

Accountable?

The Problems and Solutions of Online Ad Optimization

Benjamin Edelman | Harvard Business School

Online advertising might seem to be the most measurable form of marketing ever invented. Comprehensive records can track who clicked what ad—and often who saw what ad—to compare those clicks with users' subsequent purchases. Ever-cheaper IT makes this tracking cost-effective and routine. In addition, a web of interlocking ad networks trades inventory and offers to show the right ad to the right person at the right time. It could be a marketer's dream.

However, these benefits are at most partially realized. The same institutions and practices that facilitate efficient ad placement can also facilitate fraud. The networks that should be serving advertisers have decidedly mixed incentives, such as cost savings from cutting corners, constrained in part by long-run reputation concerns, but only if advertisers ultimately figure out when they're getting a bad deal. Legal, administrative, and logistical factors make it difficult to sue even the

worst offenders. And sometimes an advertiser's own staff members prefer to look the other way. The result is an advertising system in which a certain amount of waste and fraud has become the norm, despite the system's fundamental capability to offer unprecedented accountability.

Let me pause to offer key terminology. Online advertisers can promote almost anything, including goods and services sold through their websites, items for offline purchase, and information. To present their messages to more users, advertisers typically turn to *networks*—companies that assemble groups of websites (publishers) for advertisers to purchase en masse. This grouping simplifies payment: if n advertisers buy ads at m websites, $n \times m$ payments might be required, but when a single ad network facilitates those placements, only $n + m$ payments are needed. Grouping also facilitates optimization: if the network can measure ad performance, it can place an advertiser's

offers on the sites where the ad works best, or automatically adjust the price to reflect some measure of quality. Of course, ad networks can bring complications and facilitate abuse. This is compounded when one network hires another, which hires another. Meanwhile, large advertisers also hire ad agencies to implement and optimize their campaigns—this is an appropriate and necessary specialization, yet also a source of divergent incentives.

Throughout this article, I focus on measuring and optimizing online ad placements. There's plenty more to improve about online advertising—for example, blocking illegal and deceptive advertising and addressing the serious privacy concerns raised by pervasive tracking. But progress on those issues has been even slower, and I'll leave those tasks to others.

Display Advertising

Banner ads, among the earliest types of large-scale online advertising, are the ubiquitous groundwork of the industry. Advertisers usually pay a modest fee each time their ads are shown, such as US\$5 per thousand displays. To advertisers, this feels refreshingly familiar—not unlike paying for newspaper or radio ads, where the price is set in proportion to the number of customers reached. It's equally logical for website publishers, as they can easily compare offers from a variety of advertisers to select the ad that will yield the greatest revenue.

Yet banner ads have severe incentive problems. What stops a site from filling its pages with excessive advertising? Doubling the ads

will yield double the payments from advertisers, so publishers have every incentive to load more ads. Some stack ads on top of ads—“interstitial” popups covering standard banners, or multiple banners on top of each other. In the worst cases, sites might adjust their code to load ads invisibly. This turns out to be surprisingly easy, as seen in Figure 1.

Shrewd advertisers have some defenses against these schemes. Most obviously, an advertiser could watch for ads with low click-through rates (CTR). Whether infrequent clicks result from excessive advertising, invisible ads, or a poor match between a site’s users and an advertiser’s offer, the advertiser might be skeptical of ad placements that few users click. But as banner ads became widespread and users trained themselves not to click, low CTR quickly became the norm. Even on legitimate sites, one click per thousand impressions is now roughly average. And for “brand” advertisements that tout products sold in stores, it may be unrealistic to expect users to click. Perhaps advertisers don’t even need or want clicks: a banner ad for Tide reminds customers of that product regardless of whether they click on it.

