Tracking Fraudulent and Low-Quality Display Impressions

Patricia Callejo
IMDEA Networks Institute,
Avda. del Mar Mediterraneo 22 28918
Leganes (Madrid) Spain

Dept. Telematic Engineering
Universidad Carlos III Madrid
Avda. de la Universidad 30
28911 Leganes, Madrid, Spain
patricia.callejo@imdea.org

Ángel Cuevas
Dept. Telematic Engineering
Universidad Carlos III Madrid
Avda. de la Universidad 30
28911 Leganes, Madrid, Spain
Tel: +(34) 916246235

UC3M-Santander Big Data Institute
Calle Madrid, 135 Edif. 18,
28903, Getafe, Madrid, Spain
acrumin@it.uc3m.es

Rubén Cuevas
Dept. Telematic Engineering
Universidad Carlos III Madrid
Avda. de la Universidad 30
28911 Leganes, Madrid, Spain
Tel: +(34) 916245703

UC3M-Santander Big Data Institute
Calle Madrid, 135 Edif. 18,
28903, Getafe, Madrid, Spain
rcuevas@it.uc3m.es

Mercedes Esteban-Bravo
Dept. of Business Administration
Universidad Carlos III Madrid
C/Madrid 126
28903 Gefate, Madrid, Spain
Tel: +(34) 916248921
mesteban@emp.uc3m.es

Jose M. Vidal-Sanz
Dept. of Business Administration
Universidad Carlos III Madrid
C/Madrid 126
28903 Gefate, Madrid, Spain
Tel: +(34) 916248642
jvidal@emp.uc3m.es

Abstract: Display advertising is traded in a complex market with multiple sides and intermediaries, where advertisers are exposed to several forms of potentially fraudulent behavior. Intermediaries often claim to implement measures to detect fraud, but they provide limited information about it. Advertisers are required to trust that self-regulation efforts effectively filter out low-quality ad impressions. In this article, we propose an approach for tracking key display impression metrics, *embedding a light JavaScript code in the ad to collect the necessary information* to help detect fraudulent activities. We explain these metrics using the campaign *cost-per-mille* (CPM) and the number of impressions per publisher. We test the approach through six display ad campaigns. Our results provide a counterargument against the industry claim that it is effectively filtering out display fraud, and show the utility of our approach for advertisers.

Keywords: online display advertising, impressions fraud, auditing metrics.

US digital advertising spend reached $108.64 billion in 2018 (eMarketer 2019a), a large portion of it ($49.23 billion) bought via programmatic advertising (eMarketer 2019b) despite severe concerns about brand safety, fake news, and lack of transparency. The Interactive Advertising Bureau (IAB) estimated the total online ad fraud cost to be $7.2 billion in 2016 (IAB 2016). The Association of National Advertisers (ANA) together with cyber-security company White Ops reported a slightly smaller fraud cost of $6.5 billion in 2017 (ANA/White Ops 2019). Display ad-exchange firms identify invalid traffic using non-disclosed codes, and they do not charge for those clicks and impressions deemed invalid. But these firms have conflicting incentives regarding fraud detection (Edelman 2014a, 2014b; Edelman and Brandi 2015).

There is limited research on display advertising fraud (Edelman 2014a; Fulgoni 2016). The most widely studied type of fraudulent behavior is *click fraud*. Click fraud covers a collection of techniques for artificially inflating the number of clicks on pay-per-click Internet advertisements (Jansen 2007). It can occur for a variety of reasons. Some content publishers or their associates use it to increase their revenues, while other advertisers use it as a tactic to expel temporarily rival advertisers from the ad network, by depleting their budget and thus reducing the competition for target keywords.

Click fraud is just part of the story, though; unethical players also exploit advertisers using other practices. Some vendors monetize impressions in terms of volume due to fake traffic (inflating the count of times an ad is shown by including impressions on artificial users—automated traffic bots from datacenters and botnets formed by malware-controlled computers which mimic human browsing behavior) or fake frequency (displaying the same ad multiple times in milliseconds), distorting advertisers' achievements in terms of reach and frequency. A considerable portion of ad-tech investment likely does not actually reach the targeted audience. In addition, there are fraudulent tactics that distort contextual targeting to display impressions in sites with content that is very different from the advertisers' target (and can be even harmful for the brand).

This article presents an innovative strategy for tracking ineffective display advertising associated with different types of fraudulent activity, embedding a JavaScript code in the advertisement so that information from the impressions is downloaded directly. This information is complemented with the data provided by programmatic advertising intermediaries. We identify fraud level and factors that may exacerbate display fraud in terms of audience volume, contextual targeting, and visibility of a campaign that otherwise would

remain hidden to the advertiser. Our analysis shows the impact of fraudulent display tactics and thus justifies the need to use the systematic tool presented here. An online appendix describes the display advertising industry in detail.

MEASUREMENT OF FRAUDULENT STANCES IN DISPLAY IMPRESSIONS

Our data are grounded in three advertising measurements which we have found to be important factors of online advertising effectiveness. The key measurements that we consider are summarized in Table 1, and are: audience volume, contextual targeting, and visibility. We discuss next how these metrics are relevant for programmatic advertising intermediaries as they can manage their fraud filtering effort in connection with these measurements.

[Table 1: INSERT ABOUT HERE]

Audience Volume

Identifying the audience size in display advertising is a challenging problem; typical metrics are audience reach and frequency. In a display context, both *reach* and *frequency* counts are affected by fraud. Fraudulent traffic can increase reach, or it can increase frequency by artificially refreshing impressions in the user browser in a short time. A large portion of impression fraud can be identified by tracking the user agent and IP address arriving at the advertiser site and receiving an ad impression, then matching the IP with an *illegitimate datacenters list*. A datacenter is a physical or virtual infrastructure where a large group of computers is centralized to store, process, or distribute a large volume of data remotely. Associations such as the Media Rating Council (US) and JICWEBS (UK) include datacenter traffic as a common source of invalid traffic and recommend filtering such traffic. Hence, the IT community elaborates lists of centers with bad-behaved bots that do not identify themselves as such in their declared user-agent strings;[1] these lists focus on traffic programmed to masquerade as humans, and exclude well-behaved datacenters such as those channeling legitimate traffic originated by virtual private networks' (VPNs') secure traffic. Integral Ad Science found that 8.3% of all US digital display impressions were fraudulent (see Q1 2016 survey at https://integralads.com/). We computed two metrics for datacenter

[1] For example, the International IAB/ABC Spiders & Bots List, and the Trustworthy Accountability Group (TAG) list made available by Google.

impression fraud, both dummy variables:

- *DataCenter*, which takes a value of "1" if impressions are served to IP addresses belonging to datacenters, and "0" otherwise, using Botlab and FireHOL IP Lists (Botlab 2016; FireHOL 2017) to identify datacenter IPs; and
- *Approved*, which takes a value of "1" if impressions are approved by the ad intermediary as valid.
- *Non-excluded*, which takes a value of "1" if impressions are approved by the ad intermediary as valid but are tracked as fraudulent (to datacenter) by our auditing code. Non-excluded impressions are those for which the advertiser pays.

