Online Display Advertising: Targeting and Intrusiveness

Citation: Goldfarb, A., and C. Tucker. "Online Display Advertising: Targeting and Obtrusiveness." Marketing Science 30.3 (2011): 389-404.

As Published: http://dx.doi.org/10.1287/mksc.1100.0583

Publisher: INFORMS

Persistent URL: http://hdl.handle.net/1721.1/67901

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Online Display Advertising: Targeting and Obtrusiveness

Avi Goldfarb and Catherine Tucker*

February 2010

Abstract

We use data from a large-scale field experiment to explore what influences the effectiveness of online advertising. We find that matching an ad to website content and increasing an ad's obtrusiveness independently increase purchase intent. However, *in combination* these two strategies are ineffective. Ads that match both website content *and* are obtrusive do worse at increasing purchase intent than ads that do only one or the other. This failure appears to be related to privacy concerns: The negative effect of combining targeting with obtrusiveness is strongest for people who refuse to give their income, and for categories where privacy matters most. Our results suggest a possible explanation for the growing bifurcation in internet advertising between highly targeted plain text ads and more visually striking but less targeted ads.

Keywords: E-commerce, Privacy, Advertising, Targeting

*Avi Goldfarb is Associate 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 Assistant Professor of Marketing, MIT Sloan School of Business, 1 Amherst St., E40-167, Cambridge, MA. Tel. 617-252-1499. Email: cetucker@mit.edu. We thank Pankaj Aggarwal and seminar participants at UT-Dallas for helpful comments. Financial support from the Google and WPP Marketing Research Awards Program is gratefully acknowledged.

1. Introduction

Customers actively avoid looking at online banner ads (Dreze and Hussherr 2003). Response rates to banner ads have fallen dramatically over time (Hollis 2005). In reaction to this, online advertising on websites has developed along two strikingly distinct paths.

On the one hand, the $11.2 billion[1] online display advertising market has evolved beyond traditional banner ads to include many visual and audio features that make ads more obtrusive and harder to ignore. On the other hand, Google has developed a highly profitable non-search display advertising division (called AdSense) that generates an estimated $6 billion in revenues by displaying plain content-targeted text ads: 76% of US internet users are estimated to have been exposed to AdSense ads.[2] This paper explores how well these divergent strategies work for online advertising, and how consumer perceptions of intrusiveness and privacy influence their success or lack of it, both independently and in combination.

We examine the effectiveness of these strategies using data from a large randomized field experiment on 2,892 distinct web advertising campaigns that were placed on different websites. On average, we have data on 852 survey-takers for each campaign. Of these, half were randomly exposed to the ad, and half were not exposed to the ad but did visit the website on which the ad appeared. These campaigns advertised a large variety of distinct products, and were run on many different websites.

The advertisers in our data used two core improvements on standard banner ad campaigns to attract their audience. (1) Some web campaigns matched the product they advertised to the website content, for example when auto manufacturers placed their ads on auto websites. (2)

---

[1] Hallerman (2009).
[2] Estimates generated from Google quarterly earnings report for June 30, 2009 "Form 10-Q", filed with the Securities and Exchange Commission (SEC).

Some web campaigns deliberately tried to make their ad stand out from the content by using video, creating a pop-up, or having the ad take over the webpage. This paper evaluates whether "targeted" campaigns that complemented the website content, "obtrusive" campaigns that strove to be highly visible relative to the website content, or campaigns that did both, were most successful at influencing stated purchase intent, measured as the difference between the group that was exposed to the ad and the group that was not.

We construct and estimate a reduced-form model of an ad effectiveness function. Consistent with prior literature (e.g. Goldfarb and Tucker 2009; Wilbur, Goeree, and Ridder 2009), our results suggest that matching an ad's content to the website content increased stated purchase intent among exposed consumers. Also consistent with prior literature (e.g. Cole, Spalding, and Fayer 2009), our results suggest that increasing the obtrusiveness of the ad increased purchase intent. Surprisingly, however, we find that these two ways of improving online display advertising performance negate each other: Combining them nullifies the positive effect that each strategy has independently. These results are robust to multiple specifications, including one that addresses the potential endogeneity of campaign format by restricting our analysis to campaigns that were run both on sites that matched the product and sites that did not.

These results have important economic implications for the $664 million that we estimate is currently spent on ads that are both targeted and obtrusive. If advertisers replace ads that combine contextual targeting and high visibility with the standard ads that our estimates suggest are equally effective, we provide back-of-the-envelope calculations that suggest advertisers could cut ad spending by over 5% without affecting ad performance.

It is not obvious why increasing visibility and increasing the match to content should work well separately but not together. It does not appear to be explained by differences in ad recall. Consistent with several laboratory experiments in other settings (Mandler 1982; Heckler and Childers 1992), we find that contextually-targeted ads are recalled less, while highly visible ads are recalled more. However, we find that highly visible ads are not recalled significantly less if they are matched to the context. This suggests that the negation mechanism is explained by a difference in how successful the ad is at influencing customer behavior after customers see it, rather than because of differences in ad recall.

The literature on consumer response to persuasion attempts provides an alternative explanation: Obtrusive ads may lead consumers to infer that the advertiser is trying to manipulate them, reducing purchase intentions (Campbell 1995). Specifically, increased processing attention may lead the consumer to think about why a particular advertising tactic was used (Campbell 1995; Friestad and Wright 1994). If the tactic is perceived as manipulative, it will have a negative effect on consumer perceptions of the product advertised. Given that deception is particularly easy online, consumer awareness of manipulation is higher too (Boush, Friestad, and Wright 2009). This suggests that targeted and obtrusive banner ads, by increasing the attention paid to the tactic of targeting, may generate consumer perceptions of manipulation. Therefore, while there is a relatively high consumer tolerance to targeted ads because the information is perceived as useful (e.g. Cho and Cheon 2003; Edwards, Li, and Less 2002; Wang, Chen, and Chang 2008), this tolerance may be overwhelmed by perceptions of manipulation when the ad is obtrusive.

Why might the negative consequences of perceived manipulation from using techniques to make ads obtrusive be higher if they are targeted? Privacy concerns provide a possible answer. Both Turow et al. (2009) and Wathieu and Friedman (2009) document that customer appreciation of the informativeness of targeted ads is tempered by privacy concerns. When privacy concerns are more salient, consumers are more likely to have a prevention focus (Van Noort, Kerkhof, and Fennis 2008), in which they are more sensitive to the absence or presence of negative outcomes, instead of a promotion focus in which they are more sensitive to the absence or presence of positive outcomes. Kirmani and Zhu (2007) show that prevention focus is associated with increased sensitivity to manipulative intent, suggesting a likely avenue for the relationship between targeted ads, privacy concerns, and perceived manipulative intent that arises from making ads more obtrusive.

We explore empirically whether it is privacy concerns that drive the negative effect we observe of combining targeting and obtrusiveness. We find evidence that supports this view: Contextual targeting and high visibility (obtrusiveness) are much stronger substitutes for people who refused to answer a potentially intrusive question on income. They are also stronger substitutes in categories in which privacy might be seen as relatively important (such as financial and health products).

