

Pitfalls and Fraud In Online Advertising Metrics

Are Cheaters Hurting Your Bottom Line?

BENJAMIN EDELMAN

Harvard Business School
bedelman@hbs.edu

Editor's Note: In the December 2009 issue of the *Journal of Advertising Research* (49.4), Benjamin Edelman's "Commentary: Who Owns Metrics? Building a Bill of Rights for Online Advertisers," pp. 401–403) questioned industry fair practice on the use of metrics. Here, Dr. Edelman expands on his research with new insights on fraud and errors in advertising metrics.

How does online advertising become less effective than advertisers expect and less effective than measurements indicate? The current research explores problems that result, in part, from malfeasance by outside perpetrators who overstate their efforts to increase their measured performance. In parallel, similar vulnerabilities result from mistaken analysis of cause and effect—errors that have become more fundamental as advertisers target their advertisements with greater precision. In the paper that follows, the author attempts to identify the circumstances that make advertisers most vulnerable, notes adjusted contract structures that offer some protections, and explores the origins of the problems in participants' incentives and in legal rules.

INTRODUCTION

A century after retail pioneer John Wanamaker wondered which half of his advertising spending was wasted, advertisers have reason to think that a solution is within reach. Modern online advertising placements purport to track which users see which offers, who clicks, and who buys—a roadmap for vastly improved targeting, greater efficiency and, one imagines, reduced cost.

Despite these high hopes, the short-run reality is challenging. This paper, in fact, identifies four key challenges:

- Some perpetrators intentionally drain advertisers' budgets. When fraudsters manipulate advertising systems, their efforts often may appear highly effective but, in fact, invite advertisers to invest in placements that are ineffective.
- Measurement systems may be less reliable than they seem. When purchases are unpredictable,

even precise measurements may not identify true cause and effect. In some cases, optimization may exacerbate faulty measurement.

- Advertising brokers and networks have mixed incentives. They sometimes genuinely seek to improve advertisers' efforts, but often their interests are at odds with what's best for advertisers.
- Standard methods of accountability and dispute resolution have proven ineffective; advertisers have struggled to hold perpetrators and intermediaries accountable for apparent breaches.

Despite significant improvements over alternatives and predecessors, online marketing is far from the panacea some have expected. Through the methods proposed in the current study, however, advertisers at least can know where they are most vulnerable—and sometimes take action to protect themselves.

Despite significant improvements over alternatives and predecessors, online marketing is far from the panacea some have expected.

THE SCOPE OF THE PROBLEM

Widespread “Banner Farming”

A decade ago, display advertisers already faced the problem of invisible banners—advertisements loaded in invisible windows that users could not see. When advertisers pay “cost per mille” (CPM), in proportion to the number of times their advertisements are loaded, predictably they are vulnerable: A rogue “banner farm” site can load scores of banners invisibly and get paid accordingly.

Though shockingly simple, this technique remains widespread. For example, throughout 2013, the author tracked the French site Realtoreal.fr as it loaded eight other pages in eight separate IFRAME subwindows, each carefully configured with Web page code instructing “display:none” to make the subwindows invisible. Each of those eight other pages then loaded several banner advertisements of their own, including banners promoting top brands. As a result, a single request for this single site loaded dozens of banner ads, almost all invisibly.

One might hope that banner farms would be quickly uncovered thanks to their exceptionally low click-through rate. (Users cannot click what they do not see.) Even on legitimate sites, however, most banners attract few clicks—on average just 0.10 percent, according to Google Double-Click Display Benchmarks (2014), though click-through rates are an order of magnitude lower for certain users and sites. It is difficult to distinguish this low click-through rate from literally zero; even after 10,000 views with no clicks, it is hard to tell

whether a site’s advertisements are invisible or simply disfavored by users.

Meanwhile, banner farms can change names and domains and rotate their tainted inventory through hundreds of networks. The savviest banner farms mix in occasional clicks from genuine users, legitimate sites, botnets, or paid staff—increasing measured click-through rate and circumventing some scrutiny. Particularly when combined with networks’ incentive to look the other way (as discussed below), the most successful banner farms run for months or even years.

Paying Per Click

A natural advertiser response is to pay not when advertisements are loaded and (purportedly) seen but when users click. Indeed, search advertising almost entirely uses click-based pricing. In some respects, this approach has proven successful; certainly advertisers find many genuine leads at top search engines.