A user can be exposed to a high frequency of impressions of the same ad display in a short period of time (sometimes hundreds of them), and the ad-exchange firm may report these as different valid impressions when, in fact, they are not unique from each other. Some DSPs allow advertisers to prevent this problem by including a *frequency cap* (a limit to the number of these impressions). Frequency capping is a way to prevent over-exposure to an ad, but often it is used as a protective tool to deter datacenter-based fraud (datacenter bots generate a massive amount of traffic in a very short period of time). For example, Google Ads provides a frequency cap option, but it is not activated by default and it is not trivial to change it for non-skilled users. Our JavaScript code registers the *impression timestamp* (the time when a user arrives at an advertiser's site and receives an impression), and we again computed two metrics:

- *Inter-impression Time*, the time between two consecutive impressions reported as valid by the ad intermediary (with *Approved*=1), in seconds; and
- *User-Impressions Intensity*, the quantity of valid impressions received by a user within a minute. (A user is defined as the combination of the IP address and the user-agent; therefore, two users sharing an IP address and using the same browser would be considered as a single user in our results).

Previous research on online behavior shows the relevance of display timing. Moe and Fader (2004) and Danaher et al. (2006) study the impact of Web site visit duration and the inter-visit times on conversion behavior. Deane and Agarwal (2012) study optimal scheduling of time slots in a display campaign over a period of time.

Another concept related to audience size is the popularity of the publisher site. Alexa (https://www.alexa.com) is a company owned by Internet retailer Amazon; it ranks websites by traffic. We use Alexa's global ranking to measure the popularity of the publishers' sites

based on the number of website visits. The ranking is based on web traffic (global, by country, or by category) and is a proxy for the gross rating point (GRP), i.e. the impact, of a publisher's website (computed as the mean number of impressions in the website multiplied by the publisher's mean display time). We recorded one metric for measurement of publisher popularity:

- *Website popularity.*

We use the Global Alexa Ranking as an operationalization of this measurement.

Contextual targeting

Display advertising can target advertising in three ways: using the demographic information that users provide online, using contextual information based on matching the ad content with the website the user is seeing, or using past online behavior based on cookies. Cheap contextual targeting is one of the key advantages of online advertising compared with other traditional media (Goldfarb 2014). In traditional media, congruency between advertising and context increases ad effectiveness (see the review by De Pelsmacker, Geuens, and Anckaert 2002) and the choice of media can have a contextual effect (Dahlén 2005). In the digital context, there are some differences. Goldfarb and Tucker (2011) report experiments suggesting that, for unobtrusive displays, increasing the contextual match increases the purchase intention. On the other hand, for campaigns that are not contextually targeted (i.e. that have no match between display ad and publisher), increasing obtrusiveness results in higher purchase intentions (a rationale for this being that poor contextual matching can make the ad more noticeable, increasing attention). However, combining contextual targeting and obtrusiveness is not very effective. Thus, if an advertiser's strategy sets contextual matching, it is very important that the ad impressions satisfy the contextual matching requirement; otherwise, the ad effectiveness might be considerably diminished (especially if the ad is obtrusive). To identify possible ineffective advertising due to contextual mismatch, we considered three metrics:

- *Strict Keyword Matching*, which uses a dummy variable, *MatchingKeywords*, which takes a value of "1" if at least one of the keywords assigned to the campaign matches a URL keyword in the ad intermediary, and a value of "0" if no campaign keyword matches any of the URL's keywords. This metric evaluates the misplacement of ad impressions.[2] We

[2] See https://iabuk.net/blog/brand-campaigns-benefit-from-contextually-relevant-placement.

focus on keyword mismatching as the result of intermediary actions (not as an advertiser choice). Campaigns configured based on *keywords* follow a *contextual* strategy, where intermediaries prioritize display ads in publishers whose content is related to the targeted keyword(s) and thus contextually meaningful for the campaign. Contextual impression ads often boost the effect of any advertising;

- *L-Ch Similarity*, following Leacock and Chodorow (1998) who proposed a semantic-similarity measure between two lexical concepts in a given ontology; the more similar the two concepts are, the more closely related they are (the path between these concepts is shorter). Formally, it is defined as

$$L\text{-}Ch\ Similarity = -\log\ (length / (2 * D))$$

where *length* is the length of the shortest path between the two concepts (using node-counting) and *D* is the maximum depth of the ontology. It is commonly used because it is easily scalable for large textual analysis (see e.g. Lin and Sandkuhl 2008). We use this measurement to study the similarity between the publisher's topics and the keywords of the campaign; and
- *Brand Safety*, which categorizes websites where the impression is displayed using the web content as potentially negative for the advertiser.

Display duration

The exposure duration of stimuli has been found to be a relevant factor in allocating attention. Research by Bannerconnect found that ad impressions with a short exposure time achieved lower levels of engagement (click-through rates (CTR) decrease).[3] Impression exposures are affected by fraudulent or low-quality impressions in CPM campaigns, and we considered two metrics:

- *Display Duration*, which uses a continuous variable, *Displaytime* (impression duration), to measure how long an ad is active in a webpage (in seconds). On average, a display lasts for 71 seconds (44 seconds in the General campaign, and 101, 77, and 56 seconds in the Spain, Russia, and USA campaigns, respectively); and
- *Visual Perceptibility*, where our dummy variable, *Viewability* of impressions, takes a value of "1" when the impression display time is greater or equal to 1 second, and is "0" otherwise.

