

Standardization and the Effectiveness of Online Advertising

Avi Goldfarb and Catherine Tucker*

May 1, 2014

Abstract

The technological transformation and automation of digital content delivery has revolutionized the media industry. Increased reliance on automation has also led to requirements for standardization of content-delivery formats. This paper examines how the memorability of banner advertising changed with the introduction of new standards regularizing its format. Using data from randomized field tests, we find evidence that for most ads, ad effectiveness falls as the use of standard formats rises. The decline is smaller when a standardized ad appears to be more original (such as ads created by an ad agency). Therefore, a likely explanation is that increased use of a standard format makes it harder for basic ads to distinguish themselves from their competition because the ad format commands less attention.

Keywords: Online advertising, Standards, Marketing regulation

*Avi Goldfarb is Professor of Marketing, Rotman School of Management, University of Toronto, 105 St George St., Toronto, ON. Tel. 416-946-8604. Email: agoldfarb@rotman.utoronto.ca. Catherine Tucker is Associate Professor of Marketing, MIT Sloan School of Management, 100 Main St, E62-536, Cambridge, MA. Tel. 617-252-1499. Email: cetucker@mit.edu. We thank Michal Grajek, Timothy Derdenger, and participants at workshops at the University of Southern California, Carnegie Mellon University, and the University of Toronto for helpful comments. This research was partly supported by SSHRC grant numbers 486971 and 493140.

1 Introduction

Online ads are now delivered in under one second via real-time ad exchanges that match advertisers with webpages. This has led to substantial efficiency gains in the media industry. One side effect of the increasing reliance on electronically automated ad display and ad exchanges is that ad formats need to be standardized to fit a predetermined webpage space. Ad format standardization resolves problems of coordination between publishers and advertisers. Reducing these coordination costs is particularly important in media because of the prevalence of two-sided markets with competing platforms and potential multi-homing (Armstrong, 2006). This standardization benefits both the content producer, who can more easily fit content around advertising, and the advertiser, who can use the same creative ad design more easily across multiple advertising channels.

Standards can help firms through reduced coordination costs, but they can hurt firms by making it harder to differentiate (Shapiro and Varian, 1999; Augereau et al., 2006). In the advertising context, the inability to differentiate might manifest itself through reduced attention. As consumers adapt to a particular ad format, they pay less attention (Solomon, 1999; Pashler, 1998).

In this paper, we examine how the widespread adoption of a standard format (led by the online advertising industry association, the Interactive Advertising Bureau) influenced the effectiveness of online advertising. In doing so, we document the challenges of advertising when a cost-saving technology (that demands standard ad design) limits the ability of an ad format to grab attention. We also explore remedies to these challenges.

We use data from a large-scale database of ‘a/b’ real-time field tests of online display advertising collected by a media agency on behalf of advertisers that allows us to measure how much an ad was able to grab consumers’ attention. This data bank is one of the primary data sets used by the industry to benchmark online display advertising. In each test, people

were randomly exposed either to a focal ad (the treatment group) or a placebo ad (the control group). On leaving the website, both groups were asked whether they could pick the focal ad out of a random selection of ads as one they had recently seen. Because of the experimental design, the difference between the treatment and control groups in recognition of the focal ad can be seen as the causal impact of the ad on ad recognition. All results are robust to a measure of purchase intention; we focus on recognition because the literature on advertising and attention explicitly links attention to memory rather than to purchasing (Finn, 1988; Dreze and Hussherr, 2003).

We examine US-based online display advertising campaigns from August 2002 to August 2004, totaling 381,641 survey responses to 1,064 different advertising campaigns. We study how the difference in ad recognition between the treatment and control groups changed with the April 2003 introduction of standard formats.

We find evidence that, on average, standardization damaged ad recognition and stated purchase intention. Using a basic model specification, standardization decreased recognition by over 20 percent for standard-format ads. Before this formal standardization, ads that self-selected into what became the standard formats were no less effective than other ads.

The identifying assumption for this result is that there are no reasons why standard format ads should change in effectiveness relative to non-standard format ads other than because of the standards-setting process after the introduction of the standard. To explore the validity of this assumption we use three complementary approaches. First, we demonstrate that there was not a pre-existing trend in decreasing effectiveness of standard format ads. Second, we conduct a falsification check where we show that the effect does not occur if we only look at webpages with a single ad. This suggests that indeed it is the loss of attention, attributable to multiple ads looking more similar due to a standardized format, which is the mechanism for the effect. Third, we use instrumental variables to measure plausibly exogenous variation in the advertiser's decision to pursue standard format that stems

from the extent to which the majority of the impressions would be shown in North America (where the standard was based), or elsewhere. These instrumental variable estimates are qualitatively and statistically consistent with those where we did not control explicitly for selection (though the point estimates are higher in magnitude).

We then explore how advertisers might be able to overcome this negative effect of standardization. We show that more original ads are not affected by standardization. Specifically, ads that were designed by ad agencies and ads with content other than a simple logo have no significant change in effectiveness after standards are implemented. Therefore, our results suggest that the rewards to creativity increase in other dimensions if ad formats are standardized. This is consistent with well-established laboratory-based research on habituation and novelty (Pieters et al., 2002; Pashler, 1998).

While our paper does touch on issues highlighted in the large economics literature on standards, such as the tension between differentiation and standardization (e.g. Katz and Shapiro (1985), Farrell and Saloner (1988), Augereau et al. (2006), Simcoe (2012)), it is perhaps most directly related to the emerging literature on understanding the effectiveness of online advertising. For example, Manchanda et al. (2006); Goldfarb and Tucker (2011b) and Lambrecht and Tucker (2013) measure the effectiveness of advertising tactics in banner ads. More recently, Sun and Zhu (2013) examine the relationship between website content and ad revenue. The present paper is novel because it examines the role of attention, the challenges advertisers face in gaining attention when cost incentives suggest standardization, and the usefulness of creative ad copy in overcoming these challenges.¹

Overall, our results suggest that the increased use of standard format advertising has a cost in terms of reduced ad recognition and stated purchase intent (likely due to reduced attention as ad formats become less distinct from each other), particularly for ads that

¹This may be particularly important for environments where ads are very standardized such as in search engines (Ghose and Yang, 2009).

are not created by ad agencies. Therefore, standards lead to reduced differentiation and effectiveness, but this effect is muted when the creative content of the advertising is increased. The results do not imply that the ad format standards were bad; their widespread adoption hints otherwise. They simply suggest that in addition to any benefits related to reduced coordination costs, the standards (asymmetrically) reduced the effectiveness of advertising.

2 Voluntary Online Advertising Standards

The standards-setting organization for online advertising in the United States is the Interactive Advertising Bureau (IAB). Founded in 1996, the IAB consists of 460 leading media and technology companies who are responsible for selling 86% of online advertising in the United States. Working with its member companies, the IAB evaluates and recommends standards and practices and conducts research on interactive advertising.

