read the inequality above is that the joint surplus deriving from matching Φ for i and j' with the respective transfer is greater than the joint surplus of i and j' where they deal with each other and transfers cancel out. Now let's repeat the same reasoning for the remaining couple involved, namely advertiser i' and DMA j, using the alternative match $i'j$. The result is the inequality

$$\nu_{i'}(\Phi) + \nu^{j}(\Phi) + t_{ij} - t_{i'j'} \geq \nu_{i'}(\Phi - ij - i'j' + i'j) + \nu^{j'}(\Phi - ij - i'j' + i'j).$$

Finally, summing over the two inequalities, all the transfers cancel out and we have the following inequality:

$$\begin{aligned}
&\nu_i(\Phi) + \nu^{j}(\Phi) + \nu_{i'}(\Phi) + \nu^{j'}(\Phi) \\
&\geq \nu_i(\Phi - ij - i'j' + ij') + \nu^{j}(\Phi - ij - i'j' + i'j) + \nu_{i'}(\Phi - ij - i'j' + i'j) + \nu^{j'}(\Phi - ij - i'j' + ij').
\end{aligned}$$

The above inequality will be referred to as the sum of revenues inequality and will be paramount for estimating parameters. One such inequality can be computed for each pair of matches $ij, i'j' \in \Phi$ such that $ij', i'j \notin \Phi$.

5 Identification

5.1 Background and Econometric Assumptions

We follow the identification and estimation strategy of Fox (2018). Each agent has both observable and unobservable attributes. The observable attributes take values in the finite sets H (for advertisers) and $\hat{H}$ (for DMAs). The unobservables take values in the possibly infinite sets K (for advertisers) and $\hat{K}$ (for DMAs). Each agent ($i \in I$ or $j \in J$) is fully identified by a corresponding pair $(h, k) \in H \times K$ (or $(\hat{h}, \hat{k}) \in \hat{H} \times \hat{K}$).

The unobservables are invisible not only to the researcher, but also to other agents. Consequently, it makes sense to assume that an agent's valuation of a matching will not depend on these attributes.

Assumption 1: Valuations ν_i and ν^j do not depend on the unobservable attributes of agents other than i (respectively, j).

For example, if DMAs' past revenues are not in the data, then Assumption 1 means that advertisers do not have valuations over DMAs' past revenues.

The valuations ν_i and ν^j do depend on the unobservable attributes of i and j (the past revenue of the advertiser, for example, influences their own choice over the DMAs). However, it is assumed that valuations can be decomposed into a deterministic part that depends only on the observables of the evaluating agent and a stochastic component that is determined by the unobservables. That is, when agent i has combined type $(h, k) \in H \times K$ and agent j has type $(\hat{h}, \hat{k}) \in \hat{H} \times \hat{K}$, we have

$$\nu_i(\Phi) = \bar{\nu}_h(\Phi, \theta) + \epsilon_k(\Phi);$$
$$\nu^j(\Phi) = \bar{\nu}^{\hat{h}}(\Phi, \hat{\theta}) + \epsilon^{\hat{k}}(\Phi),$$

where:

- $\bar{\nu}_h(\Phi, \theta)$ (resp., $\bar{\nu}^{\hat{h}}(\Phi, \hat{\theta})$) is the deterministic part of the valuation function. It is a parametric function of the observable attributes h (resp., $\hat{h}$) and a vector of parameters to be estimated, θ (resp., $\hat{\theta}$);
- $\epsilon_k(\Phi)$ (resp., $\epsilon^{\hat{k}}(\Phi)$) is the unobservable or stochastic valuation part, which depends on the unobservable attributes k (resp., $\hat{k}$). The distribution of the unobserved term $\epsilon^k(\Phi)$ depends on the matching Φ, but, as previously stated, not on the unobservables of agents other than i (resp., j).

Assumption 2: Let Φ be a matching. If $i'j \in \Phi$ and $ij, i''j \notin \Phi$ then $\nu_i(\Phi) = \nu_i(\Phi - i'j)$ and $\nu_i(\Phi) = \nu_i(\Phi + i''j)$. Similarly, if $ij' \in \Phi$ and $ij, ij'' \notin \Phi$ then $\nu^j(\Phi) = \nu^j(\Phi - ij')$ and $\nu^j(\Phi) = \nu^j(\Phi + ij'')$.

The above assumption states that if advertiser i is not matched with DMA j, i is indifferent if there is a match between j and other advertisers i' and i'', and j is indifferent if there is a match between i and other agencies j' and j''. This assumption is necessary to limit the way a valuation changes over relatively similar matchings. It is nevertheless a weaker assumption than others used

in the literature such as the assumptions in Chen and Song (2013), where the valuation of agent i is affected only by i's own matches Φ_i. This assumption is necessary if the researcher is interested in studying complementarities (Baccara et al., 2012; Nosal, 2016).

Assumption 3: The deterministic components of the valuation functions are linear in the parameters θ (resp., $\hat{\theta}$):

$$\bar{\nu}_h(\Phi, \theta) = X(h, \Phi)'\theta;$$
$$\bar{\nu}^{\hat{h}}(\Phi, \hat{\theta}) = \hat{X}(\hat{h}, \Phi)'\hat{\theta},$$

where $X(h, \Phi)$ and $\hat{X}(\hat{h}, \Phi)$ are vectors of observables and θ and $\hat{\theta}$ are the vectors of parameters to be estimated.

5.2 Identification

First, consider the stochastic terms of the value functions, ϵ_k and $\epsilon^{\hat{k}}$, which depend only on the unobservable attributes k (resp. $\hat{k}$). Considering both the observables h (resp., $\hat{h}$) and the unobservables k (resp. $\hat{k}$), we have conditional distributions of the stochastic component given the observable attributes. We denote these by $P(\epsilon_k|h)$ and $\hat{P}(\epsilon^{\hat{k}}|\hat{h})$. The distributions P and $\hat{P}$ are not known and are not assumed to be the same for all h (resp., $\hat{h}$), so that heteroskedasticity is allowed.