Despite the benefits of paying per click, that model is also under attack. To increase the volume of inventory they can offer to advertisers, top advertising networks hire third-party sites to syndicate advertisements. One would not imagine Google staff clicking on advertisements shown at Google, even if that increased Google’s short-term revenue. At smaller syndicator sites, however, such practices cannot be ruled out—conceivably simple click fraud by staff, though more likely schemes such as botnets, pop-ups, simulated advertisement clicks, or adware inserting extra advertisements.

By measuring its conversion rate—the proportion of clicks yielding actual purchases—an advertiser can try to adjust its bids in light of apparent effectiveness. Or, if the network permits, the advertiser may try to reject traffic from low-performing sources. Many advertisers, however, struggle to measure conversions, particularly when selling offline or with a long sales cycle.

Cheating on Pay-per-Sale Commissions

For advertisers able to measure conversions, it is tempting to pay only when a purchase occurs—an approach often known as affiliate marketing, pay-per-sale (PPS), or cost-per-acquisition. In 2005, *The Economist* captured the promise of this approach, noting that a crisp alignment of advertisers’ interests with networks would be “the holy grail of online advertising.”

As it turns out, PPS advertising nonetheless can be gamed. Suppose a cheater can identify users likely to buy from a given merchant; perhaps the cheater placed monitoring software on users’ computers to monitor their every move and intercede when the users requested that merchant’s site, or perhaps the merchant is so popular that many users buy there with reasonable frequency. Once the cheater can find users likely to buy from that merchant, the cheater need only claim to have referred those users, then await PPS commissions. PPS payments are strikingly high: 10 percent of purchase price is normal in many sectors.

A cheater can reap significant rewards by claiming credit for even a modest number of customers. The largest cheaters take this scheme to exceptional lengths: In 2012–2013, two affiliates pled guilty to taking some \$21 million from eBay’s affiliate program via these methods. Before their scheme was uncovered, they had been eBay’s two largest marketing affiliates. Nordstrom and Lands’ End found

affiliates draining their budgets through similar schemes. For summaries of all affiliate disputes that culminated in litigation, see Edelman (2014).

Some advertisers and networks try hybrid payment models and other adjustments (Edelman and Lee, 2008; Zhu and Wilbur, 2011), but there is no silver bullet. Tinkering with payment models will not fix online advertising misbehavior or prevent low-value placements (Table 1).

MEASUREMENT PRECISION AND ACCURACY

Assessing Cause and Effect When Base Rates Are Low

Even when publishers do nothing to inflate their supposed effectiveness, advertisers often still struggle to measure campaign benefits. One key problem is that many advertisers find customers' purchases nearly random and hence exceptionally difficult to predict. Even if an advertising campaign successfully increases customers' purchase probabilities, the effect can blend with background variation in purchase events.

In one numerical example drawn from typical banner advertisers at Yahoo (Lewis, Reiley, and Rao, 2013), suppose users make purchases from a given advertiser with mean of \$7 and standard deviation of \$75. (Most users buy nothing, pulling down the mean and making the standard deviation much larger than the mean.) Suppose, further, that the advertiser is prepared to spend \$0.14 on advertising per user, an amount sufficient to buy 20 to 100 display advertisements at \$1 to \$5 CPM.

With gross margin of 50 percent, a percent increase in sales is worth \$0.35 per user of revenue and \$0.175 of profit, which would be a 25-percent return on investment over the \$0.14 spent on advertising. It is unrealistic however, to detect a \$0.35 increase in spending relative to a standard deviation of \$75 of sales. If the advertising

TABLE 1
Benefits and Weaknesses of Alternative Pricing Models

Advertising Pricing Model	Key Benefits to Advertisers	Key Risks
Pay-per-impression	Many advertisers care how many times their advertisements are seen. An advertisement may still help build a brand and facilitate a purchase even if a user doesn't click.	Advertisements loaded invisibly. Advertisements targeted to an advertiser's existing customers, who were going to buy anyway.
Pay-per-click	Helps confirm that a user is genuinely interested in an advertiser's offer.	Fake clicks. Advertisements targeted to an advertiser's existing customers, who would buy the product even if the advertiser did not advertise.
Pay-per-sale	An advertiser's expenditure directly matches purchases. Reduces apparent risk and facilitates partnerships with sites whose productivity might otherwise be difficult to verify.	This payment model invites reduced supervision of partners. Partners may systematically refer customers who would have bought anyway.

campaign achieved exactly that level of success, its R^2 would be just 0.0000054, and 2 million—evenly and randomly spread between the test and control group—users would be needed to prove the effectiveness at the 5-percent level of statistical significance. Even adjusting Lewis's estimates, it remains difficult to distinguish effective from ineffective advertisements when advertisements are shown to a broad swath of untargeted users.