In addition to our major finding that obtrusiveness and targeting do not work well in combination, and the link this has with consumer perceptions of privacy, we make several other contributions that help illuminate online advertising. Our examination of many campaigns across many industries complements the single-firm approach of the existing quantitative literature on online advertising effectiveness. For example, Manchanda et al. (2006) use data from a healthcare and beauty product internet-only firm in their examination of how banner ad exposures affect sales. Similarly, Lewis and Reiley (2009), Rutz and Bucklin (2009), Chatterjee, Hoffman, and Novak (2003), and Ghose and Yang (2009) all examine campaigns run by just one firm. Those studies have given us a much deeper understanding of the relationship between online advertising and purchasing, but their focus on particular firms and campaigns makes it difficult to draw general conclusions about online advertising. In contrast, our research gives a sense of the average effectiveness of online advertising—it boosts stated purchase intent by 3-4%—and our research suggests that at current advertising prices, plain banner ads pay off if an increase in purchase intent by one consumer is worth roughly 42 cents to the advertiser. Our results also emphasize that there is a role for cross-category and campaign advertising studies because they allow us to examine factors that vary across campaigns like campaign visibility and category characteristics.

The trade-off between perceived intrusiveness and usefulness may help explain the unexpected success of products such as Google's AdSense, which generates roughly one third of Google's advertising revenues (the other two thirds coming directly from search advertising). AdSense allows advertisers to place very plain-looking ads—they are identical in appearance to Google’s search ads—on websites with closely matched content. Our findings suggest that making these ads more visually striking could be counterproductive, or at least waste advertising budget. More generally, the results suggest two alternative, viable routes to online display advertising success: Putting resources into increasing the visibility of ads, and putting resources into the targeting of ads to context, but *not* doing both.

2. Data on Display Advertising

We use data from a large database of surveys collected by a media measurement agency to examine the effectiveness of different ad campaigns. These surveys are based on a randomized exposed and control allocation. Individuals browsing the website where the campaign is running are either exposed to the ads, or not, based on the randomized operation of the ad server. On average 198 subjects are recruited for each website running a particular campaign with an average of 852 subjects being recruited across all websites for each ad campaign. The average campaign lasted 55 days (median 49 days).

We excluded ad campaigns run on websites that were described as “other” because it was not possible to identify whether there was a contextual match. We also excluded ads for alcohol because there were no contextually targeted alcohol ads in the data. Finally, we excluded ads for prescription drugs because FTC regulations (which require reporting of side-effects) restrict their format.

Both exposed and non-exposed (control) respondents are recruited via an online survey invitation that appears after they have finished browsing the website. Therefore, our coefficients reflect the stated preferences of consumers who are willing to answer an online poll. The company that makes these data available has done multiple checks to confirm the general representativeness of this survey among consumers. In section 4.4, we find that the main negative interaction effect that we study is larger for people who refuse to answer an (intrusive) question on income, so it is even possible that the selection in our subject pool away from those who prefer not to answer surveys leads us to understate the magnitude of this negative interaction.

Because of the random nature of the advertising allocation, both exposed and control groups have similar unobservable characteristics, such as the same likelihood of seeing offline ads. The only variable of difference between the two groups is the randomized presence of the ad campaign being measured. This means that differences in consumer preferences toward the advertiser’s product can be attributed to the online campaign.

This online questionnaire asked the extent to which a respondent was likely to purchase a variety of products (including the one studied) or had a favorable opinion of the product using a five-point scale. The data also contained information about whether the respondent recalled seeing the ad. After collecting all other information at the end of the survey, the survey displayed the ad, alongside some decoy ads for other products, and the respondents were asked whether they recalled seeing any of the ads. If they responded that they had seen the focal ad, then we code this variable as one and as zero otherwise.

An important strength of this data set is that it contains a large number of campaigns across a variety of categories, including automotive, apparel, consumer packaged goods, energy, entertainment, financial services, home improvement, retail, technology, telecommunications, travel, and many others. Therefore, like Clark, Doraszelski, and Draganska's (2009) study of offline advertising, we can draw very general conclusions about online display advertising. Our data have the further advantage of allowing us to explore ad recall and purchase intent separately.

The survey also asked respondents about their income, age, and the number of hours spent on the internet. We use these as controls in our regressions. We converted them to zero-mean-standardized measures. This allowed us to "zero out" missing data.[3] In Table 3, we show robustness to a non-parametric specification for the missing data by discretizing the variables with missing information. We also show robustness to excluding the controls.

In addition to data from this questionnaire for each website user, our data set also documents the characteristics of the website where the advertisement appeared and the characteristics of the advertisement itself. We used this data to define both whether the ad was "Contextually Targeted" and whether it was obtrusive or "High Visibility."

[3] Data are missing for two reasons: sometimes that question was not asked, and sometimes respondents refused to answer. We have no income information for 28% of respondents; no age information for 2% of respondents; and no hours online information for 34% of respondents.

2.1 Defining Contextually-Targeted Advertising

To define whether a product matched the website that displayed it, we matched up the 364 narrow categories of products, with the 37 categories of websites described in the data.[4] A banner ad for a cruise would be a "Contextually-Targeted Ad" if it was displayed on a website devoted to "Travel and Leisure". Similarly, a banner ad for a new computer would be a "Contextually-Targeted Ad" if it was displayed on a site devoted to "Computing and Technology". As indicated in Table 1a, 10 percent of campaigns were run on sites where the product and website content matched.

2.2 Defining High Visibility (Obtrusive) Advertising

We define an ad as high-visibility if it has one of the following characteristics:

- Pop-Up: The ad appears in a new window above the existing window.
- Pop-Under: The ad appears in a new window that lies underneath the existing window
- In-stream Video and Audio: The ad is part of a video stream.
- Takeover: The ad briefly usurps the on-screen space a webpage has devoted to its content.
- Non-User-Initiated Video and Audio: The ad automatically plays video and audio.
- Interstitial: The ad is displayed before the intended destination page loads.
- Non-User-Initiated Background Music: The ad automatically plays background music.
- Full Page Banner Ad: The ad occupies the space of a typical computer screen.
- Interactive: The ad requests two-way communication with its user.

[4] Our data have detailed information about the categories but, to protect client privacy, the firm that gave us access to the data did not give us information on which specific advertisers sponsored particular campaigns and which specific websites displayed the campaigns.

- Floating Ad: The ad is not user-initiated, is superimposed over a user-requested page, and disappears or becomes unobtrusive after a specific time period (typically 5-30 seconds).

As indicated in Table 1a, 48 percent of campaigns used high-visibility ads. Of these ads, the largest sub-group was interactive ads, which comprised 17 percent of all ads. Pop-Under ads were the least common, representing less than 0.1 percent of all ads. In Table 3, we check the robustness of our results to our definition of high visibility.