Finding CTR an unsatisfactory method of evaluating website quality, some advertisers reverted to buying only from trusted sites. This has a certain appeal in that top-quality publishers aren’t likely to cheat. If an advertiser buys directly from those sites, the advertiser is relatively well protected, but direct buys from top publishers usually carry a premium price that scares off most advertisers. As a result, advertisers often prefer to buy placements on top publishers’ sites via networks and ad exchanges (networks of networks) that sell excess inventory at a discount. These intermediaries can lead to misrepresentation. Figure 2 shows an example

```

GET http://sharelien.fr/xtcjp131 HTTP/1.1 (a)
...
HTTP/1.1 200 OK (b)
<html><head>
<IFRAME (c)
SRC=http://infobelge.be/news3.php?mobile=1
width=800 height=800 FRAMEBORDER=0
SCROLLING=NO style="display:none;"></IFRAME> (d)
<IFRAME
SRC=http://infobelge.be/news3.php?mobile=1
width=800 height=800 FRAMEBORDER=0
SCROLLING=NO style="display:none;"></IFRAME> (e)
<IFRAME
SRC=http://infobelge.be/news2.php?mobile=1 (f)
width=800 height=800 FRAMEBORDER=0
SCROLLING=NO style="display:none;"></IFRAME>
<IFRAME
SRC=http://infobelge.be/news2.php?mobile=1 (g)
width=800 height=800 FRAMEBORDER=0
SCROLLING=NO style="display:none;"></IFRAME>
<IFRAME
SRC=http://infobelge.be/news.php?mobile=1
width=800 height=800 FRAMEBORDER=0
SCROLLING=NO style="display:none;"></IFRAME>
<IFRAME
SRC=http://infobelge.be/news.php?mobile=1
width=800 height=800 FRAMEBORDER=0
SCROLLING=NO style="display:none;"></IFRAME>
...

```

Figure 1. Loading ads invisibly. (a) Request from user’s browser to remote server. (b) Start of response from server. (c) An IFRAME inline frame loads one Web page inside another. (d) Viewing this “Sharelien” site entails loading the separate Infobelge site six different times, meaning more opportunities for the site operator to load ads and get paid. (e) The Infobelge site is loaded in windows set with `display:none`, telling the Web browser not to show the resulting content. (f) Although these size parameters seem to specify a 800 x 800 window, the `display:none` command takes precedence. (g) Advertisers may wonder why their ads aren’t clicked. The Web page appendage `mobile=1` will tell advertisers their ads are on mobile devices, where clicks are understood to be less frequent. Sharelien simply asserts `mobile=1` whether or not that’s true, but advertisers are unlikely to realize this.

of a rogue publisher falsely claiming to sell inventory from top sites like about.com and Yelp. The fraudster was actually buying cheap, indeed invisible, traffic and adding a fabricated “referrer=” parameter to the ad call, falsely telling advertisers and networks that these were high-quality placements from the best publishers. Anyone who examined the code could uncover the scheme, so the fraudster distributed it in encoded form (shown in part in the middle of Figure 2. I’ve tracked

this perpetrator for months on end, reporting his malfeasance to every advertiser and network that asked. Yet by all indications he or she is still going strong, occasionally shifting to new servers and new accounts but retaining the same scheme and even reusing most of the same code.

Paying for Results

The crux of the problem with display advertising is that advertisers pay for one thing (views of its ads)