In response to Lewis and as a standard strategy to improve advertising cost-effectiveness, most advertisers target advertisements toward users who are perceived as likely to buy—perhaps users whose demographics match an advertiser's target customers or users whose search requests match what the advertiser sells. The prevailing wisdom is that better-targeted advertisements yield

superior results—and, in turn, more sales at a lower cost.

The Problem of Excessive Targeting

As it turns out, there are multiple factors at play. One key complication is that improved targeting increases the risk of showing an advertisement to consumers who would have purchased anyway. In a controlled experiment that disabled portions of eBay's paid search campaign, returns from paid searches were a fraction of what standard reporting systems indicate because many customers who click paid advertisements would have bought from eBay even without the advertisements (Blake, Nosco, and Tadelis, 2013).

Consider a user who searches for "eBay laptop" and clicks an advertisement for eBay. That user probably would have found eBay's site even without an

advertisement: After all, the user asked for eBay specifically. Indeed, even a user who searches for “used laptop” would buy from eBay with some probability well above zero; once a user realizes that a used laptop might be a fine choice, it is no great leap to look on eBay. Yet, standard reporting consoles would attribute both of these sales solely and entirely to eBay’s search engine advertising.

In particular, standard analysis assumes, falsely, that if eBay had not bought those paid clicks, these users would not have made the specified purchases from eBay. Correcting this assumption via randomized experiments that vary which advertisements run where and for what keywords, brand-keyword advertisements—for many advertisers, purportedly the most profitable portion of a paid search campaign—actually have negative value (Blake *et al.*, 2013). Even such generic search phrases as “used laptop” often become borderline or, in any event, far less profitable.

It is tempting to write off the Blake result as an anomaly. For example, critics point out that the work is limited to a single advertiser and its campaign strategy. Indeed, a humor site, “eBads,” tracks ineffective eBay advertising. (A representative example: “Dirty Socks—Looking for Dirty Socks? Find exactly what you want today. www.ebay.com.”) If eBay’s advertising systematically is less effective, competitors might enjoy better results.

Other aspects of the conclusions ring true. For example, the better known an advertiser, the greater the chance that buyers would buy at its site even without advertising. Moreover, technological advances cause these problems to grow beyond search advertising to affect other advertising formats also.

Consider an advertiser that uses “retargeting” to show banner advertisements to users who recently have visited its

The higher the baseline purchase probability, from popular brands and better targeting, the greater the overstatement of advertising platform effectiveness.

site. Having visited, say, gap.com at some point in the last week, a given user already has an above-average chance of making a purchase from that site. Through retargeting, a display-advertising network can show Gap’s advertisements to users who recently visited the Gap site.

Retargeting may increase those users’ purchase probability somewhat, but a retargeting vendor should claim credit for, and should be paid for, only the incremental increase in purchase probability—not the overall above-average likelihood of purchase. In the realms of both retargeting and paid search, standard reporting consoles tend to claim credit for the entire purchases of all users who see or click the targeted advertisements, not just the incremental purchases.

In short: The higher the baseline purchase probability, from popular brands and better targeting, the greater the overstatement of advertising platform effectiveness.

Incentives of Providers and Brokers

Most advertising networks face conflicting incentives on fraud and measurement. In the long run, advertisers distrust networks found to offer low-quality placements and misleading measurements of advertising effectiveness. In the short run, however, advertisers are paid in proportion to advertiser spending.

For example, for placements within its Content Network, Google keeps 32 percent of an advertiser’s payment (Google, 2013), while Commission Junction takes a 23 percent share of fees in its affiliate network (Commission Junction, 2004). As

each network’s fees are proportional to its advertisers’ spending, networks have a direct short-run incentive to tolerate and retain even low-quality placements. Large-sample data confirm that networks act on this incentive. Moreover, prior research indicates that when an affiliate network manages a merchant’s affiliate marketing program, the program more often suffers certain “gray area” practices that drain the advertiser’s budget without yielding incremental sales (Brandi and Edelman, 2013).