The conditional distributions of ϵ_k and $\epsilon^{\hat{k}}$ is necessary to describe the probability that a given match occurs, since matches are chosen according to the valuation functions, which in turn depend on ϵ_k and $\epsilon^{\hat{k}}$. Indeed, the choice probability for agents with attributes h is:

$$Pr_h(\Phi) = \int_{\epsilon_k} \mathbb{1}\left[\Phi \in \arg\max_{\tilde{\Phi}}(X(h, \tilde{\Phi})'\theta + \epsilon_k)\right] dP(\epsilon_k|h) \tag{4}$$

The last formula is the usual choice probability employed by the multinomial discrete choice models literature to estimate demand side and sometimes termed as the market share equation (McFadden et al., 1973). It states that the probability that agents with observables h choose a given matching depends on how many times the matching maximizes the valuation function, when we integrate out the unobservable stochastic term (given h). Therefore, also Φ is stochastic and the source of uncertainty derives from ϵ_k. As the discrete choice literature indicates, the deterministic part of the valuation $X(h, \Phi)'\theta$ can be interpreted as the "mean" valuation for agents described by h. A crucial difference between this model and the standard discrete choice models in the literature originates from the "error" distribution, since usually ϵ^k follows an extreme value or some other known distribution while here it is left almost completely unrestricted. The crucial assumption on the stochastic component is the following:

Assumption 4 (Rank Order): let Φ and Φ^* be two matchings such that $ij, i'j' \in \Phi$, $i'j, ij' \notin \Phi$ and $\Phi^* = \Phi - ij - i'j' + i'j + ij'$. Let the observable types corresponding to i, j, i', and j' be,

respectively, h, h', $\hat{h}$, and $\hat{h}'$. Then:

$$Pr(\Phi) > Pr(\Phi^*) \iff$$
$$\bar{\nu}_h(\Phi) + \bar{\nu}_{h'}(\Phi) + \bar{\nu}^{\hat{h}}(\Phi) + \bar{\nu}^{\hat{h}'}(\Phi) > \bar{\nu}_h(\Phi^*) + \bar{\nu}_{h'}(\Phi^*) + \bar{\nu}^{\hat{h}}(\Phi^*) + \bar{\nu}_{\hat{h}'}(\Phi^*).$$

(Note that, for ease of exposition we have omitted the arguments θ and $\hat{\theta}$ from the $\bar{\nu}$ functions above.)

The assumption simply states that between two alternative matchings one has more probability to be observed in the data if and only if the deterministic components of the joint surplus obtained by advertisers and DMAs under that matching is greater than under the alternative. This assumption is necessary to make estimation for large matching markets computationally tractable. The Rank order assumption lays at the core of the maximum score identification strategy as it enables us to work directly with inequalities instead of solving the discrete choice integral. The rank order assumption explains how the unobservables enter in the choice model and it is sufficient to point identification of the parameters if at least one element of $X(h, \Phi)$ has $\mathbb{R}$ as support.

Identification of the maximum score estimator relies on inequality relations rather than equalities. The inequality appearing in the Rank order assumption allows us to characterize pairwise stability by comparing the total deterministic valuation functions arising from two alternative matchings.

From all the possible inequalities, only those that include actual matchings observed in the data are taken into consideration. Suppose Φ is the matching observed and Φ^* is the alternative matching (where Φ and.Φ^* are as in the statement of Assumption 4). As before, let the observable types corresponding to i, j, i', and j' be, respectively, h, h', $\hat{h}$, and $\hat{h}'$. The inequality that compares the joint valuations of the observed matching on the left and the alternative matching on the right is:

$$\begin{aligned}&\left((X(h,\Phi) + X(h',\Phi))'\theta + (\hat{X}(\hat{h},\Phi) + \hat{X}(\hat{h}',\Phi))'\hat{\theta}\right) > \\ &\left((X(h,\Phi^*) + X(h',\Phi^*))'\theta + (\hat{X}(\hat{h},\Phi^*) + \hat{X}(\hat{h}',\Phi^*))'\hat{\theta}\right).\end{aligned} \tag{5}$$

Rearranging terms and grouping parameters common for all regressors we obtain:

$$\begin{aligned}&\left(X(h,\Phi) + X(h',\Phi) - X(h,\Phi^*) - X(h',\Phi^*)\right)'\theta + \\ &(\hat{X}(\hat{h},\Phi) + \hat{X}(\hat{h}',\Phi) - \hat{X}(\hat{h},\Phi^*) - \hat{X}(\hat{h}',\Phi^*))'\hat{\theta} > 0.\end{aligned} \tag{6}$$

All the selected inequalities then enter a matching maximum score function, which constitutes the objective function to be maximized to estimate the true parameters. Denote the left side of the equation (6) with $Z(\theta, \hat{\theta})$ and let G be the set of inequalities arising from the observed matching, $Z(\theta, \hat{\theta}) \in G$. Then the Maximum score is:

$$F(\theta, \hat{\theta}) = \sum_G \mathbb{1}\left[Z(\theta, \hat{\theta}) > 0\right] \tag{7}$$

It is clear that the objective function does nothing more than counting true inequalities, i.e., inequalities where the observed matching is more profitable than the alternative.

Which type of parameters can be identified with such a strategy? Since identification relies on inequality between different matchings, every variable which does not differ along matchings cancels out. As a result, parameters cannot be identified for fixed attributes of players.[14] Let's take once again the advertisers and DMAs past revenue as an example: the past revenue of a single firm will be exactly the same on the left and right side of the inequality (5), making the related parameter unidentifiable. What can be identified are parameters that multiply interactions between advertisers and DMA attributes: the product (not sum) of the number of advertiser's employees and the matched DMA's employees is different for each matching, making the associated parameter identifiable, even if of little value. Interactions are a simple subset of variables that change along with the matching. Moreover it is worth noticing that simple interactions between attributes are included in one side's valuation rather than in the other just for convenience. Since these variables are symmetric by construction, they are not attributable to advertiser or DMA, but are simply informative on the joint surplus generated.

Finally, since the inequality holds true after being multiplied on both sides by a constant factor, a scale normalization on the parameters is required. This is achieved by fixing to ± 1 a single parameter θ_0. What is important is the restriction on the absolute value $|\theta_0| = 1$, while the sign can be estimated from the data. To sum up, the identification is obtained by simulating a great number of exogenous shocks at the matching assignment level. These shocks on the assignment allow us to consider different configurations that show why the observed matching actually emerges. For a more in-depth discussion, see Fox (2018).