Table 1a: Summary Statistics at the respondent level

	Mean	Std Dev	Min	Max	Observations
Likely to purchase	0.20	0.40	0	1	2464812[+]
Saw Ad	0.26	0.44	0	1	2481592[+]
Exposed (Treatment)	0.34[++]	0.47	0	1	2802333
Context Ad	0.10	0.30	0	1	2802333
High Visibility Ad	0.49	0.50	0	1	2802333
Female	0.53	9.49	0	1	2802333
Income	62,760	55,248	15,000	250,000	2019996
Age	40.9	15.6	10	100	2744149
Internet Hours	13.4	10.2	1	31	1853893

[+]A small fraction of respondents answered two out of the three questions on the measures we use as dependent variables. Therefore, the sample size varies slightly across variables.
[++]In our main specifications, we exclude observations where the random operation of the ad server meant that the respondent had been exposed to the ad multiple times. If we include these observations, the proportion of exposed to non-exposed is close to 50 percent. We confirm robustness to these excluded observations in Table 3.

Table 1b: Summary Statistics at the survey-site level

Variable	Mean	Std Dev	Min	Max	Observations
# of Subjects	198	275.1	8	8456	13121
Context Ad	0.13	0.34	0	1	13121
High Visibility Ad	0.50	0.50	0	1	13121
Difference in Purchase Intent between exposed & control	0.0083	0.15	-1	1	10468
Difference in Favorable Opinion between exposed & control	0.0100	0.16	-1	1	10362
Difference in intent-scale between exposed & control	0.040	0.54	-4	4	10468
Difference in opinion-scale between exposed & control	0.036	0.41	-3.67	4	10362
Difference in Ad Recall between exposed & control	0.046	0.18	-1	1	10750
Campaign increased Purchase Intent	0.60	0.49	0	1	13121
Campaign increased Favorable Intent	0.61	0.49	0	1	13121
Campaign increased Intent-Scale	0.64	0.48	0	1	13121
Campaign increased Opinion-Scale	0.65	0.48	0	1	13121
Campaign increased Ad Recall	0.70	0.46	0	1	13121
Campaign increased Ad Recall but not Purchase Intent	0.23	0.42	0	1	13121
Campaign increased Purchase Intent but not Ad Recall	0.13	0.11	0	1	13121

Table 1b presents some summary statistics at the level of the survey-site. Two things are immediately apparent. First, in general the effect of exposure to a banner ad is small. On average, exposure shifts the proportion of people stating they are very likely to make a purchase by less than 1%. Ad exposure increases the number of people who can actually recall seeing the ad by only 5%. Only 60% of campaigns generated an improvement in purchase intent, while 70% increased ad recall. 23% of campaigns generate ad recall but do not increase stated purchase intent, and 13% of campaigns are associated with an increase in purchase intent but not ad recall. Perhaps this lack of impact is unsurprising given that we are assessing whether seeing a banner ad once changes intended behavior. What really matters is whether this small change in intended behavior is worth the price paid (0.2 cents per view according to Khan 2009).

3. Methods

Our estimation strategy builds an effectiveness function for the individual ads. We focus on two aspects of online advertising: visibility and contextual targeting.[5] Specifically, assume the effectiveness of a campaign *c* (defined as a particular advertisement shown on a particular website) to individual *j* is:

$$Effectiveness_j^c = f_j(Context^c, HighVisibility^c, Context^c \times HighVisibility^c)$$

For estimation, we convert this effectiveness function for viewer *j* into a linear model of visibility and targeting which allows for some idiosyncratic features of the advertisement-website pair (μ^c) and some idiosyncratic characteristics of the viewer (ε_j^c). As in any research which estimates a response function, there is a concern here that viewer characteristics may be systematically correlated with the propensity to view contextually-targeted or highly visible ads (see Levinsohn and Petrin 2003 for a discussion in the context of production functions). The randomized nature of our data means that we can address this by measuring the difference between the exposed group and a control group of respondents who did not see the ad. Substituting our measure of effectiveness (purchase intent) and adding a control for potential demographic differences between the treatment and control groups, we estimate the effect of being exposed to the ad using the following difference-in-difference equation:

$$Intent_j^c = \alpha Exposed_j^c + \beta Exposed_j^c \times Context^c + \gamma Exposed_j^c \times HighVisibility^c$$

$$+\delta Exposed_j^c \times Context^c \times HighVisibility^c + X_j\theta + \mu^c + \varepsilon_j^c$$

[5] Of course, there are many features of the ad, the viewer, and the website that may influence advertising effectiveness but are not the primary focus on this study. Our analysis controls for these through the randomization of the exposed condition, the campaign-specific fixed effects, and the viewer characteristics.

Where X_j represents a vector of demographic controls (specifically, an indicator variable for whether the respondent is female and mean-standardized measures for age, income, and time spent online), μ^c represents the campaign (advertisement-website) fixed effects that control for any differences in purchase intent across products and websites, and ε_j^c is an idiosyncratic error term. The fixed effects capture the main effects (i.e. those that affect both the exposed and control groups) of context, visibility, and the interaction between context and visibility as well as heterogeneity in the response to different campaigns. Whether context and visibility are substitutes or complements is captured by the coefficient δ. This equation can be interpreted as a reduced-form effectiveness function for advertising, in which the advertising produces purchase intent (our proxy for effectiveness).

In the main specifications in this paper, we use as our main dependent variable ("$Intent_j^c$") whether the survey-taker reported the highest score on the scale ("Very Likely to Make a Purchase"). As reported in our summary statistics (Table 1a), on average one-fifth of respondents responded with an answer at the top of the scale. We discretize in this manner to avoid generating means from an ordinal scale (e.g. Malhotra 2007; Aaker, Kumar, and Day 2004). Nevertheless, we recognize that whether such scales should be treated as continuous is a gray area in marketing practice (discussed by, e.g., Fink 2009, p. 26). Bentler and Chou (1987), while acknowledging that such scales are ordinal, argue that it is common practice to treat them as continuous variables because it makes little practical difference. Johnson and Creech (1983) come to a similar conclusion. In addition, many researchers argue that properly-designed scales are in fact interval scales not ordinal scales (see Kline 2005 for a discussion). Some recent empirical research in marketing (e.g. Godes and Mayzlin 2009) has followed this interpretation and has treated ratings scales as interval scales. For these reasons, we have replicated all of our results in the appendix, treating the dependent variable as a continuous measure based on the full-scale responses. We show that this makes no practical difference.

Another issue from using purchase intention as our dependent variable is that it is a weaker measure of advertising success than purchasing or profitability (as used by Lewis and Reiley 2009 and others), because many users may claim that they intend to purchase but never do so. For our purposes, as long as people reporting higher purchase intent are actually more likely to purchase (and the treatment group is truly random), our conclusions about the directionality of the relationship between contextual targeting, high visibility, and effectiveness will hold. This positive correlation has been well established in work such as Bemmaor (1995), and in particular has been found to be higher for product-specific surveys such as the ones conducted in our data than for category-level studies (Morwitz, Steckel, and Gupta 2007). However, we do not know whether the relative size or relative significance of the results changes for actual purchase behavior.

