Callejo, P., Cuevas, Á., Cuevas, R., Esteban-Bravo, M., & Vidal-Sanz, J. M. (2020). Tracking Fraudulent and Low-Quality Display Impressions. *Journal of Advertising*, 49 (3), pp. 309-319.

DOI:10.1080/00913367.2020.1749914

Tracking Fraudulent and Low-Quality Display Impressions

Patricia Callejo
IMDEA Networks Institute and Universidad Carlos III Madrid, Madrid, Spain

Ángel Cuevas
Universidad Carlos III Madrid and UC3M–Santander Big Data Institute, Madrid, Spain

Rubén Cuevas
Universidad Carlos III Madrid and UC3M–Santander Big Data Institute, Madrid, Spain

Mercedes Esteban-Bravo
Universidad Carlos III Madrid, Madrid, Spain

Jose M. Vidal-Sanz
Universidad Carlos III Madrid, Madrid, Spain

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 provide limited information about those measures. 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 by 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 thousand (CPT) 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.

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

There is limited research on display advertising fraud (Edelman 2014b; Fulgoni 2016). The most widely studied type of fraudulent behavior is click fraud. Click fraud covers a collection of techniques for artificially inflating the

Address correspondence to Jose M. Vidal-Sanz, Department of Business Administration, Universidad Carlos III Madrid, Calle Madrid 126, 28903 Gefate, Madrid, Spain. E-mail: jvidal@emp.uc3m.es

Patricia Callejo (MSc, Universidad Carlos III Madrid) is a doctoral student in telematics engineering, IMDEA Networks and Universidad Carlos III Madrid.

Ángel Cuevas (PhD, Universidad Carlos III Madrid) is a Ramón y Cajal fellow and assistant professor in the Department of Telematic Engineering, Universidad Carlos III Madrid; an adjunct professor at Institut Mines-Telecom SudParis; and a fellow at UC3M–Santander Big Data Institute.

Rubén Cuevas (PhD, Universidad Carlos III Madrid) is an associate professor in the Department of Telematic Engineering, Universidad Carlos III Madrid and deputy director of the UC3M–Santander Big Data Institute.

Mercedes Esteban-Bravo (PhD, Universidad Carlos III Madrid) is a professor of marketing and markets research, Department of Business Administration, Universidad Carlos III Madrid.

Jose M. Vidal-Sanz (PhD, Universidad Carlos III Madrid) is a professor of marketing and markets research, Department of Business Administration, Universidad Carlos III Madrid.

TABLE 1
Framework of Analysis and Fundamental Metrics

Concept	Metrics
Audience volume	
Data center impression fraud	• *DataCenter* (IP address belongs to a data center) • *Approved* (IP address approved as valid by the ad intermediary)
High-frequency fraud	• *User-impressions Intensity* (impressions received by a user within a minute) • *Interimpression Times*
Publisher popularity	• *Website Popularity*
Contextual targeting	
Strict keyword matching	• *Matching Keywords* (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 1 second)

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 temporarily expel rival advertisers from an 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 (e.g., inflating the count of times an ad is shown by including impressions on artificial users, usually automated traffic bots from data centers and botnets formed by malware-controlled computers which mimic human browsing behavior) or fake frequency (e.g., 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, some fraudulent tactics 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 by 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. A supplemental online appendix describes the display advertising industry in detail.

MEASUREMENT OF FRAUDULENT STANCES IN DISPLAY IMPRESSIONS

Our data are grounded in three advertising measurements that we have found to be important factors of online advertising effectiveness. The key measurements that we consider are summarized in Table 1: 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.