6 Estimation

The parameters of the advertisers and DMAs' profit functions are estimated by maximizing the Maximum score objective function presented above.[15] Relative to the rapidly growing literature on the estimation of matching games, the two characterizing features of our estimation approach are to allow for many-to-many matching and for profit functions that are not additively separable on matches, but which depend instead on the whole matching (as it is necessary in order to account for complementarities).

There is a simple intuition behind the parameters' estimation: estimates are those parameter values that, once plugged into the inequalities defining the stability of the observed match, lead to the greatest number of them being satisfied. Note that, even if the parameters were the true ones, not all the inequalities would be satisfied because of stochastic terms ϵ^k. The intuition is that the true parameters describe the deterministic part of valuation that enters the inequality, but the observed matching is determined also by ϵ^k. Assumption 4, Rank Order, states that the observed matching makes the inequality likely to be satisfied, not necessarily satisfied with probability 1.

Starting from the observed matching Φ, any possible pair of matches is selected leading to a total $s_m(s_m - 1)/2$ pairs, where s_m is the number of matches counted in the observed matching Φ in matching market m. In every pair of matches, up to four players are involved. Consider, for

[14] This is analogous to the situation in discrete choice model's estimation, where any element of the utility that is invariant across choices cannot be estimated.

[15] Specifically, we implement an estimation algorithm in R starting from the MSE-R code by Theodore Chronis and Christina Tatli, `https://github.com/tatlchri/MSE-R`. Their code builds upon the maximum score estimator for matching data developed by Jeremy Fox in a series of studies, (Fox, 2010, 2018; Santiago and Fox, 2009).

instance, a case with two different advertisers (i, i') matched to two different DMAs (j, j'). For every pair of matches, a new alternative matching can be formed by switching agents' matches. For pair $ij, i'j' \in \Phi$, the alternative match is $\Phi^* = \Phi - ij - i'j' + ij' + i'j$. In the end we obtain $s_m(s_m - 1)/2$ alternative matchings which will be evaluated against the observed ones, or, equivalently, we will evaluate $s_m(s_m - 1)/2$ inequalities for each market.

The counterfactual matching used in every inequality is created in such a way that it includes the exact same number of matches of the observed matching Φ, since they are exchanged but never destroyed or created. We select the entire set of inequalities G obtaining 350,000 inequalities for the estimation.[16] The asymptotic properties of the estimator are such that increasing the number of observations from a single matching market or increasing the number of matching markets observed is equivalent for the purpose of the estimator's consistency. In fact, for both arguments we assume one continuum of agents, that is progressively revealed as the number of observations or matching markets grows large. In this case, even if the number of observation per matching market is not high, the large number of matching markets ensures consistency. Since the data covers 20 industries for 3 years, we have a total of 60 different matching markets.

For every theoretical inequality selected, the vector of regressors $X(h, \Phi)$ is computed and the empirical inequality $Z(\theta)$ is created. Maximizing the objective function, denoted by F, means evaluating it at each parameter θ, which boils down to plugging in the parameter in the first inequality and check whether it is satisfied. If it is, the score is increased by 1, otherwise it does not change. Hence, the sum of scores over each market is the value of F for a given θ:

$$F_M(\theta) = \frac{1}{M} \sum_{m \in M} \sum_{G} 1[Z(\theta) > 0] \tag{8}$$

The Maximum score estimator is: $\hat{\theta} = \arg\max_\theta F_M(\theta)$. Because the objective function is a sum of indicator functions, it is not smooth in parameters. For this reason, optimization algorithms that rely on differentiation are ill-suited for the task. Although a smooth maximum score function has been suggested (Horowitz, 1992) to bypass such a problem, we keep the original non-smooth function and maximize it by relying on a differential evolution algorithm.[17]

Finally, since the distribution of errors is not known, it is necessary to infer parameter confidence intervals directly from their empirical distribution. This is done through the subsampling procedure presented by Politis and Romano (1994). Subsampling is equivalent to the bootstrap procedure, apart from the fact that each sampling is done without replacement and is of size $b \ll g$, where g is the size of the original sample (i.e., the set G used to estimate the parameters). More precisely, the subsampling procedure needs the key assumptions that b converges to infinity at

[16] Following Fox (2018), a subset of inequalities can be chosen to enhance the computational feasibility of the estimation. In doing so, every possible inequality must be chosen with the same probability. Since there is no method to pick an optimal number of inequalities, we follow Fox (2018) in sampling 300,000 inequalities.

[17] A well known limitation of this approach is that estimates are sensible to optimization algorithm choice, especially when optimization problems are complex. (Knittel and Metaxoglou, 2014). We employ the DEoptim algorithm. The robustness of the optimization routine is checked by trying several control parameters. DE belongs to the class of genetic algorithms which use biology-inspired operations of crossover, mutation, and selection on a population in order to minimize an objective function over the course of successive generations (see Mitchell, 1998). As with other evolutionary algorithms, DE solves optimization problems by evolving a population of candidate solutions using alteration and selection operators. For the choice of control parameters, we follow the advice in Storn and Price (1997) setting initial population number (100), crossover probability (0.5) and weighting factor (0.8) and Maximum iterations (200). Boundaries are between -15 and 15. Increasing the number of iterations up to 500, lowering crossover probability and increasing weighting factor does not affect estimates. An extended set of robustness checks involving the computational aspects is presented in the appendix.

lower rate than g. Since the literature does not point to any suggested size for b, we use 10% of total inequalities for each subsample. These are drawn without replacement and used to estimate the parameter. The subsampling procedure is repeated 50 times. The empirical distribution of parameters derived from these 50 subsamples is then used to compute 95% confidence intervals. Subsampling is executed both at market level and firm level, showing no significative difference in confidence interval ranges.