Our estimation procedure is straightforward because of the randomized nature of the data collection. We have an experiment-like setting, with a treatment group who were exposed to the ads and a control group who were not. We compare these groups' purchase intent, and explore whether the difference between the exposed and control groups is related to the visibility of an ad and whether the ad's content matches the website content. We then explore how purchase intent relates to ads that are both highly visible and targeted to the content. Our core regressions are run using Stata’s commands for linear regression with panel data. To adjust for intra-website and campaign correlation between respondents, heteroskedasticity-robust standard errors are clustered at the website-campaign level using Stata's "cluster" function. Generally, our method follows the framework for causal econometric analysis provided in Angrist and Pischke (2009).

We report our main results using a linear probability model where the coefficients are calculated using ordinary least squares, though we show robustness to a logit formulation. We primarily use a linear probability model because it makes it feasible to estimate a model with over ten thousand campaign-website fixed effects using the full data set of nearly 2.5 million observations, whereas computational limitations prevent us from estimating a logit model with the full data. The large number of observations in our data means that inefficiency—a potential weakness of the linear probability model relative to probit and logit—is not a major concern in our setting. Angrist and Pischke (2009) point out that this increased computational efficiency comes at little practical cost. They show that in several empirical applications, there is little difference between the marginal effects estimated with limited dependent variables models and linear probability models.

One major concern about using the linear probability model, which is the potential for biased estimates and predicted probabilities outside the range of 0 and 1 (see Horrace and Oaxaca 2006 for a discussion), is less likely in our setting because the mass point of dependent variable is far from 0 or 1 and because our covariates are almost all binary variables. Indeed, all predicted probabilities in our model for purchase intent lie between 0.137 and 0.256. Combined with its computational advantages, this suggests that the benefits of using a linear probability model as our primary estimation technique outweigh the costs.

4. Results

4.1 The effect of combining highly visible and contextually-targeted ads

Column (1) of Table 2 reports our key results, where we include an interaction between contextual targeting and high visibility. The main effect of exposure and the additional effects of contextual targeting and high visibility of ads are all positive. Contextual targeting seems to have a slightly larger marginal effect than high visibility. The crucial result is the interaction term between exposure, contextual targeting, and high visibility. It is *negative* and significant. This suggests that firm investments in contextual targeting and highly visible ads are substitutes. A Wald test suggests that high-visibility ads on contextually-targeted sites perform worse than ads that are not highly visible (p-value=0.001, F-stat 10.83). The respondent-level controls indicate that younger, female respondents who have lower incomes and spend more time on the internet are more likely to say they will buy the product.

In columns (2)-(4) we present some evidence that reflects what a manager would conclude if they evaluated the incremental benefit of targeting or using high visibility ads independently. Column (2) reports results that allow the effect of exposure to be moderated by whether or not the website matched the product. We find a positive moderating effect of contextual targeting. That is, for campaigns where the website matches the product, there is an incremental positive effect from exposure that represents a 70% increase from the base positive effect of exposure. However, it is noticeable that without the controls for visibility that we included in column (1), we measure the effect of contextual targeting less precisely. Column (3) reports results that allow the effect of exposure to be moderated by whether the ad was high-visibility or not. We find a positive moderating effect of visibility on likelihood of purchase. Having a high-visibility ad almost doubles the effect of exposure on the proportion of respondents who report themselves very likely to purchase. Column (4) measures the mean effect of exposure. The coefficient suggests that exposure to a single ad, relative to a mean of 0.20, increases purchase intent by 3.6%. This gives us a baseline that reflects the typical metric that advertisers use when evaluating the success of a campaign (or lack of it), which we use in our calculations of the economic importance of our results.

Table 2: Influence of High Visibility and Contextual Targeting on Purchase Intent

	(1)		(2)		(3)		(4)	
	Coefficient (Std. Error)	p-value	Coefficient (Std. Error)	p-value	Coefficient (Std. Error)	p-value	Coefficient (Std. Error)	p-value
Exposed	0.00473^{***} (0.00110)	0.000	0.00565^{***} (0.00104)	0.000	0.00705^{***} (0.000814)	0.000	0.00745^{***} (0.000759)	0.000
Exposed × Context Ad	0.00941^{***} (0.00292)	0.001			0.00384^{*} (0.00215)	0.073		
Exposed × High Visibility Ad	0.00547^{***} (0.00161)	0.001	0.00421^{***} (0.00150)	0.005				
Exposed × Context Ad × High Visibility Ad	-0.0124^{***} (0.00428)	0.004						
Female	0.0116^{***} (0.00119)	0.000	0.0116^{***} (0.00119)	0.000	0.0116^{***} (0.00119)	0.000	0.0116^{***} (0.00119)	0.000
Hours on Internet (standardized)	0.0113^{***} (0.000344)	0.000	0.0113^{***} (0.000344)	0.000	0.0113^{***} (0.000344)	0.000	0.0113^{***} (0.000344)	0.000
Income (Standardized)	-0.00194^{***} (0.000406)	0.000	-0.00194^{***} (0.000406)	0.000	-0.00194^{***} (0.000407)	0.000	-0.00194^{***} (0.000407)	0.000
Age (Standardized)	-0.00957^{***} (0.000568)	0.000	-0.00957^{***} (0.000569)	0.000	-0.00957^{***} (0.000569)	0.000	-0.00957^{***} (0.000569)	0.000
Observations	2464812		2464812		2464812		2464812	
Log-Likelihood	-1062349		-1062357		-1062361		-1062363	
Variance captured by Fixed Effects	0.139		0.139		0.139		0.139	
R-Squared	0.141		0.141		0.141		0.141	

Fixed effects at ad-site level; Robust standard errors clustered at ad-site level, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

The fit of these regressions (measured by R-squared) is 0.141, with little difference across the four specifications. Given heterogeneous tastes for the products being advertised, we do not view the overall level of fit as surprising. Furthermore, most of this is explained by the fixed effects, with just 0.002 explained by our covariates for ad exposure and type. Again, we do not view it as surprising that seeing a banner ad once explains only a small fraction of the variance in purchase intent for the product. If it explained much more, we would expect the price of such advertisements to be much larger than 0.2 cents per view. Nevertheless, as we detail in section 4.3 below, our estimates still have important economic implications for the online advertising industry. It is not overall fit at the individual level that matters, but the marginal benefit relative to the marginal cost.

4.2 Robustness Checks

We checked the robustness of the negative interaction effect that we find in column (4) of Table 2 to many alternative specifications. Table 3 displays the results. Column (1) shows a logit specification to check that our linear probability specification had not influenced the results. Limitations of computing power meant that we had to take a 5% sample of the original data in order to be able to estimate a logit specification with the full set of fixed effects. Even with this smaller sample, the results are similar in relative size and magnitude.