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 Internet Protocol (IP) address arriving at the advertiser site and receiving an ad impression, then matching the IP with an illegitimate data centers list. A data center is a physical or virtual infrastructure where a large group of computers is centralized to store, process, or distribute large volumes of data remotely. Associations such as the Media Rating Council (United States) and JICWEBS (United Kingdom) include data center traffic as a common source of invalid traffic and recommend filtering such traffic. Hence, the information technology (IT) community generates lists of centers with badly behaving 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 data centers, such as those channeling legitimate traffic originating from virtual private network (VPN) secure traffic. Integral Ad Science found that 8.3% of all U.S. digital display impressions were fraudulent (see Q1 2016 survey at https://integralads.com/). We computed two metrics for data center impression fraud, both dummy variables:

- *DataCenter*, which takes a value of 1 if impressions are served to IP addresses belonging to data centers, and 0 otherwise, using Botlab and FireHOL IP Lists (Botlab 2016; FireHOL 2017) to identify data center Ips.
- *Approved*, which takes a value of 1 if impressions are approved by the ad intermediary as valid.
- *Nonexcluded*, which takes a value of 1 if impressions are approved by the ad intermediary as valid but are tracked as fraudulent (to data center) by our auditing code. Nonexcluded 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 impressions), and the ad-exchange firm may report these as different valid impressions when, in fact, they are not unique from one another. Some digital signal processers (DSPs) allow advertisers to prevent this problem by including a frequency cap (i.e., a limit to the number of these impressions). Frequency capping is a way to prevent overexposure to an ad, but often it is used as a protective tool to deter data center–based fraud, as data center 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 a simple process to change it for nonskilled users. Our JavaScript code registers the impression time stamp (i.e., the time when a user arrives at an advertiser's site and receives an impression), and we again computed two metrics:

- *Interimpression Time*, the time between two consecutive impressions reported as valid by the ad intermediary (with $Approved = 1$), in seconds.
- *User-Impressions Intensity*, the quantity of valid impressions received by a user in one minute. Here, 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, Mullarkey, and Essegaier (2006) study the impact of website visit duration and the intervisit 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), in other words, 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

TABLE 2
Main Models

Model	Explained Variable	Regressors
1 (Logit)	Nonexcluded	Constant, CPT, Number of Impressions
2 (Quantile regression)	Interimpression Times	Constant, CPT, Number of Impressions, DataCenter
3 (Exponential model)	Website Popularity	Constant, CPT, Number of Impressions, DataCenter
4 (Logit)	Strict Keyword Matching	Constant, CPT, Number of Impressions, DataCenter
5 (Linear)	Leacock–Chodorow Similarity	Constant, CPT, Number of Impressions, DataCenter
6 (Exponential model)	Display Time	Constant, CPT, Number of Impressions, DataCenter
7 (Logit)	Viewability	Constant, CPT, Number of Impressions, DataCenter

Note. CPT = cost per thousand.

(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, *Matching Keywords*, which takes a value of 1 if at least one of the keywords assigned to the campaign matches a Universal Resource Locator (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 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

$$L{-}Ch\ Similarity = -log\ \left(length/(2 \times D)\right),$$

  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 (e.g., see Lin and Sandkuhl 2008). We use this measurement to study the similarity between the publisher's topics and the keywords of the campaign.
- *Brand Safety*, which categorizes websites where the impression is displayed using the web content as potentially negative for the advertiser.

Display Duration

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; in other words, click-through rates (CTRs) decrease.[3] Impression exposures are affected by fraudulent or low-quality impressions in cost-per-thousand (CPT) campaigns, and we considered two metrics:

- *Display Duration*, which uses a continuous variable, *Display Time* (impression duration), to measure how long an ad is active in a web page (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 U.S. campaigns, respectively).
- *Visual Perceptibility*, where our dummy variable, *Viewability* of impressions, takes a value of 1 when the impression display time is equal to or greater than 1 second, and 0 otherwise.

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 CPT and the number of impressions have an impact on fraud. The 2017 study by the ANA (Benes 2019) reports that fraud protection is not free, so the lowest CPTs 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 data centers, as discussed previously, our study analyzes the relationships between our metrics and CPT, the number of impressions, and whether the impressions are served to a data center. Braun and Moe (2013) examine the impact of ad impressions on visits and conversions.