In 1996, the IAB issued its first set of guidelines for ad formats. In January 2001, among increasing concerns about the costs of effective online advertising and the perceived need for a more professional organization, the IAB hired its first CEO. A perceived problem in the industry at that time was that the existing (1996) guidelines were too broad and that “online publishers...had a tendency to go their own way [that is, introduce new ad formats] in trying to attract advertisers” (Taylor, 2001). To address this issue, in August 2002, the IAB formed the ‘Ad Sizes Task Force.’ This was created to reduce the number of ad sizes, so as to reduce costs and inefficiencies associated with planning, buying, and creating online media. In December 2002, the task force announced that they would create a universal ad package. This would consist of a set of four ad sizes that all compliant member publishers had agreed to support.

As detailed in Table 2 these standardized ad units included a Rectangle, a Medium Rectangle, a Wide Skyscraper and a Super Banner. All designs had to have a 40k initial download fileweight, and a 15 second maximum limit on any animation. Appendix Figure

A-1 presents a mock-up of how these ad formats appear on a webpage. These formats were a subset of the wide variety of formats previously suggested in IAB guidelines. After the introduction of the Universal Ad Package (UAP), compliance meant that publishers accepted these four sizes and consequently enabled the advertisers to reach their audience using standardized ads.

The intention was to establish a real standard for the industry. Jeff Bernstein, Director of MSN Ad Planning and Chairman of the IAB Ad Sizes Committee, described the aim of the initiative as follows: “This initiative....is intended to answer advertisers’ requests for a limited, core set of compelling ad units to create and plan online campaigns that will be able to run across the majority of Web sites and users. It will also enable publishers, regardless of size or niche, a common palette with which to attract advertisers and agencies, providing the framework for integrated campaigns across the Internet.” The new standards were officially launched on April 28, 2003.

As new advertising technologies arise, the IAB regularly sets new standards. For example, standards for rich media were changed substantially in 2009. We focus on the April 2003 standards for two reasons. First, they are the first set of standards to be widely adopted and therefore they provide the starkest contrast for studying before and after the introduction of standards. Second, our data contain mostly plain banner ads. The April 2003 standards are therefore the set of standards best applied to our data.

The American Association of Advertising Agencies (AAAA) agreed to support this set of ad formats and they were received broadly in the industry too. For example, Jonathan Adams, a Senior Partner at mOne Worldwide, a subsidiary of Ogilvy, said ‘The UAP affords a simpler approach to interactive media planning without impacting flexibility for advertisers to execute compelling, unique online advertising campaigns.’

Figure 1 shows the change in the proportion of display ads in the US that used these standardized units for campaigns before and after the official launch of the standards on

April 28, 2003. It indicates that there was a clear kink in the rate of standardization in April 2003, providing evidence that the standardization process was successful, and that gradually more and more publishers and advertisers used these units.

Figure 1 highlights two other points. First, as discussed by Farrell and Simcoe (2012), adjustment is not immediate after standards are established. After the announcement, both creative agencies and internet platforms needed time to adjust the design of ads and websites to accommodate the new standards. Second, even prior to the December announcement of the specific standards, the standard formats were relatively popular and growing in their use in campaigns, perhaps because these standards were chosen due to their perceived advantages by the industry.

In our empirical analysis, the difference-in-difference estimation focuses on the change in effectiveness before and after the kink. Similar to Chen et al. (2011) and Sun and Zhu (2013) we take the precise timing of the change in standards as exogenous and our source of exogenous variation. To check that our results are not simply a result of selection—that low-quality ads converted into the standard format—we also explore plausibly exogenous adoption of standard formats driven by differences in the inherent international focus of a product via instrumental variables.

3 Data on Display Advertising

We use data from a large database of field tests conducted in the United States by a media metrics agency. The aim of this database is to provide comparative guidance to advertisers about which types of ad design are effective and to benchmark current campaigns against past campaigns. Each of the 1,064 campaigns that we study follows the same methodology to evaluate effectiveness. Therefore, a key strength of this data source is that it allows comparison across many campaigns for different types of ads and different types of advertisers. This database is among the main sources used by the online media industry to benchmark



Figure 1: Standardization Over Time

online display ad effectiveness.

Specifically, for each campaign the advertiser hired the media metrics agency to assess and benchmark the effectiveness of an ongoing campaign that uses banner ads whose impact depends merely on whether a consumer sees them rather than just clickthroughs. The ability to compare is driven by the use of ‘ad recognition’ and ‘purchase intent’ as measures of ad effectiveness. These measures are weaker than data on actual purchasing, as used by Lewis and Reiley (2009). In some sense, our measures therefore trade precision on measuring effectiveness for the ability to compare across many campaigns. At the same time, to the extent that we are focused on the tradeoff between standard format and attention, the focus on the memory-based measure of ad recognition is more directly relevant to the hypothesized mechanism.

The agency integrates its services into an ongoing campaign, and randomly shows the ad for the focal product to some individuals and an ad for another product, typically a non-profit, to other individuals. It then immediately surveys these individuals (359 per campaign on average) when they leave the website on whether they saw the treatment or

control ad. The agency gathers measures of ad recognition and purchase intent for the focal product. Since these individuals are randomly assigned to the treatment and control groups, any differences in their ability to recognize ads for the focal product can be ascribed to the exposure to the focal product’s advertisement.

The field test collected data by means of an online questionnaire. This questionnaire appears as a pop-up window. Approximately half of the people who see the pop-up window will have seen the focal ad and half will have seen the dummy ad. A small fraction of users who see the pop-up window take the survey. The media metrics agency does not provide response numbers, but it seems likely that the percentage of users that takes the survey is very low. Given the random allocation of the focal and dummy ads, our results rely on the standard randomization assumption that there is no difference in the propensity to take a survey for those who saw different types of banner ads.

There is the further issue of the representativeness of the measured ‘treatment’ effect, given that the subjects may be unusual because they were willing to answer an online survey. We view it as comforting that the demographic variables reported in Table 1 appear representative of the general internet population at the time of the study as documented in the Computer and Internet Use Supplement to the 2003 Current Population Survey. Still, various forms of selection bias are possible. For example, it may be that those who are willing to answer the survey are perhaps more observant than other web users: They did notice and respond to the pop-up window. Therefore, given that the allocation to treatment and control groups is random, an accurate but cautious interpretation of our results is in terms of how standardization affects a widely-used industry measure of how well advertising performs, rather than necessarily reflecting the responses of all consumers. A less cautious interpretation assumes that the measured qualitative difference between the treatment and the control groups is not affected by the consumers who selected into the survey.

Upon agreeing to take the survey, users responded to a small number of questions about

ad recognition, purchase intention, and demographics. Given that the survey window appears shortly after the user sees the focal or dummy ad, the questions measure the immediate effect of seeing the ad.

Specifically, to measure ad recognition, the survey displayed a selection of four to six ads, one of which was the ad for the focal product and none of which was the placebo ad.² Respondents were asked to click on any ads they had ‘recently’ seen. This measure is similar to measures of print ad recognition that have a long history in marketing scholarship (Starch, 1923; Shepard, 1942; Wells, 1964; Finn, 1988; Kent and Machleit, 1990) and it is especially similar to the measure of aided ad recall used in Dreze and Hussherr (2003) . Specifically, Dreze and Hussherr’s measure is, ‘For each of the following banner ads, do you recall seeing it during the task?’, which they describe in their Appendix B as a ‘Traditional memory-based effectiveness measure’.