[3] See https://www.bannerconnect.net/exposure-time-a-new-standard-for-measuring-digital-effectiveness/.

Zhang et al. (2015) discuss measurements of display ad impression viewability. Note that display viewability does not imply that users actually look at the ads; this type of analysis requires other metrics, such as eye-tracking (Dreze and Hussherr 2003).

The industry recognizes that the CPM and the number of impressions have an impact on fraud. The 2017 study by the Association of National Advertisers (ANA/White Ops 2019) reports that fraud protection is not free, so the lowest CPMs may not include sophisticated protection measures—even the simplest, cheapest bots go unnoticed. The efforts of the advertising industry to tackle the problem justify that negligible cost impressions may show higher levels of hidden fraud. In this context, fraudsters benefit from high numbers of impressions. Based on this evidence, combined with the fact that fraudulent displays are often served to automated traffic bots from datacenters, as discussed previously, our study analyzes the relationships between our metrics and CPM, the number of impressions, and whether the impressions are served to a datacenter. Braun and Moe (2013) examine the impact of ad impressions on visits and conversions.

The CPM, the number of impressions and whether impressions are delivered to datacenters variables will be used as predictors of fraudulent impression indicators (that are, Non-excluded, Inter-impression times, Website popularity, Strict Keyword-Matching, L-Ch-Similarity, Display time, and Viewability). Table 2 describes the dependent variables and models considered in our analysis.

[Table 2: INSERT ABOUT HERE]

AN EXAMPLE FRAUD AUDIT

We ran six different display ad campaigns that aim to promote "research," as defined by keywords ("research," "universities," and/or "telematics"), target location (Spain, Russia, or USA) and CPM (0.01, 0.05, 0.10, or 0.20) in February and March of 2016. We used a leading programmatic advertising intermediary which delivers display ads using Google AdWords (the largest advertising network available on the Internet, with over 2 million publishers and reaching over 90% of all Internet users). Table 3 contains information on each display ad campaign. This resulted in 103,915 ad impressions (observation units), for which we computed the metrics discussed in the previous section.

[Table 3: INSERT ABOUT HERE]

In total, the dataset consists of 3,506 different publishers. Note that in some cases the URL is not registered (reported as URL=*null*). There are referrals from Google AdWords to publishers who want to preserve their anonymity, for which the destination URLs are not tagged (and they are reported as URL=*tpc.googlesyndication.com*). In our database, 13.48% of the impressions are associated with this type of URL, with the remaining 89,905 impressions recognized by the ad intermediary. This means that 51.38% of the publishers have not been reported. For these two URL identifiers (URL=*null* and URL=*tpc.googlesyndication.com*), we considered the URL as missing data in our analysis (so we analyzed 3,504 publishers' websites). Tables 4 and 5 show descriptive statistical data. There is no evidence of multicollinearity in the regression models described in Table 2, as the largest VIF is smaller than 1.3 (the VIF for CPM, Number of Impressions, and Datacenter are 1.12, 1.21, and 1.27, respectively) and the condition number is 6.28.

[Tables 4 and 5: INSERT ABOUT HERE]

Non-excluded

A large percentage of the impressions in our campaign are served to suspicious traffic from datacenters. Overall, in our dataset, 21,432 impressions are delivered to datacenters (20.62% of 103,916 total ad impressions). Of the traffic domain/URLs, 17.41% are identified as datacenters (610 out of the total set of 3,504 unique content website URLs).

Table 6 shows that the probability of datacenter impressions being excluded by the ad intermediary is higher when the CPM and the number of impressions are smaller. This result suggests that the ad intermediary filter is stricter with smaller values of CPM and with fewer impressions.

[Table 6: INSERT ABOUT HERE]

The ad intermediary only identifies 43,700 as valid impressions (*Approved*=1). The 8.78% of these valid impressions (3,836) are served to datacenters (31.20% in Campaign 4 and 20.81% in Campaign 3), representing 335 unique content website URLs (9.56% of the total 3,504

URLs). The total cost paid for these datacenter impressions represents 3.22% of the total investment in the six campaigns.

User-Impressions Intensity

Figure 1 shows the median number of valid impressions received by a user in a campaign during a 15-minute time window, reporting all time windows since the start of the respective campaign. We observe that in many cases a user is exposed to a high number of impressions of the same ad in a short period of time, and that the ad intermediary often reports it as a valid display.

[Figure 1: INSERT ABOUT HERE]

Inter-impression Times

The *Inter-impression Times* quantiles for all campaigns show that 10% of users receive the same ad within 5 seconds or less, 25% of users receive the same ad within a period of less than 11 seconds, and 50% of users receive the same ad with inter-impression times lower than 43 seconds. By campaign, the most dramatic case is Campaign 6, where 10% received the same ad within 4 seconds.

Inter-impression times, and even their logarithm, have an asymmetric distribution. Therefore, we considered a *quantile regression* (see Koenker and Bassett 1978) which we named Model 2 (in Table 2).

Table 6 reports that CPM has a larger positive impact on the lower quantiles of *log(inter-impression times)*. The 25th quantile of *log(inter-impression times)* is more affected by CPM than the 50th quantile. This suggests that high-frequency fraud is more prevalent when the campaign is cheaper. For the number of impressions, the effects are similar and positive on the 25th quantile and median of *log(inter-impression times).* The effect of the number of impressions is negative on higher quantiles (fewer impressions implies higher inter-impression times). In addition, the quantile regression results indicate that the effect of datacenters is much stronger at higher quantiles of *log(inter-impression times)*. This suggests that high-frequency fraud is more prevalent when the campaign is not delivered to datacenters. Note that advertisers can set up a frequency cap in their campaigns indicating the maximum number of times an ad can be shown to a user. The six campaigns we investigated did not set up any frequency cap, so we analyzed the default behavior of the ad intermediary. In this case, the datacenter filtering seems to be working properly.

Website popularity

We consider the Global Alexa Ranking as a proxy for website popularity. For our publishers' websites, the highest Alexa ranking is 1 (for www.google.com) and the lowest is 1,433,041 (for www.universalvideos.us), with the median being 12,281. The higher the global Alexa ranking number, the higher the publisher's popularity.