The CPT, the number of impressions, and whether impressions are delivered to data centers variables will be used as predictors of fraudulent impression indicators (i.e., *Nonexcluded*, *Interimpression 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 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	CPT (Euros)	Keywords	Target Location
1	Research in Spain	5,117	350	March 29	March 31	0.10	Research	Spain
2	Research in Spain	42,398	1,776	March 29	March 31	0.20	Research	Spain
3	Research in Russia	4,096	274	March 29	March 31	0.01	Research	Russia
4	Research in the United States	1,178	135	March 29	March 31	0.01	Research	United States
5	Research in general	8,767	577	February 15	February 23	0.05	Universities, Research, Telematics	Spain
6	Research in general	42,359	1,548	February 18	February 23	0.10	Universities, Research, Telematics	Spain

Note. CPT = cost per thousand.

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 the United States) and CPT (0.01, 0.05, 0.10, or 0.20) in February and March 2016. We used a leading programmatic advertising intermediary that delivers display ads using Google AdWords (the largest advertising network available on the Internet, with more than 2 million publishers 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.

In total, the data set consists of 3,506 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; they are reported as URL = tpc.googlesyndication.com. In our database, 13.48% of 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 variance inflation factor (VIF) is smaller than 1.3 (the VIF for *CPT*, *Number of Impressions*, and *DataCenter*

TABLE 4
Summary Statistics of the Key Variables

Metric	Number of Observations	*M*	*SD*	Min	Max
Approved	103,916	0.42	0.49	0	1
Individual Impressions-Intensity	103,916	81.93	142.89	1	734
Interimpression 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
CPT	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

Note. CPT = cost per thousand.

are 1.12, 1.21, and 1.27, respectively) and the condition number is 6.28.

Nonexcluded

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

Table 6 shows that the probability of data center impressions being excluded by the ad intermediary is higher when the CPT and the number of impressions are smaller. This result suggests that the ad intermediary filter is stricter with smaller values of CPT and with fewer impressions. The ad intermediary identifies only 43,700 as valid impressions (*Approved* = 1). The 8.78% of these valid impressions (3,836) are served to data centers (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 data center 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.

Interimpression Times

The *Interimpression 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 interimpression times under 43 seconds. By campaign, the most dramatic case is Campaign 6, where 10% received the same ad within 4 seconds. Interimpression 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 CPT has a larger positive impact on the lower quantiles of *log(Interimpression Times)*. The 25th quantile of *log(Interimpression Times)* is more affected by CPT 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(Interimpression Times)*. The effect of the number of impressions is negative on higher quantiles (fewer impressions implies higher interimpression times). In addition, the quantile regression results indicate that the effect of data centers is much stronger at higher quantiles of *log(Interimpression Times)*. This suggests that high-frequency fraud is more prevalent when the campaign is not delivered to data centers. 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

TABLE 5
Pearson's Pairwise Correlation Coefficients of the Key Variables

Variables	1	2	3	4	5	6	7	8	9	10	11
1. Approved	1										
2. Individual Impressions-Intensity	−0.09*	1									
3. Interimpression Times	−0.03*	−0.01*	1								
4. Website Popularity	−0.17*	−0.06*	0	1							
5. Matching Campaign and Publisher URL Keywords	0.06*	0.02*	0	0.05*	1						
6. Leacock–Chodorow Similarity	−0.16*	0.03*	0	0.01	0.25*	1					
7. Display Time	0.01*	−0.02*	0	0.01	−0.01	−0.01	1				
8. Viewability	0.02*	−0.07*	−0.01*	0.01	0.01*	−0.01	0.03*	1			
9. CPT	0.21*	−0.17*	0.03*	−0.01	−0.11*	−0.18*	0.05*	0.05*	1		
10, Number of Impressions	−0.72*	0.10*	0.03*	−0.32*	−0.09*	−0.10*	−0.01*	−0.02*	−0.23*	1	
11. DataCenter	−0.25*	0.18*	0.03*	−0.02*	0.21*	0.13*	−0.02*	−0.05*	−0.30*	0.37*	1

Note. CPT = cost per thousand.

*Denotes correlation coefficients significant at the 1% level.