Ad recognition is often used as a measure of attention (Finn, 1988; Aribarg et al., 2010). Using eye-tracking technology, Aribarg et al. (2010) examine three different measures of recognition: (1) an ‘ad noted’ measure of ‘have you seen this ad?’, similar to the measure employed in our study, (2) a measure of ‘Have you seen an ad with this brand?’, and (3) a measure of ‘Have you read more than half the text of this advertisement?’. They find (p. 397) that ‘attention to the ad predicted the ad noted measure’, but not the others. Given that the measure used in this study is similar to the one measure that they found did predict attention, and that our research design has randomized exposure (which overcomes some of their concerns about the noisiness of the measure), we view their results as support for our emphasis on the simple recognition measure as a proxy for attention to the ad. At the same time, their results suggest an important caveat: our ad recognition measure may be a good proxy for attention to the ad, but it is unlikely to be a good proxy for attention to the

²In our prior work, we referred to this measure with the data provider’s label of ‘ad recall’ . We thank an anonymous referee for pointing out that it is more like the ad recognition measures used in past literature.

content of the ad.

Thus, in the main specification in this paper, our dependent variable is whether or not the respondent was able to pick the focal product ad out of a random selection of ads. Our results are robust to using stated purchase intent (using both a five-point scale and a discrete variable for whether the respondent answered ‘likely to purchase’ or ‘very likely to purchase’ – four or five on the five-point scale). Of the 381,641 total survey responses, all provided purchase intent information but 27,266 did not provide ad recognition information.

We focus on the time span immediately surrounding the change in standards, specifically from August 2002 to August 2004. Prior to August 2002, our data become too sparse to estimate reliable effects. We end the ‘after’ period in August 2004 because the relatively short window reduces the potential of other changes in the industry to wash out our main results, such as a gradual decline in advertising effectiveness over time and the arrival of new formats. Our results are also robust to using a six-month window on either side of the April 2003 announcement.

The mean ad campaign lasted 32 days and consisted of a uniform set of ads. There were 167 separate products advertised in total on 30 different categories of websites. Products include diapers, television programs, shampoo, airlines, toys, and wireless carrier services. Website categories include personal finance websites, news websites, entertainment websites, and portals. This means that our estimates reflect the placement of ads in their natural settings. Consistent with industry norms and our prior work (Goldfarb and Tucker, 2011c), we define a ‘campaign’ as an ad shown for a specific product on a specific website.

If a respondent was in the exposed condition and returned to that particular webpage, or refreshed that webpage before exiting the website, the respondent is counted as having seen the ad again. The median exposure was to have seen the ad one time (56 percent of respondents who were in the exposed condition).

The survey also asked respondents about their gender, income, and age. We use these

variables as controls in our regressions, after converting the income and age responses to zero-mean-standardized measures. We assigned a value of zero to individuals who did not provide income or age data. We do not view the missing data on the demographic controls to be a concern, because the results are robust to a non-parametric specification of the controls that adds missing data fixed effects and to the omission of these controls entirely. Table 1 shows summary statistics for this survey data.

Table 1: Summary Statistics for Full Sample

	Mean	Std Dev	Min	Max	Observations
Ad Recall	0.29	0.45	0	1	354375
Purchase Intent	0.32	0.47	0	1	381641
Likely Scale	2.93	1.35	1	5	381641
Exposed	0.52	0.50	0	1	381641
Female	0.46	0.50	0	1	381641
Income	59520.5	49354.6	15000	250000	280943
Age	38.3	13.8	13	100	381476
Standard Format	0.47	0.50	0	1	381641
Ad Agency	0.15	0.36	0	1	381641
Copy in Ad	0.19	0.40	0	1	381641
Proportion US	0.95	0.15	0	1	381641
Observations	381641				

There were many different creative formats used for these banner ads. The database was partly designed to help guide advertisers in their creative decisions, so the format information is very detailed. For each ad, we know the precise size and various formatting decisions. We use these to determine whether or not the ad was part of the standardized ad package developed by the IAB.

As shown by Table 2 all four of the standard-format ads exhibited a sharp increase in usage after the introduction of the standard. In addition, previously popular formats (such as Banner (468x60) that were not part of the standard format ad package exhibited a decline in use from 28% to 12%. Similarly, another previously popular ad format, the Skyscraper format (120x600) declined from 19% to 10% usage.

Table 2: Standardization increased the use of the Standardized Ad Units

	Before Announcement Percentage	After Announcement Percentage
Super Banner (728x90)	4.66	21.63
Rectangle (180x150)	0.05	3.31
Medium Rectangle (300x250)	11.37	20.77
Wide Skyscraper (160x600)	5.00	16.57

Table 3: Differences in Differences: Standardized Ads

	Difference	Mean Control	Mean Exposed	T-Test
Standardized Ads				
Ad Recognition Before Standardization	-0.000	0.346	0.346	-0.063
Ad Recognition After Standardization	-0.084	0.239	0.323	-36.447
Non-Standardized Ads				
Ad Recognition Before Standardization	-0.106	0.211	0.317	-32.492
Ad Recognition After Standardization	-0.122	0.238	0.360	-44.708

Table 3 provides some initial model-free evidence on how standardization affected ad recognition. In particular, Table 3 suggests that for standardized ads there was approximately a 20 percent decline in relative recognition between the exposed and control group after the policy change. Table 3 also suggests that non-standardized ads appear to have performed somewhat better in absolute terms after the standardization process; however, this latter result is not robust to the addition of campaign-level controls. The fact that controlling for observable campaign characteristics changes this raw difference-in-difference result is suggestive that the assignment to standard formats is not random - something we address in detail when we introduce instrumental variables for selection into the standard format.

4 Empirical Analysis

Next, we document the decrease in the effectiveness of standard-format ads after standardization in an econometric framework. First, we use a difference-in-difference specification to show that the standard-format ads became less effective after April 28, 2003, relative to before that date and relative to other ads. Then, we show robustness to alternative specifications and validate the results with a falsification check that shows that there is no significant impact on ad effectiveness when we focus on situations in which the ad is the only one on the page and therefore is likely to get attention. Next, to address the endogeneity of ad selection into the standard format after standardization, we instrument for standard format using the proportion of the exposures of the campaign that were intended to be targeted outside North America where the same advantages of embracing the standard would not apply.