Table 6 shows that the website popularity increases by 1.26% per one unit increase in CPM, while holding all other variables constant; the website popularity decreases by 168.52% per increase of 1,000 impressions. Datacenter's impressions have no significant effect on website popularity.

Strict Keyword Matching

Out of 3,504 listed publishers' websites, we have data on matching keyword impressions for only 1,088 URLs (for the missing observations, either the ad intermediary excluded all impressions, or it did not report any data on exact keyword matching for the approved impressions). Out of the 1,088 URLs, only 40 (3.68%) have an exact match for some campaign. Focusing on impressions and using the ad intermediary metrics, 1.19% of impressions match a URL keyword out of 43,015 impressions for which the ad intermediary reports exact matching (the ad intermediary reports exact matching for just 41.39% of the total 103,916 impressions). If we consider only the valid impressions, 1.82% have exact matching out of 22,993 valid impressions (actually, the number of valid impressions is 43,700, but the ad intermediary reported exact matching for only 52.61%).

Table 6 (Model 4) shows that the probability of exact matching for approved impressions is higher when the CPM and the number of impressions are smaller. This result suggests that increasing CPM incentivizes the ad intermediary to display the ad in publishing sites less contextually relevant in terms of exact matching. CPM has a similar impact if we take all impressions (recognized by the ad intermediary or not), but the effect is smaller in absolute terms. The number of impressions in the URL negatively affects the probability of exact matching, suggesting that competition in the publisher site reduces the probability of exact matching. In contrast, Table 6 indicates that the probability of exact matching for approved impressions is higher when the impressions are delivered to a datacenter. This result suggests that when considering a campaign on a specific topic, e.g. "sport," the ads that appear on "sports" pages are more likely to be delivered to users who come from datacenters. The

campaigns that we have configured are based on context and therefore the user who visits the page should be irrelevant when choosing the page in which to show the ad. The results seem to suggest that the decision is made not purely on the basis of the context but on the user who visits it, implying low quality of impressions.

Dropping the impressions in publishers with high frequency (more than 500 impressions, as potentially fraudulent), the effect of the number of impressions on the probability of exact matching is positive (the coefficient estimate is 0.0027926, with a p-value of 0.000). This suggests that for low-frequency publishers, the ad intermediary is slightly more likely to do an exact matching when the number of impressions increases; while for high-frequency publishers it is the opposite.

Leacock-Chodorow Similarity

Table 6 reports the parameter estimates for Model 5 (in Table 2). These results suggest that the CPM plays a relevant role in display contextual relevance, and that the best result for the publisher is obtained for intermediate CPM levels. Note also that the effect of *DataCenter* is positive and significant on the *L-Ch Similarity* measure.

Dropping impressions in publishers' websites with more than 500 campaign impressions, the effect of CPM, number of impressions and *DataCenter* is higher on the *L-Ch Similarity* measure. As expected, for low-frequency publishers, the ad intermediary is slightly more likely to do an exact matching.

Brand Safety

In the online Appendix, we discuss Brand Safety issues related to this study. We review the contextual match of the websites with more than 500 impressions in some of the campaigns.

Display Time

Table 6 reports the parameter estimates for Model 6 (in Table 2). The results suggest that the expected *displaytime* increases by 8.33% per one unit increase in CPM, while holding all other variables constant. The number of impressions and whether they come from a datacenter or not have no significant effect on *displaytime*.

Viewability

Next, we focus on viewability. Note that 3.32% of the impressions last less than 1 second (for Campaigns 5 and 6, 5.21% and 4.38%, respectively, last less than 1 second). Table 6 reports the estimates of Model 7 (in Table 2). For each one unit increase of CPM, the estimated odds of impressions that are displayed for at least 1 second increase by 3.5392%, while holding all other variables constant. Similarly, the odds of viewability increase by 24.87% per increase of 1,000 impressions, while holding all other variables constant. As expected, the effect of *DataCenter*'s impressions on viewability is large.

CONCLUSIONS AND MANAGERIAL RECOMMENDATIONS

This article discusses several types of metric to detect fraud in ad displays. Our empirical study provides evidence of a considerable potential fraud based on impressions served to suspicious traffic from datacenters (in one campaign it reached 44.44% of all impressions). The overall level of impressions fraud might be even larger, as we do not identify impressions served to botnet computers controlled by malware. The ad intermediary charged us 3.22% of our total budget for impressions to datacenter traffic. In addition, there is a considerable level of potentially fraudulent impressions due to high frequency (50% of the total inter-impression times by users reported by the ad intermediary are lower than 43 seconds).

Our analysis also suggests that ad intermediaries fail to tackle impressions fraud and that their efforts depend on CPM. Our data suggest that the probability of (several types of) hidden fraud is related to CPM and number of impressions. We found that campaigns with the highest CPM have less risk of hidden impression fraud, leading to a recommendation for advertisers to pay more attention when running cheaper display ads. We also find evidence of contextual biases, where the impressions do not match the targeted keywords, or where there is low L-Ch similarity. This problem also varies with CPM. Moreover, there is a considerable risk to advertisers of having their brand damaged by exposure in potentially harmful contexts; in our campaigns, several potentially harmful sites (spicy humor, dating, and gaming sites) received more than 500 impressions.

Note that to establish absence of fraud, we would need a systematic large-scale study, but to prove that the self-regulation system is fallible, we need only a small counterexample. The fact that we ran just a small test and directly obtained a counterexample against the correct functioning of this industry suggests that the problem might be systemic. This could result in a range of serious concerns when considering massive investments in display advertisements. We have several recommendations for advertisers:

(1) Use intermediaries that enable you to implement a light JavaScript code to directly track different forms of fraudulent activities. The software and code are available upon request (see the online Appendix for details).

(2) Do not bid too cheaply. If the ad is displayed, the level of hidden fraud might be considerably higher if the CPM is low. Middlemen may not use sophisticated fraud detection tools when the fraud is too low.

(3) Use the frequency cap option to avoid paying for a considerable amount of high-frequency impressions with low viewability. Using the default specification for a campaign introduces the serious risk of exposure to fraud. Advertisers using our approach are likely to obtain similar insights for their own campaigns.