6.1 Profit Functions Specification

As discussed in the identification section, single attribute features not varying across advertisers-DMA pairs cannot be identified and are not included among the elements of X. All these features of the profit functions are thus captured by the stochastic term. The variable that we can include in X must all come from interactions between features of advertisers and agencies. For advertisers, the variables entering the profit function are *comp*, *compnet*, *diversif* and *exclusive*. The first variable, *comp*, is an intensity measure of competition among advertisers dealing with a common DMA. The variable is computed as follows: first, the number of advertising markets where advertiser i is active are selected. Second, we repeat the same procedure for every other advertiser connected to the DMA j. Then, we compute the share of common markets between i and other advertisers connected to it through the DMA j. The shares are the number of common markets divided by the number of participated markets. Finally, all these shares are summed. Let us remind that the markets here refer to advertising markets and not matching markets. Formally, *comp* is calculated as follows: denote by i the advertiser and index by $l = 1, 2, ..., n$ all other advertisers matched to DMA j, excluding i, by $\#market_i$ the number of markets where i is active and by $\#market_{li}$ is the number of markets shared by l and i, then:

$$comp = \left(\frac{\#market_{1i}}{\#market_i}\right) + ... + \left(\frac{\#market_{ni}}{\#market_i}\right).$$

This variable varies across all agency-DMA pair, it is not symmetric for two advertisers matched to the same DMA and can be computed for both observed and counterfactual matches. These are features that will characterize also the other elements of the profit functions presented below. Before that, however, notice that from an economics' perspective *comp* proxies well what an advertiser should consider of a potential DMA match when focusing on the extent to which this DMA is working with rival advertisers. *comp* is indeed constructed as a sum to capture every possible opportunity to overlap with competitors. In this regard, notice that variables increase when an advertiser active on many markets matches with DMAs with many clients only to the extent that these clients are active on at least some common markets. Lastly, we shall argue why it is reasonable to assume that advertisers observe the set of DMAs' clients. It turns out that most DMAs openly advertise, for instance on their web sites, their portfolio of clients. Not to mention the fact that the Redbook data exist precisely because advertisers value the possibility of learning about the advertisers-agencies connections in the market.

The second variable in the advertisers' profit function is *compnet*. This is similar to *comp*, except for the fact that it is calculated at the level of agency network, rather than DMA. Therefore, shares of a common market are computed for each advertiser connected to the same holding network. The parameters on *comp* and *compnet* are of particular interest to understand the digital

ad market: for both *comp* and *compnet*, we are agnostic about the expected sign and significance. As discussed earlier, advertisers have reasons both to like and dislike DMA's sharing: industry specialization might push toward agency sharing, but "sleeping with the enemy" motives pull advertisers toward different agencies. Similarly, sharing an agency network means sharing data and bidding algorithms with rivals through the activities of the network's agency trading desk. Nevertheless, also at the level of agency network the "sleeping with the enemy" motives might be strong enough to make advertisers preferring networks less populated by rival advertisers.

The third variable, *diversif*, measures the degree of diversification of the DMAs' property. Diversification is measured at the level of agency network through the following measure of concentration:

$$diversif = \left(\frac{\#DMAs \in network_1}{\#DMAs}\right)^2 + ... + \left(\frac{\#DMAs \in network_n}{\#DMAs}\right)^2,$$

where n is the number of agency networks dealing with the advertiser and # DMAs refers to the number of DMAs matched. Assume for example that an advertiser i deals with four DMAs: 1 belonging to WWP, 2 to Publicis and 1 which is independent. Then, $diversif_i = \left(\frac{1}{4}\right)^2 + \left(\frac{2}{4}\right)^2 + \left(\frac{1}{4}\right)^2$. The variable is used to investigate whether advertisers prefer to "put all their eggs in one basket," at the risk of mismanaging, but possibly avoiding the various costs associated with diversification. The coefficient on this variable will also be interesting to evaluate whether networks may not be able to offer an integrated package of service as required by advertisers, thus triggering diversification.

The fourth variable, *exclusive*, is a dummy that assumes value 1 if both the advertiser and the DMA match with exactly one counterpart (meaning together) and 0 otherwise. Since unmatched agents are not allowed, *exclusive* takes on the value of zero if and only if at least one parties in the match has more than one match. The aim is to understand if agents prefer to have a monogamous business relation when they can afford only one match. In other words, if they value exclusivity when they are forced to provide exclusivity. In the analysis, the number of matches for each player always remains the same, and is therefore treated as a fixed, exogenous, observable attribute. This is equivalent to state that every player decides how many matches he will undertake before starting the matching game and signals the number to the counterpart population.[18]

With regards to the DMAs' profit function, we consider two variables. They are denoted as *affilmarkets* and *affilmatches* and they are intended to capture the fact that larger advertises work with DMAs affiliated to networks, or in other term a "brand effect" of networks. The first variable, *affilmarkets*, measures whether DMAs associated to networks value advertiser who are active on multiple markets more than independent DMAs do. The variable is built by interacting a dummy variable indicating affiliation or independence (1 if affiliated, 0 if independent) with the number of markets participated in by the clients. The rationale is that affiliated intermediaries have more resources to manage complex marketing campaigns, covering many markets. The second variable, *affilmatches*, measures if affiliated intermediaries tend to match with advertiser needing multiple DMAs or not. The variable is similar to the previous one but instead of counting the number of markets, it counts the number of matches for each advertiser linked.

[18] An extension of the same variable not included in the model would be the Interaction between advertiser number of matches and DMA number of matches, trying to understand if a player who prefers multiple matches would rather choose counterparts with many matches as well. Also note that *exclusive*, while included in the advertisers' profit function could, in principle, be included in the DMAs' one as well since it is symmetric for both sides of the market.

Table 2: Profit Functions' Parameter Estimates

		Confidence interval 95%	
Advertiser	coefficient	symmmetric	asymmmetric
exclusive	-1	-	-
netcomp	12.64**	(9.82, 15.46)	(11.81, 15.46)
comp	-8.02**	(-13.25, -2.79)	(-13.25, -5.97)
diversif	-12.10**	(-16.37, -7.83)	(-16.37, -11.07)
Agency	coefficient.	symmmetric	asymmmetric
affilmarket	-0.028	(-0.094, 0.038)	(-0.095, 0.009)
affilmatches	0.338	(-1.40, 2.08)	(-0.34, 2.08)

Notes: ** reports significance at 95%. *exclusive* coefficient is normalized to ± 1.

6.2 Results

The estimation results are presented in Table 2. The top panel reports the coefficients for the parameters in the advertisers' profit function, while the bottom panel reports those for the agencies. The coefficient of *exclusive* is scale-normalized to ±1 and is therefore impossible to calculate the corresponding confidence interval. The main coefficients of interest are those on *comp* and *netcomp*. We find a negative coefficient for *comp* and a positive one for *netcomp*, both significant at 95 confidence level. The negative sign of *comp* supports the common wisdom in the industry about advertisers avoiding to share common DMAs with direct competitors. The only caveat is that competition here is measured as the number of shared advertising market where both firms look for ad slots, while little is said on actual competition over products. But these competition concepts are likely to substantially overlap.