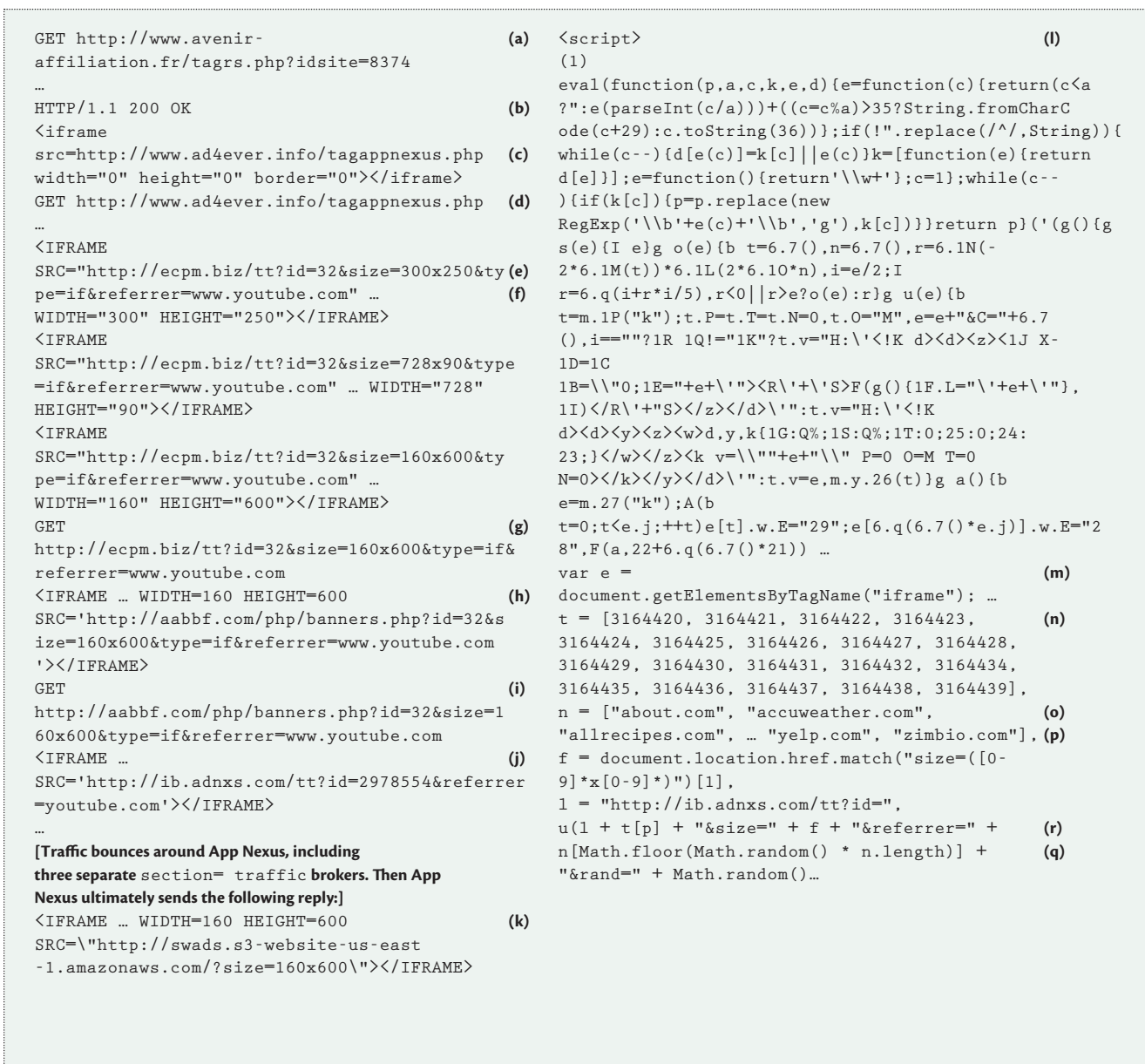


Figure 2. Misrepresenting ad placement location. (a) Request from user’s browser to remote server. (b) Start of response from server. (c) Creates invisible inline frame loading an Ad4ever page. (d) Request from user’s browser to the Ad4ever server. (e) Three ad windows created of three different sizes. (f) Falsely claiming ads are being loaded inside YouTube. (g) Request from user’s browser to the Ecpm server. (h) Forwards the user to the Aabbf server. (i) Request from user’s browser to the Aabbff server. (j) Forwards the user to the App Nexus exchange, a “network of networks.” (k) Forwards the user to a server called “swads,” hosted by Amazon Web Services. (l) Swads returns obfuscated code, intentionally encoded to make it difficult to understand. (m) Decoding and excerpting in relevant part. (n) List of App Nexus section IDs to receive this traffic, letting the perpetrator spread earnings across multiple accounts. (o) List of sites where perpetrator claims he or she is placing ads, except the placements are actually in the invisible frames described above. (p) And 86 more. (q) Random selection of section ID, supposed location. (r) A separate function (not shown here) creates the window that goes on to sell fake ad inventory.

yet care about something quite different (sales of its product). Other advertising channels have made partial progress in this regard.

Best known are pay-per-click advertisements, the text ads that have become ubiquitous at Google and beyond. In this case, an advertiser

only pays when a user clicks an ad, so users’ own actions offer a first defense against wasteful advertising. Ideally, if an ad reaches an unsuitable

consumer, the consumer won't click and the advertiser's expense will be zero. When users' search terms reveal what they're looking to buy, advertisers enjoy an interested and receptive audience. How better to sell a new Hoover than to consumers searching for vacuum cleaners?

If only it were that simple. The first complication was click fraud. Search engines pay myriad "syndicator" partners to distribute ads. Buy a search ad from Google and it will appear in searches on the New York Times, AOL, and hundreds of little-known sites. Google pays those sites a revenue share (often as much as 70%) for each click they provide. New York Times staff aren't likely to click the company's own ads, but this can be a problem for small publishers. Early click fraud detection systems checked for obvious signs like the same IP address clicking repeatedly, or a given publisher inexplicably serving (supposed) users with a particular type of web browser. (A more likely explanation: the publisher's clickbots all simulated that browser, a typical shortcut for crude bots.) But sophisticated fraudsters can easily circumvent those defenses. For example, with botnet-style control of thousands of users' computers around the globe, a fraudster can send clicks from a range of computers with minimal effort.