(4) Change the default settings exhaustively to prevent impressions in websites posing a risk for brand safety (see the Online Appendix). Some firms are already realizing about this problem; for example, JPMorgan Chase used to display ads over 400,000 websites monthly, but after recently detecting display impressions next to toxic content it has dramatically cut to 5,000 pre-approved websites.[4]

Our study was conducted using a leading company, but future research could explore other vendors, and a broad number of campaigns associated with specific types of keyword. Lack of transparency is a general problem that affects the whole ad-tech industry, and we would not be surprised to find similar problems in other ad-exchanges and intermediaries such as ANs and DSPs. Further, we used relatively simple models, but future research could consider more elaborate specifications (such as hierarchical models with fixed or random effects, models that account for measurement errors, self-selection models to handle missing data, nonparametric and machine learning methods, etc.) These approaches provide useful robustness checks, and future research might explore these avenues with larger samples.

[4] *New York Times*. A version of this article appeared in print on March 30, 2017, on page B1 of the New York edition with the headline: "A Bank Had Ads on 400,000 Sites. Then Just 5,000. Same Results." See also: https://www.nytimes.com/2017/03/29/business/chase-ads-youtube-fake-news-offensive-videos.html?smprod=nytcore-iphone&smid=nytcore-iphone-share (accessed December 8, 2017).

REFERENCES

ANA/White Ops (2019), "Five Charts: The State of Ad Fraud." Available at: www.emarketer.com/content/five-charts-the-state-of-ad-fraud (accessed December 3, 2019).

Botlab (2016), "Botlab.io Deny-Hosting IP List." Available at: https://github.com/botlabio/deny-hosting-IP (accessed October 12, 2016).

Braun, Michael, and Wendy W. Moe (2013), "Online display advertising: Modeling the effects of multiple creatives and individual impression histories," *Marketing Science*, 32 (5), 753-767.

Choi, Hana, Carl F. Mela, Santiago R. Balseiro, and Adam Leary (2017), "Online Display Advertising Markets: A Literature Review and Future Directions," Columbia Business School Research Paper No. 18-1. Available at SSRN: https://ssrn.com/abstract=3070706 (accessed August 8, 2018).

Dahlén, Micael (2005), "The Medium as a Contextual Clue: Effects of Creative Media Choice." *Journal of Advertising*, 34(3), 89–98.

Danaher, Peter J., Guy W. Mullarkey, and Skander Essegaier (2006), "Factors Affecting Web Site Visit Duration: A Cross-Domain Analysis" Journal of Marketing Research, 43(2), 182–194.

Deane, J., and A. Agarwal (2012), "Scheduling online advertisements to maximize revenue under variable display frequency," *Omega*, 40 (5), 562-570.

De Pelsmacker, Patrick, Maggie Geuens, and Pascal Anckaert (2002), "Media Context and Advertising Effectiveness: The Role of Context Appreciation and Context/Ad Similarity." *Journal of Advertising*, 31(2), 49–61.

Dreze, X., and F.-X. Hussherr (2003), "Internet advertising: Is anybody watching?," *Journal of Interactive Marketing*, 17 (4), 8–23

Edelman, Benjamin G. (2014a), "Pitfalls and Fraud in Online Advertising Metrics: Are Cheaters Hurting Your Bottom Line?" *Journal of Advertising Research*, 1–6. Available at: www.benedelman.org/publications/pitfalls-and-fraud-in-online-advertising-research-jar-jun2014.pdf (accessed November 8, 2019).

———(2014b), "Accountable? The Problems and Solutions of Online Ad Optimization," *IEEE Security & Privacy*, 12(6), 102–107.

———, and Wesley Brandi (2015), "Risk, Information, and Incentives in Online Affiliate Marketing," *Journal of Marketing Research*, 52(1), 1–12.

eMarketer (2019a), “Digital Ad Spending2019: US.” Available at: www.emarketer.com/content/us-digital-ad-spending-2019 (accessed November 3, 2019).

———(2019b), “US Programmatic Ad Spending Forecast 2019.” Available at: www.emarketer.com/content/us-programmatic-ad-spending-forecast-2019 (accessed November 3, 2019).

Fulgoni, Gian M. (2016), “Fraud in Digital Advertising: A Multibillion-Dollar Black Hole: How Marketers Can Minimize Losses Caused by Bogus Web Traffic” Journal of Advertising Research. 56(2), 122-125.

Goldfarb, Avi (2014), “What Is Different About Online Advertising?” *Review of Industrial Organization*, 44(2), 115–129.

———, and Catherine Tucker (2011), “Online Display Advertising: Targeting and Obtrusiveness.” *Marketing Science*, 30(3), 389–404.

IAB (2016), “Desktop Display Impression Measurement Guidelines.” Available at: www.iab.com/wp-content/uploads/2017/11/Desktop-Display-Impression-Measurement-Guidelines-US-MMTF-Final-v1.1.pdf (accessed November 8, 2019).

Jansen, Bernard J. (2007), “Click Fraud,” *IEEE Computer*, 40(7), 85–86.

Koenker, Roger and Gilbert Bassett Jr. (1978), “Regression Quantiles,” *Econometrica*, 46(1), 33–50.

Leacock, Claudia and Martin Chodorow (1998), “Combining Local Context and WordNet Similarity for Word Sense Identification.” *WordNet: An Electronic Lexical Database*, 49(2).

Lin, Feiyu and Kurt Sandkuhl (2008), “A Survey of Exploiting WordNet in Ontology Matching.” In: *Artificial Intelligence in Theory and Practice II* (ed. Max Bramer); vol. 276 of *IFIP AI 2008, International Federation for Information Processing* (Boston, MA: Springer), 341–350.

Moe, Wendy W., and Peter S. Fader (2004), “Dynamic Conversion Behavior at E-Commerce Sites” Management Science, 50(3), 326–335.

Zhang, Weinan; Ye Pan, Tianxiong Zhou, Jun Wang (2015), “An empirical study on display ad impression viewability measurements,” *arXiv preprint arXiv:1505.05788*.