4.1 Standard Formats and Ad Recognition

In our empirical analysis, we use a straightforward specification to capture how recognition is affected by the type of ad format. For person i who was exposed to advertising campaign j at time t , their ad recognition reflects

$$\begin{aligned} \text{Recognition}_{ijt} = & \alpha \text{Exposed}_{ij} + \beta_1 \text{Exposed}_{ij} \times \text{StandardFormat}_j \times \text{AfterStandardization}_t + (1) \\ & \beta_2 \text{Exposed}_{ij} \times \text{StandardFormat}_j + \beta_3 \text{Exposed}_{ij} \times \text{AfterStandardization}_t + \\ & \theta X_{ij} + \gamma_j + \epsilon_{ijt} \end{aligned}$$

Therefore, α captures the main effect of being exposed to an ad on recognition. β_1 captures the core coefficient of interest for the paper - whether exposure is more or less influential for ads that used one of the Standard Formats listed in Table 2 after they were recognized as the standards. β_2 and β_3 respectively control for whether the standardized ad

format is less effective even prior to the change and whether ads were generally less effective after the change. X_{ij} is a vector of controls for gender, age, income, and time online. γ_j is a series of campaign fixed effects that control for heterogeneity in baseline recognition and includes the main effect of whether or not the focal ad was standardized (*StandardFormat_j*), which is why this lower-order interaction is not included in our specification. For convenience below, we will refer to ad ‘effectiveness’ as the impact of ad exposure on ad recognition.

This is estimated with a linear probability model. We focus on the linear probability model because it allows us to estimate a model with many campaign fixed effects as these fixed effects get differenced out. In contrast, computational challenges and the incidental parameters problem limit the fixed effects we can use in a nonlinear model. We are less concerned about the potential bias of the linear probability model discussed in Horrace and Oaxaca (2006) because the predicted probabilities all lie between 0 and 1. This is likely because the mass point of the Ad Recognition variable is far from 0 or 1 and the covariates are mainly binary.³

Table 4 shows the results. Columns (1) to (2) build to the main specification for equation (1) in Column (3). Specifically, Column (1) shows the raw difference, between the exposed and control groups in ad recognition, of about 10 percentage points. This value does not control for campaign effects, category effects, or respondent demographics. The R-squared value in this column is just 0.0127. This is unsurprising given that the regression examines the effect of seeing just one online display ad once on overall recognition of advertising across individuals and campaigns. In prior work (Goldfarb and Tucker, 2011b,a), we showed that the measured ad effectiveness in this data is appropriate in light of the relatively low price of online display advertising.

Column (2) shows that standard-format ads are, on average, less effective than other ads. Column (3) shows the main specification, with campaign-level fixed effects. It shows that

³Earlier versions of the paper showed the robustness of the main specification to a logit model.

standard-format ads, when conducted after standardization occurred, are less effective than other ads. This is not true of standard-format ads prior to standardization and it is a much stronger effect than the small general reduction in the effectiveness of non-standard-format ads. As expected from a randomized field test, this specification is similar to one which includes simply category-level controls and excludes demographic controls.⁴

Table 4: Standard Format ads became less effective after standardization.

	Main (1) Ad Recall	(2) Ad Recall	(3) Ad Recall	6 Month (4) Ad Recall	Alt. Dep Var (5) Likely Scale	Falsification (6) Ad Recall	IV (7) Ad Recall
Exposed × Standard Format × After Standardization			-0.0225*** (0.00695)	-0.0284*** (0.00788)	-0.0555*** (0.0187)	-0.00951 (0.0392)	-0.0788* (0.0417)
Exposed × Standard Format		-0.0296*** (0.00302)	-0.00303 (0.00622)	-0.00312 (0.00625)	0.0136 (0.0169)	0.00554 (0.0385)	-0.0180*** (0.00399)
Exposed	0.102*** (0.00151)	0.116*** (0.00210)	0.107*** (0.00304)	0.108*** (0.00315)	0.0588*** (0.00867)	0.102*** (0.00716)	0.110*** (0.00299)
Standard Format		0.0106*** (0.00219)					
Exposed × After Standardization			-0.00644* (0.00391)	0.00554 (0.00453)	0.00421 (0.0109)	0.0187** (0.00878)	0.0289* (0.0162)
Female			-0.0329*** (0.00180)	-0.0279*** (0.00241)	-0.00942* (0.00510)	-0.0210*** (0.00421)	-0.0314*** (0.00312)
Std. Income			0.00128 (0.000915)	0.00453*** (0.00121)	-0.0231*** (0.00269)	0.000934 (0.00199)	0.00392*** (0.00126)
Std. Age			-0.0144*** (0.000802)	-0.0166*** (0.00112)	-0.0708*** (0.00227)	-0.0195*** (0.00205)	-0.0175*** (0.00132)
Constant	0.232*** (0.00109)	0.227*** (0.00150)					
Campaign Fixed Effects	No	No	Yes	Yes	Yes	Yes	Yes
Observations	354375	354375	354375	189822	381641	65109	179098
R-Squared	0.0127	0.0130	0.134	0.124	0.188	0.169	0.0160

Ordinary Least Squared (OLS) regression coefficients shown in Columns (1)-(6). IV estimates shown in Column (7). Data are for August 2002 to August 2004 in Columns (1)-(3), (5)-(6). Data are for 6-month window in Column (4). Data are for August 2002 to October 2003 in Column (7) to ensure validity of the exclusion restriction. The dependent variable is an indicator variable for whether or not the person was able to correctly pick out the ad they had just seen in Columns (1)-(4) and (6)-(7). 'StandardFormat' is collinear with campaign fixed effects in columns (3)-(7) and omitted (see page 15 for details). Robust standard errors clustered at the campaign level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The remaining columns of Table 4 provide further checks for the robustness of these results to a different data selection criteria, a different dependent variable, a falsification check, and an instrumental variables specification. Column (4) shows robustness of the main specification to using a different time window of the six months before and after the standardization announcement. Column (5) shows that the results hold using a five-point scale for purchase intent rather than ad recognition.

Another concern is that the result is not linked with standardization, increasing clutter

⁴One potential concern is that the results are driven by changing consumer responsiveness to how large ads were over time, and that this was independent of standardization. To address this, in earlier versions we showed that the qualitative results are also robust if we include additional controls for ad size. This suggests that it was not changes in the importance of ad size but rather the changes in the importance of the ads being standardized that drive the results.

and consequently reducing attention, but instead with something that we do not capture about the creative of the new standardized ad in the post-period. To investigate this, we conducted a falsification test where we examined a context where there was no potential for ad clutter, as the ad was the only one displayed on the page. Column (6) of Table 4 explores sponsorship campaigns, in which an advertiser usually takes over all advertising at a website. As discussed by Dukes and GalOr (2003), this form of exclusivity is attractive since it reduces competition from other ads. The results suggest that for these particular ad campaigns there was no measurable negative effect from standardization.⁵

Overall, Columns (1)-(6) Table 4 suggest that standard-format ads became significantly less effective after April 28, 2003, relative to the change in effectiveness of other types of advertising. Our main specification in Column (3) of Table 4 suggests that the standard-format ads become 21% less effective.