One issue of a logit probability model compared to a linear probability model is that the interpretation of interaction terms in logit and probit models is not straightforward (Ai and Norton 2003), as they are a cross-derivative of the expected value of the dependent variable. Specifically, for any nonlinear model where $E[y|x_1, x_2] = F(\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + Z\gamma) = F(.)$, Ai and Norton point out that the interaction effect is the cross-derivative of the expected value of y: $\frac{\partial^2 F(.)}{\partial x_1 \partial x_2} = \beta_1 F'(.) + (\beta_1 x_1 + \beta_{12} x_1 x_2)(\beta_2 x_2 + \beta_{12} x_1 x_2) F''(.)$. The sign of this marginal effect is not necessarily the same as the sign of the coefficient β_{12}. We therefore used the appropriate formula for three-way interactions to calculate the marginal effect in this setting. The marginal effects at the means of the estimation sample were 0.008 (p-value=0.312) for Exposed × Context Ad; 0.010 (p-value=0.015) for Exposed × High Visibility Ad; and -0.037 (p-value=0.024) for Exposed × Context Ad × High Visibility Ad. These estimates for our key interaction and the effect of visibility were therefore larger than those reported in Table 1, though the coefficient estimate for contextual targeting is no longer significant. This robustness check strongly supports our core finding that contextually targeted ads that are highly visible are less effective than ads that are *either* contextually targeted or highly visible but not both.

We also use the logit results to conduct an out-of-sample prediction on the remaining 95% of the data. Our model correctly predicts 61.4% of successful campaigns (defined as ad campaigns on a specific website where the average purchase intent for the treatment group is higher than the control group). This is significantly better than the 52% of successful campaigns that we would predict by chance using random assignment.[6]

In column (2), we use the full (1 to 5) scale of the dependent variable rather than discretizing it. Our results are robust. In the appendix, we show robustness of all results to this specification. In column (3), we show that our results do not change if we exclude demographic controls for gender, age, income, and internet usage. This provides supporting evidence that the experiment randomized the treatment between such groups and that unobserved heterogeneity

[6] Since about 60% of campaigns are successful, the hit rate under random assignment is 52%, i.e., (40% *x* 40%) + (60% *x* 60%.)

between control and exposed groups does not drive our results. Column (4) shows robustness to a non-parametric specification for the gender, age, income, and internet usage controls, where we include a different fixed effect for different values for age, income, and internet usage. This also allows us to include fixed effects for missing values or intentionally non-reported values of these controls.

In column (5), we see whether our results are echoed in another measure of ad effectiveness: Whether or not the respondent expressed a favorable opinion of the product. The results in column (5) are similar to those previously reported, though the point estimate for the effect of a contextually-targeted ad is slightly smaller. This may be because contextually-targeted advertising is more successful at influencing intended actions than at influencing opinions (Rutz and Bucklin 2009). In column (6), we show that our results are robust to our definition of ad visibility. One concern with including pop-up ads is the development of pop-up blockers as a feature of internet browsers. This could influence our results, for example, if people who browsed sites with specific content were more likely to have pop-up blockers. However, the results in column (6) show the same pattern as before, even when we exclude such ads.

In Table 3 in column (7), we report results for all respondents, including those who due to the randomized nature of the ad server saw the ad more than once. We excluded these 730,000 respondents from the data we use to report our main specifications in order to simplify interpretation of what ‘exposed’ means in our setting. The results are similar to before, though the effect of baseline exposure is slightly higher.

Table 3: Robustness Checks

Dependent Variable: Respondent very likely to purchase

	(1) Logit+	(2) Use full scale—i.e. treat rating scale as interval	(3) No Controls	(4) Non-Parametric Controls	(5) Favorable Opinion	(6) No Pop-Ups	(7) Multiply Treated	(8) Only ads that appear on both matched and unmatched context websites
Exposed	0.0788***	0.0220***	0.00492***	0.00513***	0.00573***	0.00479***	0.00582***	0.00451***
	(0.0268)	(0.00512)	(0.00107)	(0.00112)	(0.00103)	(0.00110)	(0.000983)	(0.00107)
Exposed × Context	0.177**	0.0455***	0.00931***	0.00948***	0.00623**	0.00921***	0.0106***	0.0102***
	(0.0863)	(0.0109)	(0.00290)	(0.00293)	(0.00302)	(0.00290)	(0.00282)	(0.00330)
Exposed × High Visibility	0.0861**	0.0263***	0.00566***	0.00534***	0.00579***	0.00536***	0.00514***	0.00510***
	(0.0401)	(0.00679)	(0.00159)	(0.00161)	(0.00160)	(0.00161)	(0.00143)	(0.00163)
Exposed × Context × High Visibility	-0.365***	-0.0560***	-0.0125***	-0.0123***	-0.0118***	-0.0120***	-0.0130***	-0.0144***
	(0.125)	(0.0163)	(0.00427)	(0.00428)	(0.00451)	(0.00428)	(0.00395)	(0.00471)
Female	0.0612***	0.0146***			0.0136***	0.0116***	0.0117***	0.0118***
	(0.0196)	(0.00519)			(0.00116)	(0.00119)	(0.00111)	(0.00112)
Hours on Internet (standardized)	0.0763***	0.0414***			0.0126***	0.0113***	0.0110***	0.0112***
	(0.00964)	(0.00123)			(0.000375)	(0.000344)	(0.000319)	(0.000366)
Income (standardized)	-0.0253**	-0.0366***			-0.00170***	-0.00194***	-0.00150***	-0.00199***
	(0.0100)	(0.00169)			(0.000476)	(0.000406)	(0.000372)	(0.000439)
Age (standardized)	-0.0671***	-0.0880***			-0.00435***	-0.00957***	-0.00893***	-0.00957***
	(0.00952)	(0.00259)			(0.000605)	(0.000568)	(0.000557)	(0.000569)
Age, Income, Internet Use Fixed Effects	No	No	No	Yes	No	No	No	No
Observations	102414	2464812	2464812	2464812	2443939	2464812	3196474	1943489
Log-Likelihood	-40251.8	-4167674.1	-1063992.6	-1061035.0	-1193844.2	-1062349.5	-1380581.9	-856614.7
R-Squared	N/A	0.200	0.140	0.142	0.136	0.141	0.142	0.140

Fixed effects at ad-site level; Robust standard errors, clustered at ad-site level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

+Based on a 5% sample of data to overcome computational limitations imposed by estimating fixed 13,000 fixed effects.

While the individual-level randomization of exposure to ads addresses the usual concerns about unobserved heterogeneity at the individual level, our result for this negative interaction between high visibility and contextual targeting relies on the assumption that contextually-targeted highly visible ads are not of worse quality than either highly visible ads that are not contextually-targeted or contextually-targeted ads that are not highly visible. This therefore assumes that the endogeneity of advertising quality will not systematically affect our core result. In Table 3 column (8), we explore the appropriateness of this assumption in our context. We look only at ads that were run on more than one website and where the websites sometimes were contextually targeted and sometimes not. Our core results do not change. Therefore, holding campaign quality fixed, we still observe substitution between visibility and contextual targeting.

4.3 Economic implications

Section 4.2 establishes that our results are statistically robust to many specifications, but we have yet not established that they are economically meaningful. First, we note that column (1) of Table 2 suggests that seeing one plain banner ad once increases purchase intention by 0.473 percentage points. Relative to a price of 0.2 cents per view, this means that online advertising pays off if increasing purchase intention to "very likely to purchase" is worth 42 cents to the firm (95% confidence interval 29 to 78 cents). Therefore, while the coefficient may seem small, it suggests an economically important impact of online advertising. The effects of targeted ads and obtrusive ads are even larger, though the prices of such ads are also higher.