Figure 1: Evolution in the median number of impressions per user during 15-minute windows

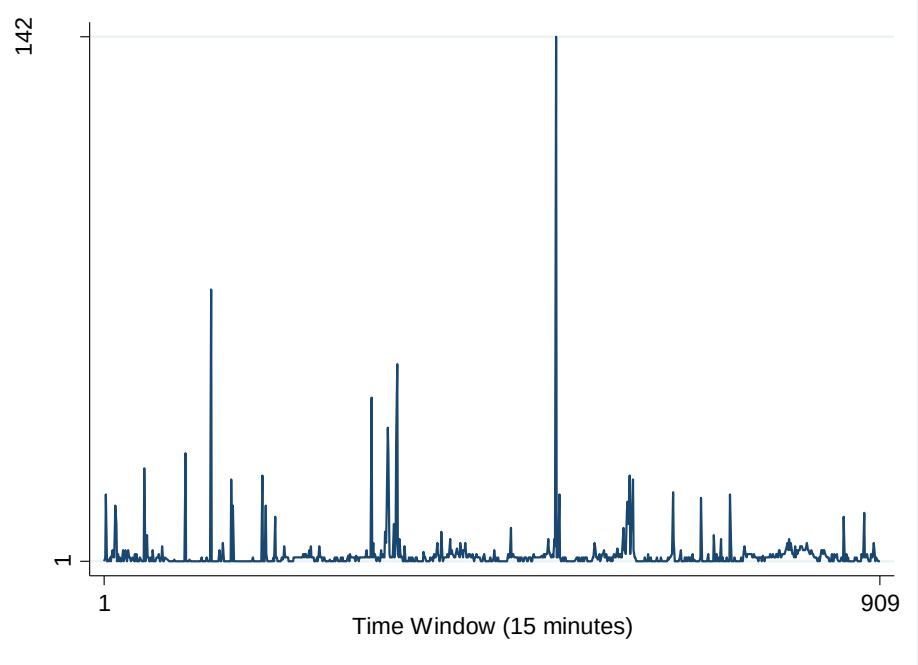


Table 1: Framework of analysis and fundamental metrics

	Concept	Metrics
Audience volume	Datacenter impression fraud	• *DataCenter* (IP address belongs to a datacenter) • *Approved* (IP address approved as valid by the ad intermediary)
	High-frequency fraud	• *User-Impressions Intensity* (impressions received by a user within a minute) • *Inter-impression Times*
	Publisher popularity	• *Website popularity*
Contextual targeting	Strict keyword matching	• *MatchingKeywords* (matching campaign and publisher URL keywords)
	Similarity	• *L-Ch Similarity* (Leacock-Chodorow similarity)
	Negative context	• *Brand Safety*
Impression exposure	Display duration	• *Display Time* (duration in seconds)
	Visual perceptibility	• *Viewability* (display time longer than a second)

Table 2: Main models

	Explained variable	Regressors
Model 1 (Logit)	*Non-excluded*	*Constant, CPM, Number of impressions*
Model 2 (Quantile Regression)	*Inter-impression times*	*Constant, CPM, Number of impressions, Datacenter*
Model 3 (exponential model)	*Website popularity*	*Constant, CPM, Number of impressions, Datacenter*
Model 4 (Logit)	*Strict Keyword-Matching*	*Constant, CPM, Number of impressions, Datacenter*
Model 5 (Linear)	*L-Ch-Similarity*	*Constant, CPM, Number of impressions, Datacenter*
Model 6 (exponential model)	*Display time*	*Constant, CPM Number of impressions, Datacenter*
Model 7 (Logit)	*Viewability*	*Constant, CPM, Number of impressions, Datacenter*

Table 3: Description of the six ad campaigns used to test our auditing methodology

TACTIC CAMPAIGN	TOPIC	Number of impressions (observations)	Number of publishers	Start date	End date	CPM (Euros)	Keywords	Target Location
1	Research in Spain	5117	350	29 March	31 March	0.10	Research	Spain
2	Research in Spain	42398	1776	29 March	31 March	0.20	Research	Spain
3	Research in Russia	4096	274	29 March	31 March	0.01	Research	Russia
4	Research in USA	1178	135	29 March	31 March	0.01	Research	United States
5	Research in General	8767	577	15 February	23 February	0.05	Universities, Research, Telematics	Spain
6	Research in General	42359	1548	18 February	23 February	0.10	Universities, Research, Telematics	Spain

Table 4: Summary statistics of the key variables

Metric	Number of observations	Mean	Std. Dev.	Min	Max
Approved	103,916	0.42	0.49	0	1
Individual impressions-intensity	103,916	81.93	142.89	1	734
Inter-impression times	71,087	9,183.25	124,491.40	0	3,549,681
Website popularity	43,015	0.01	0.11	0	1
Matching campaign and publisher URL keywords	20,537	2.27	0.38	1.34	4
Leacock-Chodorow similarity	61,763	120,824.80	1,078,480.00	0	83,600,000
Display time	103,916	0.97	0.18	0	1
Viewability	103,916	13.20	6.06	1	20
CPM	30,645	52,763.19	103,599.30	1	1,433,041
Number of impressions	103,916	13,950.63	15,325.66	1	34,690
Datacenter	103,916	0.21	0.40	0	1

Table 5: Pearson's pairwise correlation coefficients of the key variables

Variables											
Approved	1										
Individual impressions-intensity	-0.09*	1									
Inter-impression times	-0.03*	-0.01*	1								
Website popularity	-0.17*	-0.06*	0	1							
Matching campaign and publisher URL keywords	0.06*	0.02*	0	0.05*	1						
Leacock-Chodorow similarity	-0.16*	0.03*	0	0.01	0.25*	1					
Display time	0.01*	-0.02*	0	0.01	-0.01	-0.01	1				
Viewability	0.02*	-0.07*	-0.01*	0.01	0.01*	-0.01	0.03*	1			
CPM	0.21*	-0.17*	0.03*	-0.01	-0.11*	-0.18*	0.05*	0.05*	1		
Number of impressions	-0.72*	0.10*	0.03*	-0.32*	-0.09*	-0.10*	-0.01*	-0.02*	-0.23*	1	
Datacenter	-0.25*	0.18*	0.03*	-0.02*	0.21*	0.13*	-0.02*	-0.05*	-0.30*	0.37*	1

* Denotes correlation coefficients significant at the 1% level.