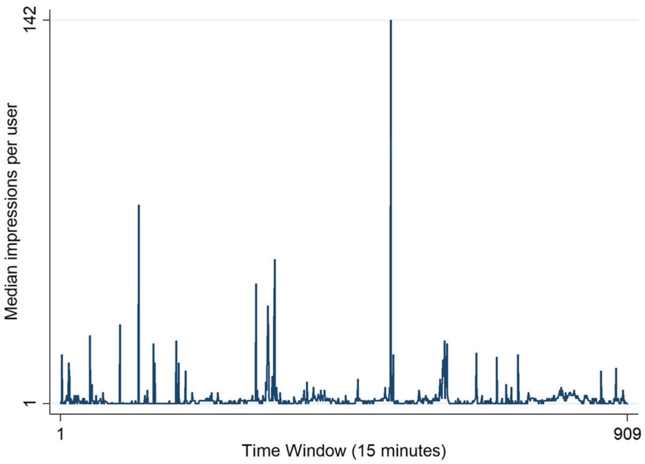

FIG. 1. Evolution in the median number of impressions per user during 15-minute windows.

not set up frequency caps, so we analyzed the default behavior of the ad intermediary. In this case, the data center 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 CPT, while holding all other variables constant; the website popularity decreases by 168.52% per increase of 1,000 impressions. Data center'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 CPT and the number of impressions are smaller. This result suggests that increasing CPT incentivizes the ad intermediary to display the ad in publishing sites less contextually relevant in terms of exact matching. CPT 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

TABLE 6
Estimates of the Main Models (Table 2)

		Model 2							
	Model 1	(Quantile 0.25 Regression)	(Quantile 0.50 Regression)	(Quantile 0.75 Regression)	Model 3	Model 4	Model 5[a]	Model 6	Model 7
Regressors	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
CPT	−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)
Number of 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. CPT = cost per thousand.

[a] We use White's robust to heteroskedasticity standard errors.

[b] Likelihood ratio chi-square test. The p value is 0.00; the null is rejected.

[c] F statistics test. The p value is 0.00; the null is rejected.

indicates that the probability of exact matching for approved impressions is higher when the impressions are delivered to a data center. This result suggests that when considering a campaign on a specific topic, for example, "sports," the ads that appear on "sports" pages are more likely to be delivered to users who come from data centers. The campaigns that we have configured are based on context; 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; 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 match when the number of impressions increases; 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 CPT plays a relevant role in display contextual relevance, and that the best result for the publisher is obtained for intermediate CPT 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 CPT, 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 supplemental 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 *Display Time* increases by 8.33% per one unit increase in CPT, while holding all other variables constant. The number of impressions and whether they come from a data center or not have no significant effect on *Display Time*.

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 CPT, 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 data centers (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 data center traffic. In addition, there are considerably high levels of potentially fraudulent impressions due to high frequency (50% of the total interimpression 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 CPT. Our data suggest that the probability of (several types of) hidden fraud is related to CPT and number of impressions. We found that campaigns with the highest CPT 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 CPT. 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 from the authors upon request (see the supplemental online appendix for details).
2. Do not bid too low. If the ad is displayed, the level of hidden fraud might be considerably higher if the CPT is low. Intermediaries may not use sophisticated fraud detection tools when the fraud cost 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 supplemental online appendix). Some firms are already finding out about this problem. For example, JPMorgan Chase used to display ads on more than 400,000 websites monthly, but after recently detecting display impressions placed next to toxic content it dramatically cut its displays to 5,000 preapproved 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 ad networks 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, and nonparametric and machine learning methods) These approaches provide useful robustness checks, and future research might explore these avenues with larger samples.

FUNDING

This work is partly supported by the European Union through SMOOTH (786741) and PIMCITY (871370); the European Social Fund through Ramón y Cajal (RYC-2015-17732); and the Spanish Ministry of Economy and Competitiveness through ECO2015-67763-R and PGC2018-096083-B-I00 projects.

SUPPLEMENTAL MATERIAL

A supplemental online appendix is available on the publisher's website.