The remaining identification issue that we face is that the decision to opt into the standard after the policy change may be endogenous. It is not clear how this biases our results as the bias could go both ways. On the one hand, firms may have chosen to embrace the standard for the campaigns which they had put the least effort in designing in order to automate the process and not have to expend extra effort on their distribution. If this is the case, then our estimates would overstate the extent to which the standard and the true estimate of the effect of standardization would be lower. On the other hand, it may be firms with the most online technical expertise and by association the best-designed online campaigns who adopted the standard. If this is the case, then our estimates would understate the true extent to which the standard negatively affects ad performance. To tease apart this potential bias,

⁵An additional concern in interpreting the results as driven by a causal relationship from the increased use of standard format ads to the decreased effectiveness of those ads is whether the standard format ads had begun to lose effectiveness prior to their increased usage from standardization. To examine this possibility, we also examined whether there was an increasing trend towards the standard format ads prior to the announcement of the formats. We found that including a dummy that placed standardization either six or nine months prior to standardization did not suggest a significant trend towards loss of effectiveness prior to standardization.

we turn to instrumental variables.

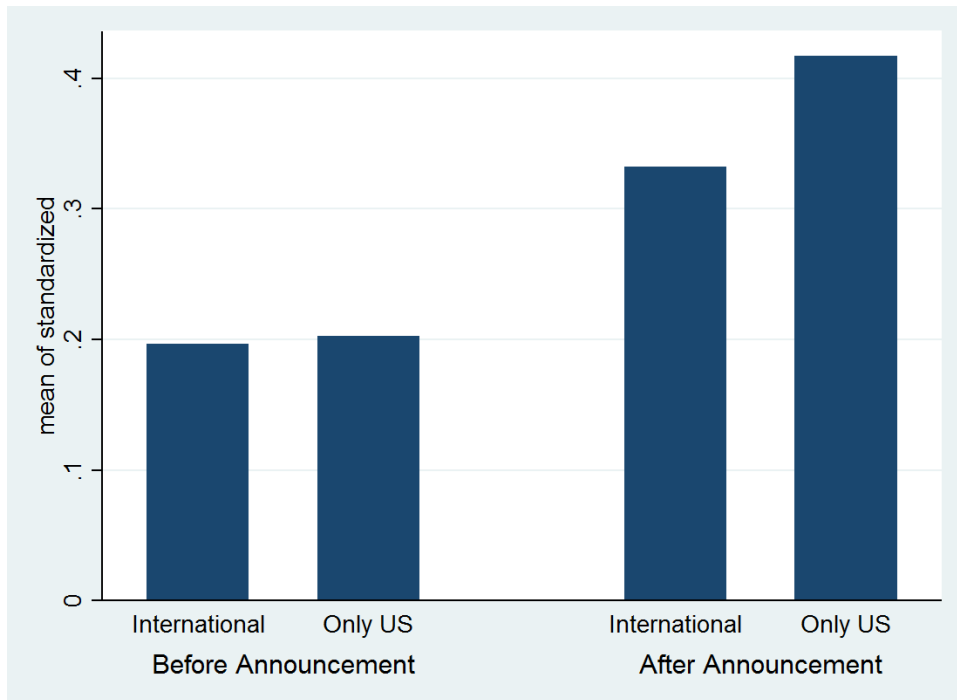


Figure 2: Correlation of Instrument with Endogenous variable

We need to identify an instrument which drives the decision of advertisers to adopt the standardized format or not, but does not directly affect how well online ads for that campaign perform. To do so, we identify an instrumental variable that is likely to affect the relative costs of adoption of the standard for each campaign.

Specifically, we use the proportion of international exposures relative to domestic (US and Canada) exposures as an instrument for the decision to embrace the standard after standardization. For each campaign, we define this as the fraction of international relative to domestic ads for that particular creative shown for the time span the creative was run. We believe this is a valid instrument because the standardized format was a North American format for the first seven months (it was adopted in Europe at the end of October 2003), so the costs of adopting it are correlated with the extent to which the firm is planning to use that creative in North America relative to other countries during those seven months.

For the instrumental variables analysis, therefore, we restrict our sample to end in October 2003.

The key assumption is that adopting the standard poses a fixed cost for the firm and they decide whether or not to adopt the standard partly based on how much of their online advertising they can spread the cost over. If a firm focuses on North America (for example they make fuel-inefficient SUVs whose main customer base is in North America), they will find it relatively less costly to adopt a standard, since the standard would be used for all their advertising, than would a firm that is selling a fuel-efficient compact vehicle in multiple international markets. The international firm potentially faces multiple different standards, and their cost of adopting the standard would only be spread among a small proportion of their overall advertising.

The idea that international advertising campaigns are more costly to coordinate is long-established in the marketing literature. For example, Peebles et al. (1978) discuss several steps they view as necessary to enable international campaigns to succeed. These steps are in addition to what is required for domestic campaigns, and ignoring differences in details of how the advertising industry works across countries can be costly. Takada et al. (2012) (p. 508) emphasize that media purchasing across countries is subject to location conditions, increasing the cost of international campaigns.

More direct support for our identifying assumption that it is costly to design for multiple ad sizes comes from the reasoning that European countries used for adopting the North American standards. Specifically, the UK Internet Advertising Bureau writes, “We carefully selected ad sizes that were already standard sizes in other markets...making international campaigns easier to run.”⁶ Similarly, on announcing the standards in October 2003, the European Interactive Advertising Association press release stated the standards are to “fa-

⁶<http://www.iabuk.net/resources/standards-and-guidelines/iab-display-ad-standards-2010>. Accessed April 11, 2014.

cilitate the growth of truly pan-European ad campaigns” by reducing differences in formats across countries.⁷

Figure 2 displays model-free evidence that shows how the share of domestic ad exposures is positively correlated with the decision to adopt the standardized formats after standardization. The underlying argument regarding the exclusion restriction is that we would not expect the relative share of international advertising to affect the lift of an ad relative to a placebo in the US, except through the extent that it led the firm to embrace US ad standards.

Column (7) of Table 4 presents our instrumental variable results.⁸ The magnitude of the estimate is larger and less precisely estimated, perhaps suggesting that the OLS estimates underestimated the extent of the effect due to selection. A possible explanation is that it was the most digitally able and best connected firms and ad agencies who adopted the standard, and since their ads were high-performers naturally, we underestimate the extent of the negative effect of standardization. However, we also note that the instrumental variables results are not significantly different from the OLS results as measured by a Hausman test. Therefore any discussion of differences is necessarily speculative, based on point estimates that are not significantly different.

The first-stage results indicate that there is (as expected given Figure 2) a strong correlation between the tendency to embrace the standardized online format after it becomes the standard, and the proportion of the campaign which was directed to the domestic market. Indeed, the coefficient of 0.09 for the instrument in the first stage is significant and has a standard error of .003 and a t -statistic of 28.8. Consistently with this, the first stage F-test suggests that the instruments have power.

The next question for instrumental variable validity is whether or not the instrument satisfies the exclusion restriction - that is, that the only way the ratio of international to

⁷<http://www.iab.it/docs/news/nuovi-formati-standard-advertising-1.pdf>. Accessed April 11, 2014.