Table 6: Estimates of the main models (Table 2)

	Model (1)	Model (2) (Quantile 0.25 regression)	Model (2) (Quantile 0.5 regression)	Model (2) (Quantile 0.75 regression)	Model (3)	Model (4)	Model (5)[a]	Model (6)	Model (7)
	Coef.	*Coef.*	*Coef.*	*Coef.*	*Coef.*	*Coef.*	*Coef.*	*Coef.*	*Coef.*
CPM	-0.0619	0.0230	0.0091	-0.0042	0.0126	-0.1010	-0.0086	0.0728	0.0304
	(0.0026)	(0.0024)	(0.0024)	(0.0044)	(0.0015)	(0.0110)	(0.005)	(-0.013)	(0.0046)
Nº impressions	-0.0002	0.0001	0.0001	-0.0002	-0.0020	-0.0022	-0.0001	-0.0000	0.0002
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0002)	(0.0000)	(0.0000)	(0.0000)
Datacenter		0.7696	1.5461	3.2351	-0.1990	2.8616	0.1462	-0.9091	0.5733
		(0.0512)	(0.0520)	(0.0947)	(0.0472)	(0.1135)	(0.0100)	(0.2233)	(0.1205)
Intercept	-1.6586	1.9625	3.4277	5.4459	11.2439	-2.9705	2.3594	10.7134	2.8393
	-0.0351	(0.0355)	(0.0361)	(0.0658)	(0.0271)	(0.1463)	(0.0088)	(0.2371)	(0.0667)
Global significance	5,356.77[b]					1,394.85[b]	243.29[c]		192.16[b]
McFadden's Pseudo R^2	0.16	0.01	0.01	0.04	0.33	0.33	0.05	0.02	0.02
Number of observations	103,916	28,132	28,132	28,132	27,386	22,993	14,265	25,382	43,700

Note: Standard errors are in parentheses. [a] We use White's robust to heteroskedasticity standard errors.

[b] Likelihood Ratio Chi-Square test. The p-values are 0.00; the null is rejected.

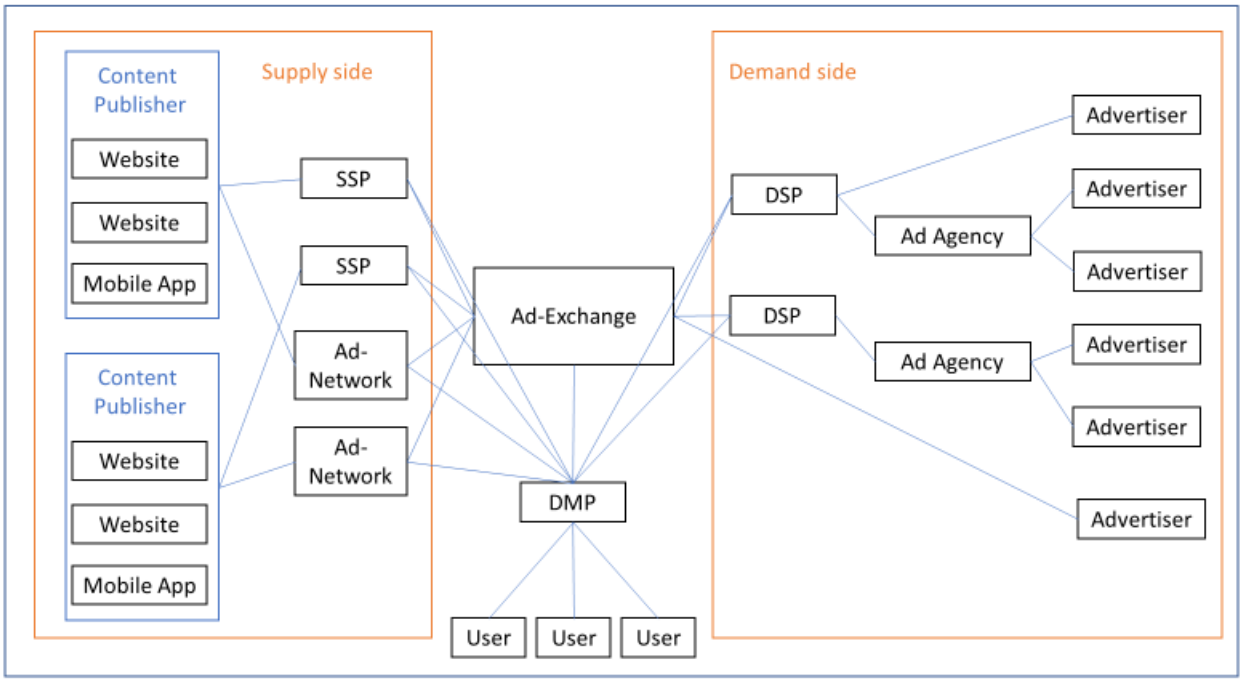
[c] F-statistics test. The p-values are 0.00; the null is rejected.

ONLINE APPENDIX TO: "TRACKING FRAUDULENT AND LOW-QUALITY DISPLAY IMPRESSIONS"

THE DISPLAY ADVERTISING INDUSTRY

The display ads market has a rich structure (Choi et al. 2017). On the demand side, advertising agencies (mediating with advertisers) and large advertisers generally contract with programmatic intermediaries known as *demand side platforms* (DSPs). A DSP allows advertisers to manage several ad-exchange accounts through one interface (such as DoubleClick, MediaMath, TubeMogul, or DataXu). On the supply side, content publishers embed requests for advertisements in their websites and applications. This ad space is offered through supply intermediaries: *advertising networks* (ANs) and *supply side platforms* (SSPs) sell online display inventory in multiple websites and apps (for example, Rubicon Project, Pubmatic, or Appnexus Publisher SSP). ANs typically sell inventory in bulk at a common price, whereas SSPs generally sell per impression at differentiated prices. Demand and supply intermediaries generally operate in an *ad-exchange* (a market clearing house which matches demand and supply orders through multiunit real-time bidding auctions over just a few tens of milliseconds). In addition, *data management platforms* (DMPs) provide information about Internet users to demand and supply middlemen. When a user arrives to a content website, at the same time that the publisher sends the content html code, it also submits an ad request to an SSP which can match the request to a cookie it has previously stored on the user's computer. This typically enables the SSP to track the user IP and retrieve some behavioral data (usually stored on an external DMP). The SSP submits an ad request to an ad-exchange; this generates a real-time auction, forwarding the request to several DSPs which place bids for the offered display impression on behalf of advertisers. Advertisers compete for targeted publishers in a *generalized second price auction* (a common type of auction; see e.g. Edelman, Ostrovsky, and Schwarz 2007; Varian 2009). Traditionally, the bid is per-click. However, it is possible to bid on a per-impression or per-click basis for the same website space. DSPs can also access a DMP to retrieve user information. These ad-tech markets are generally run by algorithms, and they are referred to as *programmatic advertising*. Figure A1 describes this ecosystem.