Major advertisers might hope that sophisticated media buyers should protect them both from fraud and from errors in judgment. In many instances, however, this is an overly optimistic assumption. For one, media-buying organizations operate on slim margins. A generation ago, media buyers often enjoyed a standard 15-percent commission—a large fee relative to the effort required to oversee a limited number of large advertising purchases such as national television networks. More recently, media-buying services are fiercely competitive, with fees typically driven to low, single-digit percentages. With such low fees, ad buyers struggle to allocate sufficient staff to examine—not to mention evaluate—the myriad placements advertisers expect.

Moreover, online campaign management is much more complicated: A typical display advertiser buys advertisements from thousands of sites via at least half a dozen networks, whereas a paid search advertiser often bids on thousands of keywords. Measuring and optimizing these campaigns requires significant

data-processing capabilities—a sea change from agencies’ modest technical capabilities even a decade ago.

In parallel, media-buying services face pressure to demonstrate success in measured performance: Advertisers inevitably want more views, more clicks, or more reported purchases, and they look to ad buyers to deliver these results. Savvy media buyers and advertisers are broadly aware of the fraud and measurement issues flagged in the preceding sections. Indeed, when circumstances permit, most advertisers examine multiple metrics—not just advertisement views but clicks and apparent purchases.

When under pressure to meet promised metrics, the advertiser might focus on a single metric, perhaps pushing the media buyer to increase views without increasing cost. A site selling cheap but invisible advertisements might come up short under multi-faceted scrutiny, which would reveal its suspect practices. Once the advertiser focuses only on measured views, however, that site looks great: It is willing to sell more impressions at a lower price. If the advertiser favors or accepts these measurements, the media-buyer has every incentive to give its support, even if it knows the numbers are not truly reliable.

Advertising network incentives also invite advertisers to repeat the blunder (Blake *et al.*, 2013) found in eBay’s search marketing efforts. In principle, a network’s reporting console could examine the true incremental effect of an advertising campaign—automating the A/B testing that lets the advertiser check whether advertising brings in new customers. Google would be uniquely positioned to offer such tools due to the spillover from algorithmic results to sponsored search and the company’s overall emphasis on measurement and testing.

By all indications however, the results would be unfavorable to Google and

would lead advertisers to reduce their spending on search advertising. No wonder Google’s widely used Analytics platform and standard AdWords reporting console include no such function and, indeed, never mention or even suggest the risk of advertising reaching customers who would have bought anyway.

Accountability and Dispute Resolution

One might hope that contracts and potential legal liability would hold advertising networks to the highest standards and, at the very least, to their written commitments.

Experience reveals otherwise.

When pressed, networks successfully have disavowed their written statements to advertisers. For example, in 2011 litigation,¹ advertisers questioned Google’s decision to withhold “smart pricing” discounts that had been understood to reduce the cost of low-performing ad placements. They also challenged Google’s placement of advertisements through certain controversial and undesirable partners.

And Google offered two notable responses:²

- Google claimed it had never “promised” to provide the smart pricing discounts. Though Google staff and its Web sites often discuss these discounts, statements often were ambiguous. In a representative example, a Google training document stated that “smart pricing ... may reduce your cost”—a promise which, in Google’s view, did not require discounts in any particular circumstances or, indeed, any discounts at all; rather, Google said the word “may” meant discounts were at its sole discretion.

¹ *Woods v. Google*. U.S. District Court, Northern District of California. Case No. 5-11-CV-01263-EJD. Complaint filed March 15, 2011.

² *Woods v. Google*. U.S. District Court, Northern District of California. Case No. 5-11-CV-01263-EJD. Motion to Dismiss First Amended Complaint. October 8, 2011.

Even when an ad network’s statements are written, clear, and unambiguous, they still may not be legally enforceable.

- Google denied that its Web-site statements—in the AdWords Help site, technical documentation, and otherwise—formed part of its contract with advertisers. Rather, Google argued that its only binding obligations were those within the AdWords Terms and Conditions (T&C’s) document.

In particular, Google claimed that “the [T&C’s] agreement does not incorporate” the additional statements Google made elsewhere and, thus, such statements could not be grounds for advertisers’ breach of contract claims.

Judge Edward Davila substantially accepted Google’s arguments.³

Advertisers naturally rely on statements from ad networks—the natural and obvious way to understand network policies. But Google’s successful arguments show a danger of this approach: Statements on help pages and technical documentation may not be legally enforceable even if they are written, preserved, and unambiguous.