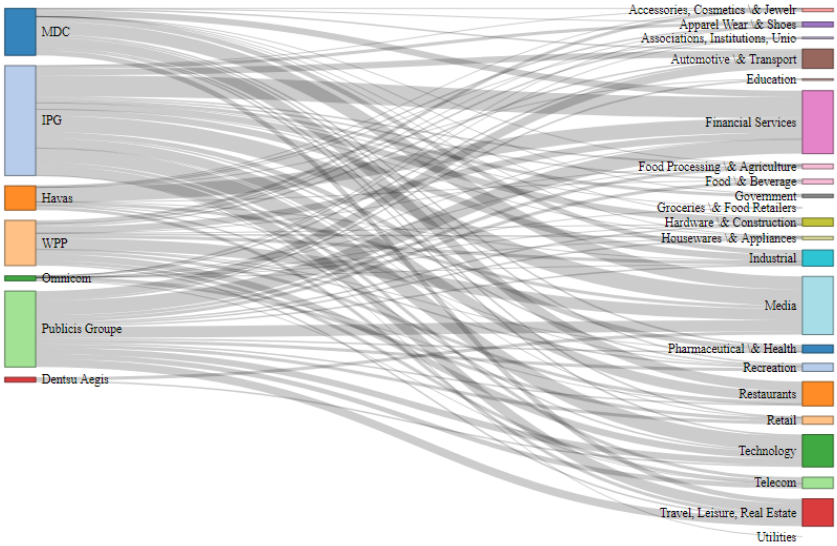
The positive sign of *netcomp* is also interesting. First of all, contrary to the case of DMAs, advertisers are not worried to share the same corporate holding with rivals. But this would not be enough to rationalize the positive and large coefficient which is instead the likely result of the perceived benefits of sharing data and bidding algorithms via the agency trading desks. As discussed earlier, there could be multiple benefits, from lower ad prices through coordinated bidding, to more effective ad targeting and internalization of externalities.

Our analysis alone cannot distinguish between these motives. Nevertheless, the descriptive evidence in Figure 4 shows that the advertisers' preference for a common network is not merely due to the networks' specialization in particular markets. The figure represents, for each of the 7 agency networks (on the left), the industry to which their clients belong (on the right): none of the 7 networks appears to be particularly specialized in any specific industry. This visual evidence is also supported by the results in Table 3, where we report the Krugman index, the inequality in productive structure index and number of industries covered to measure the relative specialization of Networks in different industries.[19]. Both indexes point to a very low level of specialization apart from one network whose activities are circumscribed to three industries due to its limited presence in the digital segment. Therefore network specialization is not taking place on the whole digital

[19]Krugman index is calculated as $\sum_{i=1}^{I}(|(b_i - \bar{b}_i)|)$ and takes value between 0 (indicating that the two distributions are the same) and $2(I-1)/I$. The higher the index, the more dissimilar are the two distributions, with a value of 2 representing a situation where the two distributions have nothing in common. The Inequality in Production index is calculated as $\sum_{i=1}^{I}(b_i - \bar{b}_i)^2$ and takes value between 0 and $(I-1)/I$, where I is the number of industries, b_i is the share of network client belonging to industry i and $\bar{b}$ the share of advertiser belonging to industry i.

advertising market but only within industries. At the same time, since the effect of *comp* is negative, we can exclude that specialization is due to single affiliated DMAs. Even excluding some sort of direct coordination for the purchase, the pattern observed is a clear sign that complementarities between competitors arise at a narrow level inside a given industry. Taking into account the high automation in this sector, it is more likely that such advantage stems from a data and bidding integration rather than to some general acknowledged expertise regarding particular markets.

Figure 4: Network clients divided by industry in 2015



Notes: the figure represents the (lack of) specialization of the 7 agency networks (on the left) in the advertisers' industries (on the right). Each line corresponds to one advertiser.

Table 3: Network Specialization at Industry Level

Networks	Industries Covered	Krugman Index	IP Index
Omnicom	11	0.878	0.067
Dentsu Aegis	3	1.232	0.221
WPP	16	0.333	0.007
Havas	12	0.692	0.047
MDC	17	0.511	0.031
Publicis Groupe	19	0.458	0.016
IPG	19	0.852	0.005

Notes: Krugman index is calculated as $\sum_{i=1}^{I}(|(b_i - \bar{b}_i)|)$ and takes value between 0 and $2(I-1)/I$ while Inequality in Production index is calculated as $\sum_{i=1}^{I}(b_i - \bar{b}_i)^2$ and takes value between 0 and $(I-1)/I$. Where I is the number of industries, b_i is the share of network clients belonging to industry i and $\bar{b}$ the share of advertiser belonging to industry i. Both indexes show that networks are not specialized in particular industries.

Regarding the other coefficients of the advertisers' profit function, for *exclusive* we can only discuss its sign. It is estimated to be negative and this squares well with the presence of many-to-many matches in the data. The *diversif* parameter is negative and significant. In this case the sign suggests that advertisers dealing with multiple DMAs prefer to hire agencies belonging to different

networks, and that this is accepted by DMAs. It is worth noticing that the diversification measure treats affiliated and independent DMAs in the same way, therefore we do not know if networks tolerate clients dealing with other networks or just other independent DMAs. At the same time, another interesting interpretation is possible: networks are not interested in providing an integrated offer to their clients, otherwise we would see advertisers dealing with multiple agencies within the same network. If an integrated service is not available, inevitably the advertisers are forced to turn to multiple networks for their needs. This interpretation might rationalize the frequent acquisitions of independent agencies by the networks as an attempt to expand their portfolio of services offered.

In terms of the agencies' profit function, neither *affilmarket* and *affilmatches* are significant and they present opposite signs. The two measures of matches and markets are intended to be proxies of the firm's marketing campaign dimension, which seems to be unrelated to the choice of the affiliated DMAs. We cannot conclude that the advertisers looking for more DMAs prefer at the same time the ones belonging to networks and that the firms active on more markets are looking for independent ones. Hence, a network's "brand effect" is not detected. Overall, a simple interpretation of our findings about agencies' preferences is that they would be willing to match with any advertiser. If this is certainly not true when considering the universe of US advertisers, this interpretation is not unreasonable in our sample where advertisers are all large US advertisers interested in digital ads.