Figure A1: Ad-tech market structure



There exist several ad-exchanges. Intermediaries in the ad-tech market are often vertically integrated. For example, Google, one of the largest, owns an ad-exchange; on the supply side it also owns the SSP *DoubleClick for Publishers* (DFP) and the advertising network *AdSense*, while on the demand side it owns the DSPs *AdWords* for relatively low-budget operations and *DoubleClick Bid Manager* (DBM) for operations of higher budget.

This market is affected by multiple problems related to *asymmetric information* and *moral risk* which generate fraudulent behavior. Edelman (2009, 2010) reviews some of the pathologies in online display advertising platforms. A major problem in this market is that the lack of transparency exacerbates the information asymmetry. Ad-exchanges as well as other intermediaries generally use proprietary technologies that prevent advertisers from independently assessing the quality of advertising campaigns. To provide a credible signal, ad-tech middlemen usually promise some level of fraud detection, based on proprietary technology or authentication by third-party companies known as *verifiers* tracking the volume of fraud-free viewable impressions (e.g. Integral Ad Science, WhiteOps,

DoubleVerify, etc.[5]). But these auditors can have conflicts of interest (their main customers are ad-tech middlemen and content publishers), and their transparency is limited because they use undisclosed proprietary algorithms.

Display middlemen undertake some efforts to control click fraud. For example, Google AdWords has created an anti-click fraud program, filtering "invalid" clicks in real time, before publishers are charged. Google AdWords states that it classifies as invalid traffic accidental clicks providing no value (double-click by a user), clicks from robots, and manual clicks (to deplete the budget of an advertiser, to lower its click-through rate, or to artificially increase a publisher's revenues).[6] Besides click fraud, there is also impressions fraud. The process is transparent only to the extent that advertisers have the option to view the volume and percentage of invalid interactions over a period. The rigor of these preventive efforts is unclear (Tuzhilin 2006).

BRAND SAFETY

Table A1 lists the websites with more than 500 impressions in some of the campaigns, regardless of the *Approved* value. Some of these publishers no longer exist, but the description of their contents can be checked in the *Internet Archive Wayback Machine* (https://archive.org/web/).

[5] The respective URLS are: https://www.whiteops.com, https://integralads.com, and http://doubleverify.com.
[6] Google Ads Help, "About Invalid Traffic." Available at: https://support.google.com/adwords/answer/2549113?topic=10625 (accessed November 3, 2019).

Table A1: List of URLs with more than 500 ad impressions served

PUBLISHERS' WEBSITES	NUMBER OF IMPRESSIONS	PERCENTAGE OVER THE TOTAL	DESCRIPTION
vimeo.com	3904	3.76%	For hosting, sharing, and streaming videos
dailymotion.com	3082	2.97%	News site
sabiasdisso.com	3068	2.95%	Gossip news in English
euw.op.gg	2011	1.94%	Gaming site
fdating.com	1635	1.57%	Free online dating site
hltv.org	1040	1.00%	Gaming site
fortadpays.com	860	0.83%	Sell advertising site with revenue share.
educationuni.com	719	0.69%	Community college news
tecnologiala.info	716	0.69%	Tech news in Spanish
futwatch.com	612	0.59%	Gaming site
tiosarcasmo.com	583	0.56%	Peruvian web page, sarcastic and spicy humor
voip-service-providers-business.cf	576	0.55%	*Business VoIP* provider site
magesy.be	569	0.55%	News site of Audio Pro®
es.match.com	543	0.52%	Dating site
worldgamers.org	534	0.51%	Gaming site

The list of websites includes a sarcastic and spicy humor site, two dating sites, and four gaming sites. These publishers' websites implicitly entail a brand safety risk, as it is not unreasonable to expect that the context damages a consumer's perception of the advertiser launching the campaign. The only websites with impressions not recognized by the ad intermediary are the dating site *es.match.com* and the gaming site *euw.op.gg*. Impressions to the other sites are marked as *Approved*=0.

AUDITING CODE

The software and code are available upon request (Email to Rubén Cuevas, rcuevas@it.uc3m.es). The specific conditions to share the code will be studied on a case-by-

case basis and, will depend on the specific use of the code/software. For instance, commercial applications of the software/code will be subject to a contract and an economic compensation. Non-commercial use (e.g., for research purposes) would not require an economic compensation, but a non-disclosure agreement would be a prerequisite to share the software/code).

REFERENCES

Choi, Hana, Carl F. Mela, Santiago R. Balseiro, and Adam Leary (2017), "Online Display Advertising Markets: A Literature Review and Future Directions," Columbia Business School Research Paper No. 18-1. Available at SSRN: https://ssrn.com/abstract=3070706 (accessed August 8, 2018).

Edelman, Benjamin G. (2009), "Securing Online Advertising: Rustlers and Sheriffs in the New Wild West." In: *Beautiful Security* (ed. John Viega, O'Reilly Media, Inc.). Available at SSRN: https://ssrn.com/abstract=1267311 and http://dx.doi.org/10.2139/ssrn.1267311.

———(2010), "The Pathologies of Online Display Advertising Marketplaces," *CM SIGecom Exchanges*, 9(1), 1–5.

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

FireHOL (2017), "FireHOL IP Lists." Available at: http://iplists.firehol.org (accessed December 11, 2017).

Tuzhilin, Alexander (2006), "The Lane's Gifts v. Google Report." Available at: https://googleblog.blogspot.com.es/pdf/Tuzhilin_Report.pdf (accessed October 23, 2017).

Varian, Hal R. (2009), "Online Ad Auctions," *American Economic Review*, 99(2), 430–434.