We combine information from our data with various industry sources to develop "back-of-the-envelope" calculations on the total magnitude of wasted advertising spending due to ads that are both targeted and obtrusive. These estimates are all clearly rough; the purpose of this subsection is simply to give a sense of the order of magnitude of the importance of our results.

In order to conduct this calculation, we require estimates of (a) the total size of online display advertising spending, (b) the percentage of these campaigns that are both targeted and obtrusive, (c) the cost of the targeted and obtrusive ads relative to plain banner ads, and (d) the effectiveness of targeted and obtrusive ads relative to plain banner ads.

For (a) "total ad spending", we use Khan (2009)'s estimate that US firms spent $8.3 billion on online display advertising in 2009. For (b), we rely on the fact that in the data we study 6.4% of the campaigns are both targeted and obtrusive. Our dataset is the most commonly used source in industry for information on trends in online display advertising. For (c), we use our results on the relative effectiveness of plain banner ads and targeted and obtrusive ads in columns (2) and (3) of Table 2 to derive estimates of their relative costs. We use these estimates rather than actual industry costs because estimates vary widely about the relative cost of different types of ads (as do the list prices websites provide to advertisers).[7] Our estimates suggest that advertisers should pay 74% more for high visibility ads (95% confidence interval 16%-197%) and 54% (95% confidence interval -6%-147%) more for context-based ads. We conservatively assume that ads that are both targeted and highly visible have the same premium as highly visible ads (74%) relative to plain banner ads.[8] For (d), we use our estimates from Table 2 to generate the relative effectiveness of targeted and obtrusive ads relative to plain banner ads.

Combining these data suggests that 8% (95% confidence interval 7.9-9.2%) of total ad spending, or $664 million, is being spent on campaigns that combine high visibility and

[7] For example, Khan (2009) estimates the typical cost of a plain banner ad at $2 per 1000 impressions or 0.2 cents per impression in 2009. Targeted ads and visible ads can cost much more. Confidential estimates from a large media company report that banner prices on web properties that allow contextual targeting are priced 10 times higher than regular "run of the network" ads (Krauss et al 2008). Similarly, there are industry reports that basic video ads cost four times as much as regular ads; premium ads cost as much as 18 times regular banner ads (Palumbo 2009).
[8] Advertisers appear to pay a *premium* for ads that are both contextually targeted and highly visible. For example, technologyreview.com, a website owned by MIT to distribute its research on technology, charges a price that is already 35 times higher than average for 1000 impressions because it can attract technology advertisers who want to advertise to technology professionals. Then, it charges a further 50% premium for advertisers who want to use video or audio ads.

targeting. Because these ads are no more effective than standard banners, if advertisers replaced redundantly targeted and visible ads with cheaper standard banner ads they could cut ad spending by 5.3 percent (95% confidence interval 3.5%-7.4%) without affecting ad performance.

There are some obvious caveats to the validity of this number. Using the estimates in Table 2 to calculate (c), the implied price premiums have the advantage of allowing us to provide a range of statistical confidence, but these estimated price premiums are likely too low given posted industry prices. Therefore, it is likely that we understate the cost savings. Second, while our data are collected by a major advertising company on behalf of its clients and are the one of the most used data sources for evaluating campaign performance in industry, it is not certain that they are representative of the display advertising sector in general when we use these data to project the fraction of campaigns that are targeted and obtrusive. This could bias the results in either direction. Therefore, despite reporting confidence intervals, validity concerns may overwhelm the reliability measures.

4.4 Underlying Mechanism

Having established the robustness of the result, we then explore the underlying mechanism. We first rule out differences in ad recall. Specifically, one potential explanation for why highly visible ads and contextually-targeted ads work poorly together is that highly visible ads are less successful at distinguishing themselves if they appear next to similar content. For example, a non-user-initiated video ad that is placed on a movie website may be less noticeable there than if it is placed on a cooking site.

Table 4 reports the results of a specification similar to that in Table 2, but where the dependent variable is whether or not the survey respondent said they could recall seeing the ad on the website. Column (1) suggests that the effect of being exposed to the ad is larger for this measure than in Table 2. This is unsurprising, as the effect is more direct; there is a 20 percent increase from the baseline proportion remembering seeing the ad for people who saw the ad, as compared to those who did not. Column (2) suggests that if an ad's content matches the website's content, it is less likely to be recalled. So, contextually-targeted ads are associated with higher purchase intent for the product, but people who see contextually-targeted ads are less likely to recall those ads than are people who see other kinds of ads. This is consistent with findings in the product placement literature (Russell 2002) that contextually appropriate ads can blend into the content, reducing recall but increasing purchase intent. Column (3) is consistent with our prior results: Visible ads are more likely to be recalled. Column (4) suggests that there is no significant negative interaction between contextual targeting and high visibility for recall.

The difference in these results from Table 2 suggests that the combination of high visibility and contextual targeting does not predominantly affect recall, but instead affects how effective the ad is at persuading the consumer to change their purchase intent. Indeed, making contextually-targeted ads too ‘visible’ seems to drive a wedge between the advertiser and the consumer. One interpretation of this result is that contextually-targeted ads are not as directly noticeable, and hence are viewed more favorably as background information; however, when contextually-targeted ads are highly visible, they no longer blend into the underlying website.

Table 4: Influence of High Visibility and Contextual Targeting on Ad Recall

Dependent Variable: Respondent recalls seeing the advertisement

	(1)	(2)	(3)	(4)
Exposed	0.0458***	0.0465***	0.0410***	0.0412***
	(0.00116)	(0.00126)	(0.00160)	(0.00173)
Exposed × Context		-0.00665**		-0.00219
		(0.00298)		(0.00409)
Exposed × High Visibility			0.0104***	0.0115***
			(0.00233)	(0.00253)
Exposed × Context × High Visibility				-0.00972
				(0.00597)
Female	-0.0230***	-0.0230***	-0.0230***	-0.0230***
	(0.00114)	(0.00114)	(0.00114)	(0.00114)
Hours on Internet (standardized)	0.0230***	0.0230***	0.0230***	0.0230***
	(0.000425)	(0.000425)	(0.000425)	(0.000425)
Income (standardized)	-0.00260***	-0.00260***	-0.00260***	-0.00260***
	(0.000427)	(0.000427)	(0.000427)	(0.000427)
Age (standardized)	-0.0131***	-0.0131***	-0.0131***	-0.0131***
	(0.000586)	(0.000588)	(0.000586)	(0.000588)
Observations	2481592	2481592	2481592	2481592
Log-Likelihood	-1307660.5	-1307699.2	-1307704.8	-1307660.5
R-Squared	0.123	0.123	0.123	0.123

Fixed effects at ad-site level; Robust standard errors, clustered at ad-site level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We next explore what drives the negative interaction between contextual targeting and high visibility when purchase intent is the dependent variable. We show that privacy is an important moderator, both in terms of whether the respondent is particularly sensitive to intrusive behavior, and in terms of the nature of the product itself. This is consistent with highly visible ads increasing consumer inferences of manipulative intent in targeted advertising (Campbell 1995; Friestad and Wright 1994; Kirmani and Zhu 2007).