The overall percentage of inequalities satisfied indicates the fitness of the model: the more inequalities are satisfied, the higher the explanatory power. At our parameter estimates, about 90% of inequalities are satisfied. This percentage is strikingly high if compared to the literature, especially given how parsimonious the profit function specification is. This is a very important result for the analysis that follows where we are going to evaluate counterfactual scenarios. A poor model fit would have meant that the stochastic component of the utility explained the observed matching patters. This would have been problematic since that component is not estimated and would not be usable in the counterfactual analysis.

7 Evaluating Mergers between Intermediaries

In the lively policy debate about how the tech giants are shaping the current economic and social landscape, competition policy has received nearly unprecedented attention. The two pillars of competition policy, merger review and antitrust, have both been invoked to respectively prevent further concentration in the tech sector and monitor, or sanction, anti-competitive behaviors. Since digital advertising is the source of financing for most of the tech sector, it is obviously interesting to point out an application of our analysis to the problem of merger evaluation. This type of perspective analysis is conducted for both horizontal and vertical mergers when the firms involved are large enough to meet certain requirements that imply the compulsory notification of the intended merger to the public authority in charge (i.e., either the DOJ or the FTC in the US). Horizontal mergers involving products or services that are directly sold to consumers are assessed through methods that are by now well established and that, when the data allows, involve full blown merger simulations through the estimation of structural models of demand and supply.

Our estimation approach provides a natural counterpart for the problem of assessing the effects on clients of an upstream horizontal merger between agencies. To the best of our knowledge, the

idea of using a many-to-many matching estimator to evaluate the effects of mergers is novel in the literature and has never been done so far by competition authorities. The closest approach is likely that presented in the recent literature on the estimation of multilateral bargaining games and especially the works of Crawford and Yurukoglu (2012), Gowrisankaran, Nevo and Town (2015) and Ho and Lee (2019). While the methods in these studies could be applied to the evaluation of mergers between agencies, that would require observing the transfers. This is often difficult, even for competition authorities. Hence, we see our method as a viable alternative that allows overcoming this important shortcoming in the data. Therefore, while a full blown merger simulation passing the high standards of a competition case is beyond the scope of this study, the remaining part of this section lays down the key elements entailed by such an approach.

7.1 Empirical Implementation

The advertising agency sector is characterized by increasing concentration. For instance, during 2017 Dentsu Aegis (one of the 7 networks) acquired Merkle agency for around 1.5 billion dollars, Gyro agency for an undisclosed price and several other minor firms. In the same year, WPP, the largest of the networks, acquired DeepLocal Inc., Zubi advertising and others. Publicis group, instead, lost an agency that it had previously acquired, Moroch, due to a buyback of shares, which made Moroch an independent agency again.

We will use our model to evaluate all the M&A cases that we observe in our data for 2017: six acquisitions and two divestitures. We identified these cases by observing changes in the affiliation of DMAs to networks in the Redbooks data and then verifying that these were indeed M&A operations through Zephyr data, the Bureau van Dijk's dataset on M&A deals and rumours. We use the parameters estimated in the previous section to compute the valuation under two different scenarios. The first scenario is the one observed in 2016, called pre-merger scenario, while the counterfactual scenario is identical to the previous one apart from the fact that the DMA affiliation considered is the one the post-acquisition one, realized in 2017. The analysis is conducted in partial equilibrium: in both the factual and counterfactual scenarios, the matching remains the same, what changes is only the DMAs' affiliation to networks. Therefore, only the variables concerning the affiliation characteristics are different.

We focus on the percentage difference in valuation under the two scenarios, which is:

$$\frac{\Delta\hat{\theta}X_i}{\hat{\theta}X_i^{pre-merger}}, \tag{9}$$

where $\Delta\hat{\theta}X_i$ is the difference between factual and counterfactual over every match. Advertisers and DMAs are affected by M&A activity in two ways, directly and indirectly. Directly affected players are those undergoing the change in property and their clients, while indirectly affected players are DMAs affiliated to the acquiring (or divesting) network and their clients.

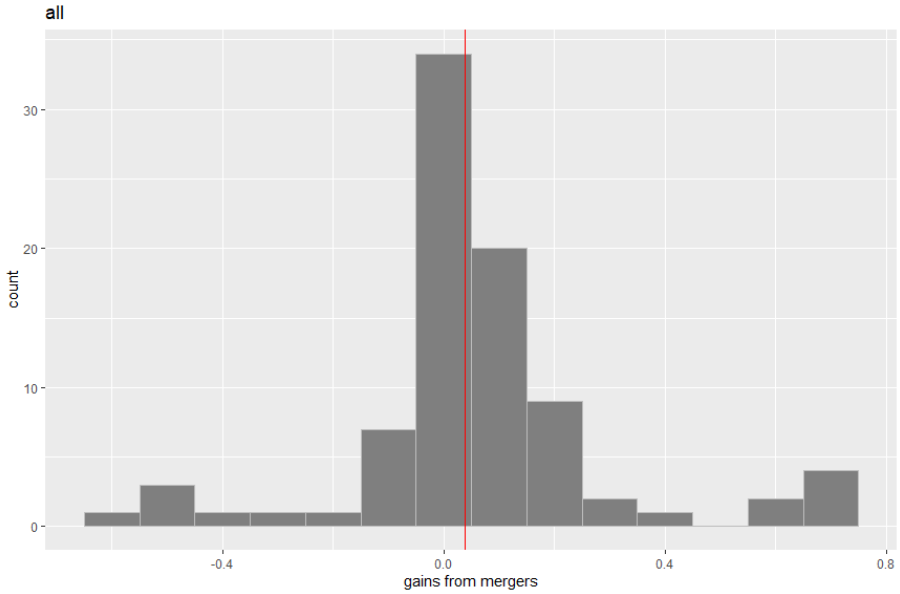
If the reason why valuations of directly-affected players vary is obvious, the indirect effect is less so. There are two channels through which DMAs are indirectly affected. The first channel is the network effect represented by the positive coefficient of the *netcomp* variable: new affiliated DMAs bring fresh data about clients and markets, as well as enhanced possibilities for bid coordination, which help to improve the performance of all partner DMAs, benefiting all clients. This

network effect is a synergy pushing independent DMAs towards networks. The opposite effect, pulling independents and affiliates DMAs apart is the diversification effect represented by the negative coefficient for the variable *diversif*. Negative difference in profits arises when firms already connected to some network face a decreased diversification of intermediaries' affiliation.