⁸The technical details of how we implemented the instrumental variables in our three-way interactions setting are presented in the Appendix.

national coverage of the campaign affects banner ad effectiveness is through its effect on the likelihood of the ad being standardized or not. The key assumption is that an increase in international coverage for the ad campaign does not directly affect the effectiveness of the online banner ads that we study, through any other mechanism than the likelihood of that campaign having embraced a standard or not.

It is not possible to test that an instrumental variable meets the exclusion restriction (Angrist and Pischke, 2009). Instead, all that is possible to do is to look for evidence that the exclusion restriction fails.

One issue is that could potentially undermine the exclusion restriction, is that the kind of ads that had a larger international presence were improving in other dimensions that increased their appeal relative to ads with a large domestic presence. We do not have many measures of quality that are not directly related to size (and the standard format). Perhaps the best measure we have is whether the advertiser used an ad agency for the campaign. Therefore, we examine whether international brands were hiring better ad agencies who were devising better copy in order to explore the face validity of our exclusion restriction. Importantly, we find very little difference in mean usage, or change in mean usage, of advertising agencies across internationally and domestically focused campaigns before and after the implementation of the standard format ad (less than .7 of a percentage point).

Similarly, we checked that there was no increasing trend in baseline awareness of the products being advertised across internationally focused campaigns and domestically focused campaigns. If there was increased awareness, for example for domestic brands, then there is the potential that advertising would be less effective at increasing awareness inherently in such a mature market. However, we found no statistically significant difference in average levels, or change in average levels, of baseline product awareness for internationally focused and domestically focused campaigns.

Another separate concern is that companies might have selected higher quality campaigns

to run internationally. Our identification argument is based on the idea that it is inherent product characteristics rather than campaign characteristics which drive the decision to run campaigns internationally. In other words, the argument assumes that firms do not choose to make better campaigns international and off-standard.

We explore the validity of this assumption in three ways. First, we note that the proportion of international campaigns does not fall after standardization. Second, we use the same argument as above: there is no evidence of a change in quality for international campaigns on observable characteristics. Third, and most directly, we checked that it was indeed product characteristics which are driving the decision to run international campaigns. In particular, we reran our estimation using the same instrument, that is the proportion of domestic impressions, but calculated on the product-level (for example the car model type) rather than the campaign level. The product-level and campaign-level instruments are highly correlated (.92) suggesting indeed that it is the product characteristics which are driving the decision for whether the campaign is international. Given this high correlation, it is not surprising that the results were similar.

Though this is encouraging evidence that there were not obvious factors that could explain why internationally focused campaigns increased relatively in effectiveness to domestic campaigns after the introduction of standards, we recognize that it is impossible to prove the validity of an instrument. Our identification argument depends on the extent to which there is a lack of change in any direct interaction between the international span of an ad campaign and the inherent appeal of the campaign, after the introduction of standards.

5 Overcoming the Challenges of Standard Formats and Habituation

Next, we explore how advertisers can overcome the challenges associated with getting user attention on a standard format ad once users become habituated to the ad format. To

do so, we draw on the psychology of attention literature to argue that originality increases attention (Pieters et al., 2002). In light of this, we expect that original ad content will help overcome the reduced effectiveness of standard format ads after standardization. These results demonstrate the role of attention in overcoming the change in standardized ad effectiveness after April 2003, so we argue that they help to validate a causal interpretation that the introduction of standards drove the observed reduction in the effectiveness of standardized ads. Perhaps more importantly, they provide advertisers with a strategy to overcome the challenges of habituation.

In Table 5, we examine whether more original ads are hurt less by standardization. We identify more original ads in two ways. First, our data contain information on whether the ads were designed by an advertising agency or whether they were done in-house. We argue that advertising agency ads are likely to be more creative and original, given that the business of ad agencies is, at some level, to create more distinctive ads than could be done in-house. Second, our data contain limited information about the content of the ads. Specifically, we have a variable on whether the ad contains words beyond just the company logo. We argue that ads that contain more than the company logo are relatively original and creative, though we admit this is a low threshold for creativity.

Columns (1) and (2) compare the effectiveness of ads that were designed by a named ad agency compared to those that were not. It is clear that the ads that were not designed by an ad agency were more negatively affected by standardization, suggesting that the original input of the ad agency helped them stand out. Columns (3) and (4) compare ads that had explicit ad copy compared to those that did not. This suggests that focusing the advertising appeal on a different dimension (wording) as opposed to format helped such ads stand out after standardization.

These results also serve as a further validation test (in addition to our instrumental variables results) against the possibility that firms only move their worse ads to the standard

Table 5: Originality: Ad content

	(1)	(2)	(3)	(4)
	Ad Agency	No Ad Agency	Ad Copy	No Ad Copy
Exposed \times Standard Format \times After Standardization	0.0232 (0.0247)	-0.0258*** (0.00739)	-0.0107 (0.0509)	-0.0215*** (0.00726)
Exposed \times Standard Format	-0.0351 (0.0239)	-0.00153 (0.00646)	-0.0508 (0.0502)	0.00793 (0.00641)
Exposed	0.189*** (0.0127)	0.101*** (0.00313)	0.156*** (0.00709)	0.0940*** (0.00337)
Exposed \times After Standardization	-0.0976*** (0.0134)	0.00127 (0.00426)	-0.0132 (0.00944)	-0.00488 (0.00432)
Demographic controls	Yes	Yes	Yes	Yes
Campaign controls	Yes	Yes	Yes	Yes
Observations	70931	283444	55815	298560
R2	0.126	0.136	0.124	0.137

OLS regression coefficients shown, data are for August 2002 to August 2004, and the dependent variable is an indicator variable for whether the person recognized the ad. The coefficients on non-instrumented variables in the first stage are not shown. ‘StandardFormat’ is collinear with campaign fixed effects and omitted (see page 15 for details). Robust standard errors clustered at the campaign level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

format. If our results were driven by increasingly poor ads showing up in the standard format, we should not see a change in the effectiveness of poor quality ads after standardization, and no difference in the change in relative effectiveness of low and high quality ads.

Broadly, these results are consistent with an interpretation of the reduced effect of standard-format ads after standardization being causally due to the standardization, most likely because of a change in attention given to the ads. Originality, however, nullifies the effect of a lack of differentiation due to this standardized format, suggesting a useful strategy that advertisers can use to overcome these challenges.

6 Implications and Conclusion

This paper uses rich field experiment data from real online advertising campaigns to investigate how the increasing use of a standard format for advertising affects the memorability of advertising and how it affects purchase intent. We examine the effects of the first attempts to set formal format standards for online display advertising in the US by the Interactive Advertising Bureau. We find evidence that the sharp rise in the use of standard formats reduced the effectiveness of standard format ads by over 20%. This reduction in effectiveness

was mitigated by adding original content to the ads as would likely be accomplished through an ad agency.

Our results also have important implications for media platforms, for advertisers, and for evaluating the benefits of moving to a standard format. Generally, the reason that media platforms try to set cross-platform standards is that standards facilitate the placement and use of a single format across multiple platforms. In a world of automatic advertising content delivery and real-time ad exchanges, this has become particularly important. We present evidence that while such standards-setting may be beneficial in terms of efficiency for both advertiser and platform, it reduces the ability of ads to attract attention.