Meanwhile, networks have important advantages in their legal relationships with advertisers. For one, a large network typically insists that the contract follow

³ *U.S. District Court, Northern District of California. Case No. 5-11-CV-01263-EJD. Order Granting-In-Part and Denying-In-Part Motion to Dismiss First Amended Complaint. August 24, 2012.*

a non-negotiable, prewritten agreement that the network drafted. The largest and most sophisticated advertisers seek alternatives but, in online search, options are particularly limited: Most advertisers need Google far more than Google needs them. Furthermore, networks have the benefit of specialization and repeat transactions; a network can pay attorneys to assess its vulnerabilities, then write its ideal contract once.

In contrast, most advertisers' attorneys are generalists who must handle myriad other matters. The Internet Advertising Bureau (IAB; 2009) offers "Standard Terms and Conditions for Internet Advertising for Media Buys" that reasonably apportion rights and responsibilities between advertisers and publishers. If a network insists on its own contract and rejects the IAB's standard, advertisers should be appropriately skeptical.

MOVING FORWARD

With opportunities for both malfeasance and faulty measurement, it is tempting to renounce online advertising. Yet, the efficiencies are genuine and important. One ought not lose sight of the promise of advertisements better targeted to the right people, with easier adjustment and testing and greater efficiency throughout. These are real benefits, and every forward-thinking advertiser is moving in this direction.

As to the problems the author has identified, advertisers should begin by recognizing their vulnerabilities. Long-standing principles of incentives and ethics have

not disappeared merely because some behaviors can be measured more precisely. Ambiguities in cause and effect do not disappear when data are collected in more granular form. These problems need not be fatal; for generations, managers devised mechanisms to evaluate and optimize performance despite these challenges. Yet, some online advertising practices reflect an aura of infallibility: Advertisers often accept exceptional risks, as if online ad markets are invulnerable to long-known weaknesses. In fact, these problems persist, and advertisers that ignore the situation can find themselves all the more vulnerable. **JAR**

.....
BENJAMIN EDELMAN is an associate professor at the Harvard Business School. He holds a PhD degree from the department of economics at Harvard University and a JD degree from the Harvard Law School, and he is admitted to the Massachusetts Bar. His writings are posted at www.benedelman.org.

REFERENCES

- BLAKE, T., C. NOSKO, and S. TADELIS. (2013). "Consumer Heterogeneity and Paid Search Effectiveness: A Large-Scale Field Experiment." Mimeo. Retrieved from <http://faculty.haas.berkeley.edu/stadelis/Tadelis.pdf>
- BRANDI, W., and B. EDELMAN. (2013). "Information and Incentives in Online Affiliate Marketing." HBS Working Paper 14-041.
- COMMISSION JUNCTION. (2004). "About CJ Access." Retrieved from <http://www.cj.com/>

[solutions/adv_access.jsp](http://www.cj.com/solutions/adv_access.jsp) (archived page was posted from 2004 to 2006 and is now preserved by Archive.org).

"eBADS—THE eBAY AD SEARCH GAME." Retrieved from <http://www.dealcafe.com/funnies/searchgame.html>.

EDELMAN, B. (2014). Affiliate Litigation Index. Retrieved from <http://www.benedelman.org/affiliate-litigation>.

EDELMAN, B., AND H. LEE. (2008). "CPC/CPA Hybrid Bidding in a Second Price Auction." HBS NOM working Paper 09-074.

GOOGLE DOUBLECLICK DISPLAY BENCHMARKS. (2014). Retrieved from <http://www.richmediagallery.com/resources/benchmarks/>.

GOOGLE. AdSense Revenue Share. (2013). Retrieved from <https://support.google.com/adsense/answer/180195?hl=en>.

INTERNET ADVERTISING BUREAU. (2009). "Standard Terms and Conditions for Internet Advertising for Media Buys." Retrieved from <http://www.iab.net/guidelines/508676/tscs3>.

LEWIS, R., D. REILEY, and J. RAO. "Measuring the Effects of Advertising: The Digital Frontier." In *NBER Economics of Digitization*, S. Greenstein, A. Goldfarb, and C. Tucker, eds. Chicago: University of Chicago Press, 2014.

ZHU, Y., and K. WILBUR. "Hybrid Advertising Auctions." *Marketing Science* 30, 2 (2011): 249–273.