Figure 5 summarizes the findings by showing the distribution of the gains (and losses) from the merger. The median gain is +0.04, meaning that more than half of the firms affected by the mergers experience revenue gain greater than 4%. More specifically, 63% of the affected firms show a positive gain. Clients of DMAs affected by M&A activity are relatively few compared to the firms affected indirectly. The ratio of directly vs indirectly affected is 28%. Table 4 reports quintiles for the distribution of gains, distinguishing among directly affected, indirectly affected or both.

-

Figure 5: Distribution of the Gains from the M&As



Notes: the figure displays the distribution of the gains from the merger. The red, vertical line is the median of the distribution.

The gain variance is quite high since it ranges from a loss of 400% to a gain of 90%. The fact that acquisition effects are so mixed should ring a bell on the predictive power of the profit function so specified as it might miss some relevant elements, required to evaluate the effects of DMAs' mergers. It is indeed plausible that the acquisitions having a negative impact on valuation because of the *diversif* valuation are the ones motivated by acquisition of strategic capability. A decrease in valuation is more pronounced when merging DMAs force clients to entertain multiple relations with the same network.

In conclusion, we find that our approach highlights that players win and lose from the intense M&A activity in the DMA sector. However, the results also seem to indicate that these M&As are dictated by considerations not captured by the profit function. This is, however, well aligned with what discussed in industry reports. For example, Gyros (acquired by Dentsu Aegis) is an agency specialized in business-to-business advertising suggesting that Dentsu was looking precisely

Table 4: Percentage Valuation Change by DMAs Affected by Acquisitions

Quantiles	All	Acquired	Indirectly
0	-0.615	-0.615	-0.476
0.2	-0.024	-0.153	-0.024
0.4	+0.016	+0.037	+0.012
median	+0.040	+0.060	+0.039
0.6	+0.057	+0.070	+0.050
0.8	+0.155	+0.460	+0.126
1	+0.728	+0.728	+0.710
% Positive Gain	63%	60%	64%

Notes: This table uses the point estimates from Table 2 to calculate the joint valuations from DMAs that are acquired or divested in 2017 and their clients. The valuations are calculated before and after the acquisitions: in the model ownership change under the future scenario. We calculate $\frac{\Delta \hat{\theta} X_i}{\hat{\theta} X_i^{pre-merger}}$. Acquired column reports change in valuation from DMAs acquired and their clients. Indirectly column reports DMAs not acquired but affected from other DMAs joining or leaving their network.

for such a type of DMA to fill a gap in its service offer. In light of this, it seems that the witnessed sorting concerning diversification is due to external constraint on the inability to provide a full package service, rather than a desire captured by the valuation function. Similarly, it is well known that the aggressive expansion strategy of certain network explicitly targeted DMAs experiencing financial or management problems, for instance linked to the deaths or illness of their mangers or founders. Hence, our approach, while useful to assess some implications of the M&As, is clearly insufficient to also explaining their strategic drivers.

8 Conclusion

This work studies the digital advertising market by estimating the profit functions of advertisers and DMAs as a many-to-many matching game with unobservable transfers. The estimates indicate that direct competitors avoid sharing DMAs, although they prefer sharing the same networks. We argue that this is rationalized by the simultaneous presence of multiple trade-offs that lead advertisers to avoid sharing the same set of creatives with rivals, but to value the benefits allowed by pooling data and bidding algorithms with rivals under the same agency network. The algorithmic nature of the auctions where internet ad space is sold seems to be the driving force of this latter effect, rather than some specialization effect which is not witnessed in the industry. In addition, the estimates reveal the preference of advertisers to diversify among networks or the inability of networks to provide integrated offers. The proposed approach has also been used to analyze mergers between agencies. When applied the M&A operations occurring in 2017, we find that two-thirds of DMAs acquisitions are justified by the synergies stemming from market expertise and data internalized by the whole network.

One of the main limitations of the work is that the valuation function overlooks important aspects, making it more similar to an "externalities" function rather than a faithful model of profits. The flexible assumption on error terms should moderate the concerns regarding the parsimonious specification as all unspecified elements will be captured by it. Nevertheless, using a more nuanced function could yield sharper results. Moreover, only joint surplus is determined, while splitting it between advertiser and DMA is impossible. In other words, it is impossible to state whether

competitors use same DMAs because it directly benefits them or because it benefits the DMAs. This is a straightforward consequence of the model and the fact that we do not observe transfers (i.e., the DMAs' fees) between the parts. This limits the usage of the approach, but also reveals its potential for settings like the ones faced by researchers in which business-to-business transaction prices are not observed. Similarly, the merger analysis – at least in its current formulation – is not a counterfactual assessment of the merger effects, but we are optimistic about the possibility of improving on this in future iterations of this study. Finally, another limitation of the current study which could however be solvable is that the creation and destruction of new matches is not allowed. That is, the number of matches for each player is fixed and regarded as an observable attribute. The analysis, however, could be enriched by relaxing this assumption by allowing agents to match to the empty set and appropriately normalizing the associated profits.

References

Abdulkadiroğlu, Atila, Nikhil Agarwal, and Parag A Pathak. 2017. "The welfare effects of coordinated assignment: Evidence from the New York City high school match." *American Economic Review*, 107(12): 3635–89.

Agarwal, Ashish, and Tridas Mukhopadhyay. 2016. "The impact of competing ads on click performance in sponsored search." *Information Systems Research*, 27(3): 538–557.

Agarwal, Nikhil. 2017. "Policy Analysis in Matching Markets." *American Economic Review*, 107(5): 246–50.

Athey, Susan, and Denis Nekipelov. 2010. "A structural model of sponsored search advertising auctions." Vol. 15.

Azevedo, Eduardo M, and John William Hatfield. 2018. "Existence of equilibrium in large matching markets with complementarities." *Available at SSRN 3268884*.

Baardman, Lennart, Elaheh Fata, Abhishek Pani, and Georgia Perakis. 2019. "Learning Optimal Online Advertising Portfolios with Periodic Budgets." *Available at SSRN 3346642*.

Baccara, Mariagiovanna, Ayşe İmrohoroğlu, Alistair J Wilson, and Leeat Yariv. 2012. "A field study on matching with network externalities." *American Economic Review*, 102(5): 1773–1804.