To catch click fraud and optimize their campaigns more generally, advertisers usually track "conversions" in which clickers actually make purchases. If one keyword yields purchases at a much higher rate than another, an advertiser is likely to increase investment in the former and reduce the latter. The same goes for differences across search engines and their various configuration options. This is the bedrock of modern search advertising and the guiding principle for tens of billions of dollars of search advertising, yet it can be strikingly

inapt. Many advertisers struggle to link online advertising to users' purchases. Anyone selling in retail stores (be it candy bars, dog food, or automobiles) will struggle to connect online ad clicks to users' purchases. Similar problems arise for sellers with long sales cycles and distributor networks.

One might imagine tracking a user's physical location to assess ad performance. If a user sees a car ad one day, then visits a suitable dealer the next, the ad may have had an effect. But tracking users day in and day out yields invasive records of routine activities, which can alarm many users. Even connecting desktop behavior to mobile walking isn't always easy; Google can do it thanks to powerful positions on both devices, but regular sites often struggle. And this approach is no help to advertisers whose products are sold widely. You may have seen an ad for a candy bar and then visited a grocery store later that week. Did you buy the product? Even Google doesn't know.

Most serious is the assumption that the purchase wouldn't have occurred were it not for the advertising. Consider a user who searches for Dell on Google and clicks on the first ad, invariably promoting Dell. Dell is likely to divide the cost of that click by the likelihood of the user making a purchase to get the expected cost per purchase. Then Dell will compare that cost to the gross margin and, in all likelihood, find the ad quite cost-effective. But the fact is that a user who searches for Dell has some chance of making a purchase from that company even if no ad appears. Dell's analysis mistakenly assumes that the probability is zero, sharply overstating the effectiveness of the ad. The state-of-the-art is random experimentation: showing ads to some consumers but not others and measuring purchases in the two groups. In a 2013 paper, Steve Tadelis and colleagues

ran an experiment on eBay's ad campaign that demonstrates this mistake. They found that brand-name advertising at Google actually has negative returns (the benefits are worth less than the costs), although a naive analysis would have found it highly profitable.¹

Companies may also buy ads for their own trademarks in fear of competitors poaching their customers. In trademark lawsuits, results are mixed: Google wins most cases that go to trial, though plaintiffs who settle often get what they want. Perhaps Dell is willing to pay to prevent an HP ad from appearing when a user searches for Dell—but existing measurement systems invite Dell to conclude, mistakenly, that this keyword is sending incremental customers.

Dissatisfied with search ads, some advertisers turn to other types of performance-based advertising. A common choice is affiliate marketing, in which advertisers pay only when a user makes a purchase. (The payment is usually a predefined percentage of the user's purchase, chosen to be below the advertiser's average marketing expense.) At first glance, this may seem risk-free to advertisers, and affiliate networks often tout it as such. But if a fraudster can find customers who are likely to buy from a given merchant, the fraudster can claim commission on purchases that the user would have made anyway. Fraudsters often claim commission by loading invisible ads (cookie-stuffing), placing adware on a user's computer (to see where the user is going and then intercede to claim to have referred the user there), and typosquatting (anticipating a user's misspelling of merchants' domain names and then invoking the affiliate links to send the user onward). My forthcoming article "Risk, Information, and Incentives in Affiliate Marketing" examines these schemes in greater detail.²

Limited Tools and Uncertain Counterfactuals

For anyone concerned about advertising effectiveness, one challenge is the limitation of available tools. It's typically straightforward for an advertiser to check how much was spent on advertising and how many sales were attributed to those ads. This data is often neatly organized in an automatically updated report—one row per search term, website showing display ads, or affiliate. But the sales data is less clear-cut. Are these sales that would have happened anyway, even without the advertising expenditure? The report invites advertisers to assume that every sale happened only thanks to advertising. Tadelis's result shows how wrong this can be.¹

What would a better tool look like? Ad platforms widely encourage advertisers to experiment with alternative versions of an ad to test the impact of changing color, words, or layout. But standard tools never invite advertisers to test the lack of an ad. The smarter test for advertisers concerned about the incrementality of their spending would be to run a control ad—showing a fraction of users an offer for UNICEF, for example—to see how many people buy from the advertiser anyway. The answer won't be zero.

See No Evil, Speak No Evil

It's tempting to imagine an advertiser's managers, staff, and contractors diligently working for the advertiser's genuine benefit, but the reality is less clear-cut. Performance is typically measured through automated reports—the same systems that often fail to evaluate whether traffic is truly incremental. If an ad network or agency is tasked with delivering 2,000 orders, its success will be evaluated based on what the measurement system reports. Savvy vendors will quickly realize that it's much easier to claim credit on orders that would have happened

anyway, rather than hustle to find new customers.