NOTES

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

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

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

[4] *New York Times*. A version of this article appeared in print on March 30, 2017, on p. 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

Benes, Ross (2019), "Five Charts: The State of Ad Fraud," *eMarketer*, May 20, www.emarketer.com/content/five-charts-the-state-of-ad-fraud (accessed December 3, 2019).

Botlab (2016), "Botlab.io Deny-Hosting IP List," 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), 753–67. doi:10.1287/mksc.2013.0802

Dahlén, Micael (2005), "The Medium As a Contextual Clue: Effects of Creative Media Choice," *Journal of Advertising*, 34 (3), 89–98. doi:10.1080/00913367.2005.10639197

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), 182–94. doi:10.1509/jmkr.43.2.182

Deane, J., and A. Agarwal (2012), "Scheduling Online Advertisements to Maximize Revenue under Variable Display Frequency," *Omega*, 40 (5), 562–70. doi:10.1016/j.omega.2011.11.001

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. doi:10.1080/00913367.2002.10673666

Dreze, X., and F.-X. Hussherr (2003), "Internet Advertising: Is Anybody Watching?," *Journal of Interactive Marketing*, 17 (4), 8–23. doi:10.1002/dir.10063

Edelman, Benjamin G. (2014a), "Accountable? The Problems and Solutions of Online Ad Optimization," *IEEE Security and Privacy*, 12 (6), 102–107. doi:10.1109/MSP.2014.107

——— (2014b), "Pitfalls and Fraud in Online Advertising Metrics," *Journal of Advertising Research*, 54 (2), 127–32 (accessed November 8, 2019).

———, and Wesley Brandi (2015), "Risk, Information, and Incentives in Online Affiliate Marketing," *Journal of Marketing Research*, 52 (1), 1–12. doi:10.1509/jmr.13.0472

Enberg, Jasmine (2019), "US Digital Ad Spending 2019," *eMarketer*, March 28, https://www.emarketer.com/content/us-digital-ad-spending-2019 (accessed November 3, 2019).

Fisher, Lauren (2019), "US Programmatic Ad Spending Forecast 2019," *eMarketer*, April 25, www.emarketer.com/content/us-programmatic-ad-spending-forecast-2019 (accessed November 3, 2019).

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

Fulgoni, Gian M. (2016), "Fraud in Digital Advertising: A Multibillion-Dollar Black Hole: How Marketers Can Minimize Losses Caused by Bogus Web Traffic," J*ournal of Advertising Research*, 56 (2), 122–25. doi:10.2501/JAR-2016-024

Goldfarb, Avi (2014), "What Is Different about Online Advertising?," *Review of Industrial Organization*, 44 (2), 115–29. doi:10.1007/s11151-013-9399-3

———, and Catherine Tucker (2011), "Online Display Advertising: Targeting and Obtrusiveness," *Marketing Science*, 30 (3), 389–404. doi:10.1287/mksc.1100.0583

IAB Tech Lab (2016), "Desktop Display Impression Measurement Guidelines," https://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. doi:10.1109/MC.2007.232

Koenker, Roger, and Gilbert Bassett, Jr. (1978), "Regression Quantiles," *Econometrica*, 46 (1), 33–50. doi:10.2307/1913643

Leacock, Claudia, and Martin Chodorow (1998), "Combining Local Context and WordNet Similarity for Word Sense Identification," in *WordNet: An Electronic Lexical Database*, Christiane Fellbaum, ed., Cambridge, MA: MIT Press, 265–83.

Lin, Feiyu, and Kurt Sandkuhl (2008), "A Survey of Exploiting WordNet in Ontology Matching," in *Artificial Intelligence in Theory and Practice II*, Max Bramer, ed., vol. 276, Boston, MA: Springer, 341–50.

Moe, Wendy W., and Peter S. Fader (2004), "Dynamic Conversion Behavior at E-Commerce Sites," *Management Science*, 50 (3), 326–35. doi:10.1287/mnsc.1040.0153

Zhang, Weinan, Ye Pan, Tianxiong Zhou, and Jun Wang (2015), "An Empirical Study on Display Ad Impression Viewability Measurements," https://arxiv.org/pdf/1505.05788.pdf.