We stratify our results by whether or not the respondent checked the box “I prefer not to answer that question” when asked about income. As documented by Turrell (2000), people who refuse to answer questions on income usually do so because of concerns about privacy. A comparison between column (1) and column (2) of Table 5 shows that survey respondents who do not respond to an intrusive question also display stronger substitution between contextually-targeted and highly visible ads than other respondents. When contextual targeting is highly visible, these respondents react negatively to the ads. Interestingly, privacy-minded people do not react differently than others to ads that are highly visible or contextually-targeted but not both. The estimates in column (2) suggest that these privacy-minded people respond positively to high visibility ads so long as they are not also contextually-targeted. To assess the relative significance of the estimates in these two columns, we also ran a specification where we pooled the sample and included a four-way interaction between Exposed × Context Ad × High Visibility and whether the person kept their income secret. The results are reported in full in Table A4. This four-way interaction suggested that indeed there was a large and statistically significant additional negative interaction for customers who kept their income secret.

In columns (3) through (6), we explored whether this potential role for intrusiveness as an underlying mechanism was also echoed in the kind of products that were advertised. We looked to see whether the effect was stronger or weaker in categories which are generally considered to be personal or private.

Table 5: Privacy-sensitivity is associated with stronger substitution between contextual targeting and high visibility

Dependent Variable: Respondent very likely to purchase

	(1) Does not Reveal Income	(2) Reveals Income	(3) Private	(4) Not Private	(5) Health CPG	(6) Other CPG
Exposed	0.00391* (0.00206)	0.00505*** (0.00116)	0.00184 (0.00288)	0.00517*** (0.00119)	-0.00480 (0.00522)	0.00840** (0.00344)
Exposed × Context	0.0102 (0.00669)	0.00916*** (0.00311)	0.0135** (0.00561)	0.00857** (0.00344)	0.0423*** (0.0100)	0.0135 (0.0138)
Exposed × High Visibility	0.00643** (0.00328)	0.00502*** (0.00168)	0.0102** (0.00464)	0.00480*** (0.00172)	0.0177** (0.00795)	0.0108** (0.00517)
Exposed × Context × High Visibility	-0.0214** (0.00932)	-0.0101** (0.00466)	-0.0274*** (0.00968)	-0.00979** (0.00484)	-0.0788*** (0.0186)	-0.0168 (0.0181)
Female	0.0128*** (0.00181)	0.0123*** (0.00124)	0.00922** (0.00367)	0.0119*** (0.00125)	0.0343*** (0.00641)	0.0652*** (0.00419)
Hours on Internet (standardized)	0.00750*** (0.000716)	0.0119*** (0.000376)	0.0101*** (0.000946)	0.0115*** (0.000369)	0.0125*** (0.00160)	0.0138*** (0.00124)
Income (standardized)	N/A	-0.00181*** (0.000407)	-0.00126 (0.000944)	-0.00201*** (0.000444)	-0.00609*** (0.00166)	-0.00644*** (0.00142)
Age (standardized)	-0.0112*** (0.000916)	-0.00910*** (0.000568)	-0.0120*** (0.00146)	-0.00923*** (0.000615)	-0.0195*** (0.00280)	0.00196 (0.00239)
Coefficient on four-way interaction in alternative specification	-0.0369*** (0.00919)		-0.0183* (0.0109)		-0.0677*** (0.0258)	
Observations	390608	2074204	325376	2139436	114922	192609
Log-Likelihood	-152023.4	-903402.0	-119800.1	-941137.4	-57285.5	-102271.0
R-Squared	0.157	0.143	0.145	0.139	0.139	0.146

Fixed effects at ad-site level; Robust standard errors, clustered at ad-site level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Full results for four-way interaction specification on pooled data used to assess significance reported in Table A4.

In columns (3) and (4), we looked across all categories and identified several that are clearly related to privacy. Specifically, these categories are Banking, Health, Health Services, OTC Medications, Insurance, Investment, Mutual Funds, Retirement Funds, Loans, and Other Financial Services. We identified health and financial products as categories where there may be privacy concerns on the basis of two factors. First, Tsai et al (2010) offer survey evidence of customers about which products they considered private, and confirm that health and financial information are categories where privacy is particularly important. Second, actual privacy policies tend to specify health and financial products as privacy-related. For example, Google's privacy policy states that "Google will not associate sensitive interest categories with your browser (such as those based on race, religion, sexual orientation, *health, or sensitive financial categories*), and will not use such categories when showing you interest-based ads." Column (3) includes the categories where privacy concerns are readily apparent; column (4) includes all others. Substitution between visibility and contextual targeting is highest for categories related to privacy. We estimated a four-way interaction specification that indicated that this incremental negative effect is significant at the 10% level.

As noted in the information security and privacy literature (e.g. Kumaraguru and Cranor's (2005) extensive survey results), consumers consider health information and use of medications to be especially personal information, and something that should be protected online. This is also echoed in the special protections given to health information in the European Union data privacy law 2002/EC/58. In columns (5) and (6), we focus on health and look for differences within a single category, in order to compare health-related products to other similar products. Specifically, we compare estimates for over-the-counter medications to food-based consumer packaged goods categories. The results again suggest that the substitution between contextual targeting and high visibility was significantly larger for the privacy-related products. The incremental negative effect is evaluated to be significant at the 1% level in a four-way interaction specification.

5. Interpretation and Conclusion

Our study of 2,892 online display advertising campaigns across a variety of categories and website yields three core conclusions:

First, we find that while obtrusive (or highly visible) online advertising and context-based online advertising work relatively well independently, they appear to fail when combined. This result is more pronounced in categories of products likely to be more private, and for people who seem to guard their privacy more closely. This suggests that the ineffectiveness of combining contextual targeting and obtrusiveness in advertising is driven by consumers' perceptions of privacy. Consumers may be willing to tolerate contextually-targeted ads more than other ads because they potentially provide information; however, making such ads obtrusive in nature may increase perceptions of manipulation (Campbell 1995). When privacy concerns are salient, a prevention focus may increase sensitivity to manipulative intent of the targeted ads (Kirmani and Zhu 2007). This suggests a role for privacy concerns in models that optimize advertising content in data-rich environments by minimizing viewer avoidance (e.g. Kempe and Wilbur 2009).

Second, the result that contextually-targeted ads work at least as well if they do not have features designed to enhance their visibility can help to explain the unexpected success of products such as AdSense by Google. It also explains why online advertising is becoming increasingly divided between plain text-based highly targeted ads, and more visually striking but less targeted ads. The results also suggest more generally the importance of more nuanced empirical models for advertising success. Not only is it important to separately model how ads affect awareness and preferences, but is also important to include controls for the features and placement of ads in such specifications.