A key novel finding of our study is that standardization did not affect all ads negatively. Instead, it appeared to have had little effect on ads that were made by specialists in ad design. Ads which were reasonably generic in design were the ones most negatively affected. This leads to a somewhat unexpected conclusion. The process of standardization might, by virtue of standardizing some design elements, promote greater creativity in other design elements, as in Sellier and Dahl (2011). Our results suggest that the constraints imposed by standardization led to a clear strategic response: An increase in ad creativity as the marginal return to ad creativity has risen.

Speculatively, this suggests another benefit of standardization for the industry: it reduces the role of plain ads that require less insider expertise. This relative benefit to insiders over outsiders suggests that the socially optimal set of standards may differ from those chosen by an industry standards body in important ways, perhaps favoring creativity over simplicity (Simcoe, 2012).

There are of course limitations to our study that suggest potential avenues for future research. First, we focus exclusively on the adoption of standards for online display advertising. We do not know exactly the extent to which our results will generalize to standards-setting processes in general and to other media in particular. Second, we measure the effect of

standardization on the type of measures that advertisers themselves use to measure ad effectiveness, but do not have data about how the pricing of ads, or their true effectiveness, changed as a result of this standardization process. Therefore, we do not know how this process affected media platform revenues. Third, the standards were known well in advance. This means that our treatment of the rise of standard format as a natural experiment is driven by the seemingly sudden acceptance of the format coincident with the April 2003 official launch rather than any new information about the formats available on the market. Fourth, we measure aided ad recognition and purchase intent rather than actual purchases as a result of advertising. Given that our data come from a key source for benchmarking ad effectiveness, a weaker interpretation of our results is that the increasing use of a standard format reduced a key measure of ad effectiveness used by the industry. Fifth, we focus on measuring the (asymmetric) costs of the increasing use of standards, leaving the analysis of the benefits to other work (including Rysman and Simcoe (2008), David and Greenstein (1990) and others). Given that the standards were widely adopted, the benefits likely outweighed the costs in total. Finally, while our results are suggestive of a role for reduced attention through habituation, the field setting means we do not directly measure either reduced attention or habituation. Instead, we proxy for attention with ad recognition measures and we proxy for habituation with the increased use of the standard formats over time.

Notwithstanding these limitations, we believe our study does represent an important step in understanding how the increased use of a standard format for advertising affects consumer response, and therefore how it might affect advertiser and media platform strategies.

References

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton Press.

- Aribarg, A., R. Pieters, and M. Wedel (2010). Raising the bar: Bias adjustment of recognition tests in advertising. *Journal of Marketing Research* 47(3), 387–400.
- Armstrong, M. (2006). Competition in two-sided markets. *RAND Journal of Economics* 37(3), 668–691.
- Augereau, A., S. Greenstein, and M. Rysman (2006). Coordination vs. differentiation in a standards war: 56k modems. *RAND Journal of Economics* 37(4), 887–909.
- Chen, Y., Q. Wang, and J. Xie (2011). Online social interactions: A natural experiment on word of mouth versus observational learning. *Journal of Marketing Research* 48(2), 238–254.
- David, P. and S. Greenstein (1990). The economics of compatibility standards: An introduction to recent research. *Economics of Innovation and New Technology* 1, 3–42.
- Dreze, X. and F.-X. Husherr (2003). Internet advertising: Is anybody watching? *Journal of Interactive Marketing* 17(4), 8–23.
- Dukes, A. and E. GalOr (Spring 2003). Negotiations and exclusivity contracts for advertising. *Marketing Science* 22(2), 222–245.
- Farrell, J. and G. Saloner (1988). Coordination through committees and markets. *RAND Journal of Economics* 19(2), 235–252.
- Farrell, J. and T. S. Simcoe (2012). Choosing the Rules for Consensus Standardization. *RAND Journal of Economics* 43(2), 235–252.
- Finn, A. (1988). Print ad recognition reeadship scores: An information processing perspective. *Journal of Marketing Research* 25(2), 168–77.

- Ghose, A. and S. Yang (2009). An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management Science* 55(10), 1605–1622.
- Goldfarb, A. and C. Tucker (2011a, April). Advertising Bans and the Substitutability of Online and Offline Advertising. *Journal of Marketing Research* 48(2), 207–227.
- Goldfarb, A. and C. Tucker (2011b). Online display advertising: Targeting and obtrusiveness. *Marketing Science* 30(3), 389–404.
- Goldfarb, A. and C. Tucker (2011c). Privacy regulation and online advertising. *Management Science* 57(1), 57–71.
- Horrace, W. C. and R. L. Oaxaca (2006). Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economic Letters* 90, 321–327.
- Katz, M. L. and C. Shapiro (1985). Network externalities, competition, and compatibility. *American Economic Review* 75(3), 424–440.
- Kent, R. J. and K. A. Machleit (1990). The differential effects of within-brand and between-brand processing on the recall and recognition of television commercials. *Journal of Advertising* 19(2), 4–14.
- Lambrecht, A. and C. Tucker (2013). When does retargeting work? Information specificity in online advertising. *Journal of Marketing Research* 50(5), 561–576.
- Lewis, R. and D. Reiley (2009). Retail advertising works! Measuring the effects of advertising on sales via a controlled experiment on Yahoo!”. Working Paper, Yahoo! Research.
- Manchanda, P., J.-P. Dube, K. Y. Goh, and P. K. Chintagunta (2006). The effect of banner advertising on internet purchasing. *Journal of Marketing Research* 43(1), 98 – 108.
- Pashler, H. E. (1998). *The Psychology of Attention*. MIT Press, Cambridge MA.

- Peebles, D. M., J. K. R. Jr., and I. R. Vernon (1978). Coordinating international advertising. *Journal of Marketing* 42(1), 28–34.
- Pieters, R., L. Warlop, and M. Wedel (2002). Breaking through the clutter: Benefits of advertisement originality and familiarity for brand attention and memory. *Management Science* 48(6), 765–781.
- Rysman, M. and T. Simcoe (2008). Patents and the Performance of Voluntary Standard-Setting Organizations. *Management Science* 54(11), 1920–1934.
- Sellier, A.-L. and D. Dahl (2011). Focus! creative success is enjoyed through restricted choice. *Journal of Marketing Research* 48(6), 996–1007.
- Shapiro, C. and H. R. Varian (1999). *Information Rules: A Strategic Guide to the Network Economy*. Boston: Harvard Business School Press.
- Shepard, T. M. (1942). The starch application of the recognition technique. *Journal of Marketing* 6(2), 118–24.
- Simcoe, T. (2012). Standard setting committees: Consensus governance for shared technology platforms. *American Economic Review* 102(1), 305–36.
- Solomon, M. R. (1999). *Consumer Behavior, Fourth Edition*. Prentice Hall, Upper Saddle River, New Jersey.
- Starch, D. (1923). *Principles of Advertising*. A.W. Shaw Company, Chicago.
- Sun, M. and F. Zhu (2013). Ad revenue and content commercialization: Evidence from blogs. *Management Science* 59(10), 2314–2331.
- Takada, H., M. Mizuno, and L. Bith-Hong (2012). *Handbook of Research on International Advertising*, Chapter Analysis of the relationship between advertisers and advertising agencies in the global market, pp. 497–518. Edward Elgar Publishing.