My research reveals the risk of moral hazard among marketing managers.² Consider the varying incentives of in-house program managers—who usually have greater reason to support company objectives including long-term relationships, bonuses, and physical and social proximity—versus outsourced staff whose performance is judged more narrowly, primarily based on reported metrics. In my forthcoming article, I discuss my finding that in-house staff members are considerably more effective at protecting their employers from “gray area” malfeasance.² Unfortunately this benefit comes at a cost, as in-house staff members have less information about clear-cut violations than outsourced specialists who have access to greater amounts of data collected across multiple advertisers. A sophisticated advertiser must navigate this tradeoff; for example, let outsourcers handle the clear violations but retain the right to make decisions on more difficult cases.

Law to the Rescue?

In an era of cybercrime, the legal system's limitations have become increasingly apparent. Jurisdictional boundaries often dilute incentives—victims tend to be American advertisers, while perpetrators can be anywhere. Voters in, say, Russia are understandably hesitant to allocate government resources to catching perpetrators whose victims are largely American. Even when perpetrators are identifiable and local, law enforcement is often hesitant to go after them. United States enforcers like the Federal Trade Commission and Department of Justice often prefer to focus on schemes where victims are individuals rather than companies. In business-to-business fraud, law enforcement expects the company, not taxpayers, to cover the costs.

This leaves advertisers to fend for themselves, but it's usually not realistic to expect them to do so. Suppose Expedia realizes a fraudster has stolen \$10,000. Between attorney time and diversion of management attention, Expedia would probably incur ten times that cost in pursuing the loss. It's easier just to ignore the problem. Yet this leads to a tragedy of the commons. If a fraudster steals \$10,000 from each of a hundred companies, the total loss is large—yet none of the companies is likely to pursue the matter. In the analogous context of consumer claims, class actions can solve a portion of the problem, but company class actions are less common and typically face procedural challenges.


Companies generally hesitate to pursue malfeasance in online advertising. Often, staff members regret allowing the problem to occur in the first place—a “blame the victim” analysis that, to my eye, severely misunderstands the nature of fraud. (Successful fraud always has a victim.) Other companies perceive that there will be embarrassment, cost, difficulty, or distraction in pursuing the problems. There can be countervailing benefits to these concerns. After eBay brought suit against two rogue affiliates who jointly took more than \$20 million of commission on nonincremental sales, many would-be fraudsters on “black hat” discussion boards elected to forego future attacks on eBay. Yet eBay's experience is the rare exception: the company's loss was so large that it was compelling and cost-effective to take action. In addition, eBay is a sophisticated repeat actor with longstanding ties to law enforcement. Most companies are not as well-positioned.

A final challenge comes from the long chains of intermediaries—there were seven separate brokers in the full ad placement summarized in Figure 2, and more in other observations. If A is promised by B that

C had promised D that E would do some quantum of work, who should A blame when something ultimately goes wrong? Even a bit of shirking by each intermediary will yield a final product that's much worse than promised. Historically, high transaction costs encouraged companies to deal directly with advertisers; if each intermediary introduced a 20% overhead, companies had strong incentive to simplify their supply chains. But in online ad relationships, efficient "ad exchange" brokers can facilitate trades at a lower cost. Suddenly, long chains become normal and even expected, and accountability predictably wanes. Lawyers and contracts have yet to catch up.

Online advertising is probably the largest service that's sold, distributed, and delivered electronically. With no printed records, no in-person transactions, and a never-ending onslaught of brokers as well as buyers and sellers, perhaps it's no surprise that the field has become both complex and risky. The smartest advertisers are appropriately skeptical of all contract counterparts: questioning what a report truly means, which data and analyses support the figures, and how they can verify that the benefit was actually provided. When these questions are sufficiently answered, the efficiencies of online advertising can be staggering—yet inattention to these questions can bring losses that are equally severe. ■

Benjamin Edelman is an associate professor at Harvard Business School. Edelman received a PhD in economics from Harvard University and a JD from the Harvard Law School, and is admitted to the Massachusetts Bar. His research explores the public and private forces shaping Internet architecture and business opportunities, with a focus on online advertising. Contact him via www.benedelman.org.

 Selected CS articles and columns are also available for free at <http://ComputingNow.computer.org>.

References

1. T. Blake, C. Nosko, and S. Tadelis, "Consumer Heterogeneity and Paid Search Effectiveness: A Large Scale Field Experiment," to be published in *Econometrica*.
2. B. Edelman and W. Brandi, "Risk, Information, and Incentives in Affiliate Marketing," June 2014, to be published in *J. Marketing Research*.