Balseiro, Santiago R, and Ozan Candogan. 2017. "Optimal contracts for intermediaries in online advertising." *Operations Research*, 65(4): 878–896.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile prices in market equilibrium." *Econometrica: Journal of the Econometric Society*, 841–890.

Berto Villas-Boas, Sofia. 2007. "Vertical relationships between manufacturers and retailers: Inference with limited data." *The Review of Economic Studies*, 74(2): 625–652.

Chen, Jiawei, and Kejun Song. 2013. "Two-sided matching in the loan market." *International Journal of Industrial Organization*, 31(2): 145–152.

Chiappori, Pierre-André, and Bernard Salanié. 2016. "The econometrics of matching models." *Journal of Economic Literature*, 54(3): 832–61.

Choi, Hana, Carl F Mela, Santiago Balseiro, and Adam Leary. 2017. "Online display advertising markets: A literature review and future directions." *Columbia Business School Research Paper*, , (18-1).

Choo, Eugene, and Aloysius Siow. 2006. "Who marries whom and why." *Journal of political Economy*, 114(1): 175–201.

Condorelli, Daniele, Andrea Galeotti, and Ludovic Renou. 2016. "Bilateral trading in networks." *The Review of Economic Studies*, 84(1): 82–105.

Crawford, Gregory S., and Ali Yurukoglu. 2012. "The Welfare Effects of Bundling in Multichannel Television Markets." *American Economic Review*, 102(2): 643–85.

Day, D. 2014. "Matching with Conflicts: An Application to the Advertising Industry." mimeo.

Decarolis, Francesco, and Gabriele Rovigatti. 2019. "From Mad Men to Maths Men: Concentration and Buyer Power in Online Advertising." *CEPR Discussion Paper No. DP13897*.

Decarolis, Francesco, Maris Goldmanis, and Antonio Penta. 2020. "Marketing agencies and collusive bidding in online ad auctions." *Management Science*, forthcoming.

Donna, Javier, Pedro Pereira, Tiago Pires, and Andre Trindade. 2018. "Measuring the Welfare of Intermediaries in Vertical Markets." *Available at SSRN 3285205*.

Ellickson, Paul B, Stephanie Houghton, and Christopher Timmins. 2013. "Estimating network economies in retail chains: a revealed preference approach." *The RAND Journal of Economics*, 44(2): 169–193.

Fox, Jeremy T. 2010. "Identification in matching games." *Quantitative Economics*, 1: 203–254.

Fox, Jeremy T. 2018. "Estimating matching games with transfers." *Quantitative Economics*, 9(1): 1–38.

Fox, Jeremy T, and Patrick Bajari. 2013. "Measuring the efficiency of an FCC spectrum auction." *American Economic Journal: Microeconomics*, 5(1): 100–146.

Goldman, Matt, and Justin M Rao. 2014. "Position Auctions in Practice." *Available at SSRN 2524688*.

Gomes, Renato, and Sweeney Kane. 2014. "Bayes-Nash Equilibria of the Generalized Second Price Auction." *Games and Economic Behavior*, 86: 421–437.

Gowrisankaran, Gautam, Aviv Nevo, and Robert Town. 2015. "Mergers When Prices Are Negotiated: Evidence from the Hospital Industry." *American Economic Review*, 105(1): 172–203.

Hausmann, Ricardo, and César A Hidalgo. 2011. "The network structure of economic output." *Journal of Economic Growth*, 16(4): 309–342.

Ho, Kate, and Robin S Lee. 2019. "Equilibrium provider networks: Bargaining and exclusion in health care markets." *American Economic Review*, 109(2): 473–522.

Horowitz, Joel L. 1992. "A smoothed maximum score estimator for the binary response model." *Econometrica: journal of the Econometric Society*, 505–531.

Jackson, Matthew O, and Asher Wolinsky. 1996. "A strategic model of social and economic networks." *Journal of economic theory*, 71(1): 44–74.

Knittel, Christopher R, and Konstantinos Metaxoglou. 2014. "Estimation of random-coefficient demand models: two empiricists' perspective." *Review of Economics and Statistics*, 96(1): 34–59.

Manski, Charles F. 1975. "Maximum score estimation of the stochastic utility model of choice." *Journal of econometrics*, 3(3): 205–228.

McAfee, R. Preston. 2011*a*. "The Design of Advertising Exchanges." *Review of Industrial Organization*, 39(3): 169–185.

McAfee, R Preston. 2011*b*. "The design of advertising exchanges." *Review of Industrial Organization*, 39(3): 169–185.

McFadden, Daniel, et al. 1973. "Conditional logit analysis of qualitative choice behavior."

Nosal, Kathleen. 2016. "Two-Sided Matching in Physician-Insurer Networks: Evidence from Medicare Advantage." *Available at SSRN 2727272*.

Pakes, Ariel, Jack Porter, Kate Ho, and Joy Ishii. 2015. "Moment inequalities and their application." *Econometrica*, 83(1): 315–334.

Politis, Dimitris N, and Joseph P Romano. 1994. "Large sample confidence regions based on subsamples under minimal assumptions." *The Annals of Statistics*, 2031–2050.

Santiago, David, and Jeremy T. Fox. 2009. "A Toolkit for Matching Maximum Score Estimation and Point and Set Identified Subsampling Inference." *miemo*.

Silk, Alvin J, and Charles King III. 2013. "How Concentrated is the US Advertising and Marketing Services Industry? Myth versus Reality." *Journal of Current Issues & Research in Advertising*, 34(1): 166–193.

Spulber, Daniel F. 1996. "Market microstructure and intermediation." *Journal of Economic perspectives*, 10(3): 135–152.

Storn, Rainer, and Kenneth Price. 1997. "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces." *Journal of global optimization*, 11(4): 341–359.

Train, Kenneth, and Melvyn Weeks. 2005. "Discrete choice models in preference space and willingness-to-pay space." In *Applications of simulation methods in environmental and resource economics*. 1–16. Springer.

Uzzi, Brian. 1996. "The sources and consequences of embeddedness for the economic performance of organizations: The network effect." *American sociological review*, 674–698.

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

Villas-Boas, J. Miguel. 1994. "Sleeping with the Enemy: Should Competitors Share the Same Advertising Agency?" *Marketing Science*, 13(2): 190–202.