Taylor, C. (2001). IAB guidelines on units may revive the Net. *Advertising Age Feb 26*.

Wells, W. D. (1964). Recognition, recall and rating scales. *Journal of Advertising Research* 4(3), 2–8.

A Press Release

April 28, 2003 IAB ANNOUNCES FINAL INTERACTIVE UNIVERSAL Industry Survey
Feedback Supports Four Large Ad Sizes

Today, the Interactive Advertising Bureau (IAB) announced the new Universal Ad Package (UAP), a creative suite of four ad sizes that will enable advertisers to reach the majority of each online publisher's audience. Designed in response to advertiser demand for more standard online advertising guidelines, this creative suite will ensure a greater consistency with online ads regardless of where they are published on the Web. The UAP is intended to improve the efficiency and ease to planning, buying and creating online media. The UAP has the support of the American Association of Advertising Agencies (AAAA).

The Universal Ad Package interactive units (IU) include: IU 728 x 90 IU 300 x 250 IU 160 x 600 IU 180 x 150

If a publisher is UAP compliant, an advertiser can buy, plan and create around four units knowing they can reach the majority (51%) of that publisher's audience. Buyers can identify IAB member UAP compliant sites by the UAP Compliance Seal (attached).

The UAP offers a win-win for agencies and their clients. Now, as with a 30-second spot for TV, agencies can plan against a standard set of ad units. The UAP affords a simpler approach to interactive media planning without impacting flexibility for advertisers to execute compelling, unique online advertising campaigns, said Jonathan Adams, Senior Partner, Group Media Director, mOne Worldwide (Chairman, AAAA, Eastern Interactive Marketing and New Media Committee).

In December 2002, the IAB Ad Sizes Task Force recommended the four interactive UAP ad sizes and solicited feedback on the proposed units from industry stakeholders including agencies, advertisers and online publishers. The results of this survey confirmed that the chosen sizes recognized and conformed to the needs of the media buying community. The

initial four sizes were chosen based on customer feedback, extensive usability studies and brand and traditional click performance tests.

The current IAB member companies that are currently, or plan to be compliant in the next 12-18 months, include: 24/7 Real Media Inc., About, Inc., America Online, Inc., CBS SportsLine.com, Classmates Online, Inc., CNET Networks, CondeNet, Edmunds.com, Inc., The Excite Network, Forbes.com, Inc., iVillage Inc., MarketWatch.com, Inc., Meredith Corporation, MSN, New York Times Digital, Terra Lycos, Univision Online, USAToday.com, Wall Street Journal Online, The Walt Disney Internet Group, Washingtonpost/Newsweek Interactive, The Weather Channel Interactive, Inc. and Yahoo!.

The buying community implored publishers to simplify the planning process for interactive media, and we did just that. We listened and are reshaping our sites to accommodate these needs. This industry is determined to prove our commitment to our advertising clients. We are set to take the industry to a new level that offers advertisers best practices and leads to equal or greater results than other media vehicles such as TV and print, said Joanne Bradford, MSN vice president and chief media revenue officer. MSN is firmly committed to making online a better environment for advertisers to reach consumers and interact in a meaningful way.

From an agency perspective, the UAP exploits the best aspects of interactive advertising but does not impact flexibility and creative option. With the UAP, an advertiser retains the ability to develop an ad in any size, shape or form they wish, whether it be a half-page or otherwise, and at the same time, they have the reassurance that they can create against the UAP and reach their desired audience, said Matt Freeman, CEO, Worldwide, Tribal DDB.

In less than 9 months, the IAB Ad Sizes Task Force has delivered a program that will ultimately result in one of the most meaningful changes in this industry to date. The UAP presents a uniform platform against which advertisers and agencies can develop campaigns with maximum efficiency thus reducing the barriers to entry for the media buying community,

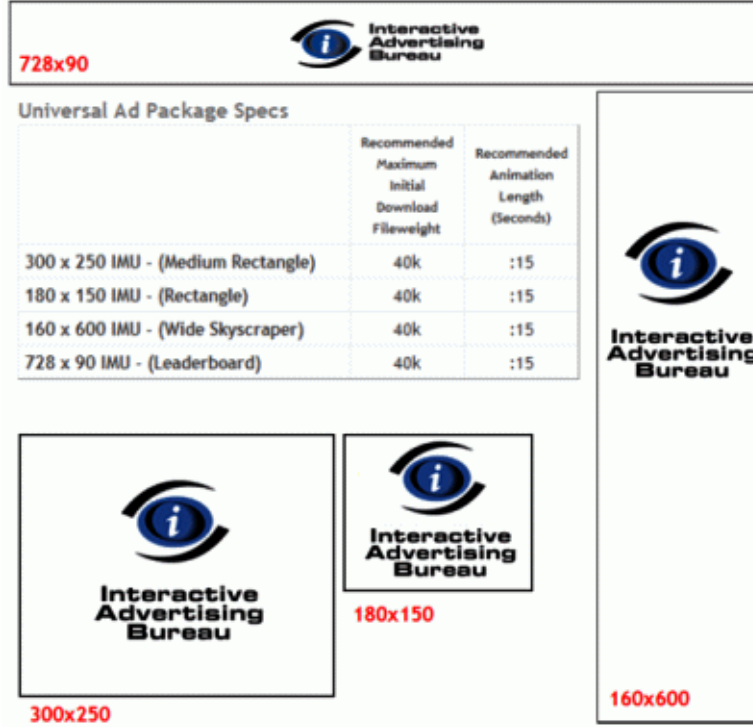


Figure A-1: Mock-Up of Compliant Ads

said Adam Gelles, Director, Industry Initiatives, IAB.

At long last, this medium is using ad units that are the size people have come to expect in offline media. For that reason, I think they will generate attention and response, said Mike Donahue, Executive Vice President, AAAA.

B Details of Instrumental Variables Strategy

For estimation, this instrument should directly affect the probability that a firm adopts a standard format after standardization, $StandardFormat_j \times AfterStandardization_t$. This generates a instrumental variables approach with the following first stage for the key endogenous variable.

$$\begin{aligned}
& StandardFormat_j \times AfterStandardization_t = \quad (A-1) \\
& \alpha AfterStandardization_t \times ProportionInternational_j + \\
& \beta_1 AfterStandardization_t + \beta_2 ProportionInternational_j + \theta X_{ij} + \epsilon_{ijt}
\end{aligned}$$

The endogenous variable of interest in the second stage is the three-way interaction $Exposed_{ij} \times StandardFormat_j \times AfterStandardization_t$. Though $StandardFormat_j \times AfterStandardization_t$ is endogenous, in the main equation it is collinear with the campaign fixed effects so is not reported.