Third, our findings on privacy and intrusiveness have policy implications for the direction of government policy. There is mounting pressure in the US and Europe to regulate the use of data on browsing behavior to target advertising (Shatz 2009). Our research suggests that regulators may need to consider a potential trade-off from such regulation. If regulation is successful at reducing the amount of targeting that a firm can do by using a customer's browsing behavior, then firms may find it optimal to invest instead in highly obtrusive ads. Customers may dislike having data collected about their browsing behavior, but they have also expressed dislike of highly visible ads in surveys (Chan, Dobb, and Stevens 2004). Therefore, regulators should weigh these two potential sources of customer resistance towards advertising against each other.

As with any empirical work, this paper has a number of limitations which present opportunities for future research. First, we rely on stated expressions of purchase intent and not actual purchase data. It is possible that the type of advertising used may have a different effect on actual purchases. Second, we present evidence that suggests that the negative effect of combining both visibility and targeting is more pronounced in situations where privacy is important, but there are still further questions about what other triggers of privacy concerns there may be for ad viewers. In general in academic marketing research, there have been few studies about how customer privacy concerns are triggered and what the implications are for firms. This study highlights the need for a better understanding of the behavioral processes that generate consumer privacy concerns. Third, we look only at targeting of ads based on the customer's current website browsing patterns. We do not look at reactions to ads that are targeted on the basis of past browsing behavior, though these ads have become very controversial in recent discussions at the Federal Communications Commission. Given the growing commercial importance of behavioral targeting, this is an area where academic research could potentially help answer questions both about the usefulness of such ads and how customer perceive and react to being targeted by ads in this way.

More generally, we think there are several promising avenues for research to build on our findings. First, while there is a formal theoretical literature that discusses behavioral targeting of pricing and its implications for privacy (e.g. Acquisti and Varian 2005; Chen, Narasimhan, and Zhang 2001; Fudenberg and Villas-Boas 2006; Hermalin and Katz 2006; Hui and Png 2006), the literature on behavioral targeting of advertising remains sparse. What are the benefits of such targeting? What are the costs? How can we formalize the concept of privacy in the context of behaviorally targeted advertising? Second, it would be interesting to explore how our findings on privacy apply to websites where visitors reveal detailed personal information, such as social networking websites. Are visitors to such sites more aware of manipulation attempts? Is there a way to leverage the social network and overcome perceptions of manipulative intent? This latter question may apply to all online advertising settings. Once we have a richer formal theoretical framework for understanding how behavioral targeting and privacy concerns interact, we will develop a better understanding of when it is appropriate to target the increasingly visible ads generated by advances in online advertising design.

References

Aaker, David A., V. Kumar, and George S. Day (2004). *Marketing Research: Eighth Edition*. Wiley: New York.

Acquisti, Alessandro, and Spiekermann, S. (2009). Do Pop-ups Pay Off? Economic effects of attention-consuming advertising. *Working paper, Carnegie Mellon University*.

Ai, C., and E. Norton. (2003). Interaction terms in logit and probit models. *Economic Letters* 80(1), 123-129.

Angrist, Joshua D. and Jörn-Steffen Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press: Princeton NJ.

Bemmaor, Albert C (1995). "Predicting Behavior from Intention-to-Buy Measures: The Parametric Case" *Journal of Marketing Research*, Vol. 32, No. 2, pp. 176-191.

Bentler, P.M., and Chih-Ping Chou (1987). Practical Issues in Structural Modeling. *Sociological Methods and Research* 16(1), 78-117.

Campbell, Margaret C., and Amna Kirmani (2000). Consumers' Use of Persuasion Knowledge: The Effects of Accessibility and Cognitive Capacity on Perceptions of an Influence Agent. *Journal of Consumer Research* 27(1), 69-83.

Chen, Yuxin, Chakravarthi Narasimhan, and Z. John Zhang. 2001. Individual Marketing with Imperfect Targetability. *Marketing Science*, 20(1), 23-41.

Cho, C.-H., and Cheon, H. J. (2003). Why do people avoid advertising on the internet? *Journal of Advertising*, 33(3), p. 89-97.

Clark, Robert, Ulrich Doraszelski, and Michaela Draganska. (2009). The Effect of Advertising on Brand Awareness and Perceived Quality: An Empirical Investigation Using Panel Data. *Quantitative Marketing and Economics* 7(2), 207-236.

Cole, Sally G., Leah Spalding, and Amy Fayer. 2009. The Brand Value of Rich Media and Video Ads. DoubleClick Research Report, June.

Drèze, X., and Hussherr, F.-X. (2004). Internet Advertising: Is Anybody Watching? *Journal of Interactive Marketing* , 8-23.

Edwards, S. M., Li, H., and Less, J.-H. (2002). Forced Exposure and Psychological Reactance: Antecedents and Consequences of the Perceived Intrusiveness of Pop-Up Ads. *Journal of Advertising, 31* (3), 83-95.

Fink, Arlene. (2009). *How to Conduct Surveys: A Step-by-Step Guide, 4th edition*. SAGE Publications: Thousand Oaks, CA.

Friestad, Marian and Wright, Peter (1994). The Persuasion Knowledge Model: How People Cope with Persuasion Attempts. *Journal of Consumer Research* 21(1), 1-31.

Fudenberg, Drew, and J. Miguel Villas-Boas (2006). Behavior-based price discrimination and customer recognition. *Handbooks in Information Systems, Vol. 1: Economics and Information Systems,* Ed. Terrence Hendershott, p. 377-436.

Ghose, Anindya and Sha Yang (2009). An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets. *Management Science* 55(10), 1605-1622.

Godes, David, and Mayzlin, Dina (2009). Firm-Created Word-of-Mouth Communication: Evidence from a Field Test. *Marketing Science* 28: 721-739.

Goldfarb, A., and Tucker, C. (2009). Search Engine Advertising: Pricing Ads to Context. *Working Paper, MIT.*

Hallerman, David (2009). “US Advertising Spending: The New Reality”, April, Emarketer Technical Report.

Hermalin, Benjamin E. and Michael L. Katz. (2006). Privacy, Property Rights & Efficiency: The Economics of Privacy as Secrecy. *Quantitative Marketing and Economics*, 4(3) 209-239.

Hollis, Nigel. (2005). Ten Years of Learning on How Online Advertising Builds Brands. *Journal of Advertising Research* 45(2), 255-268.

Horrace, William C., and Ronald L. Oaxaca. 2006. Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economics Letters* 90, 321-327.

Hui, Kai-Lung and I.P.L Png (2006). The Economics of Privacy. *Handbooks in Information Systems, Vol. 1: Economics and Information Systems*, Ed. Terrence Hendershott, 471-497.

Johnson, D.R., and Creech, J.C. (1983), Ordinal measures in multiple indicator models: a simulation study of categorization error, *American Sociological Review*, 48, 398-407.

Kempe, David, and Kenneth C. Wilbur (2009). What can Television Networks Lean from Search Engines? How to Select, Order, and Price Ads to Maximize Advertiser Welfare. *Working Paper, Duke University*.

Khan, Imran (2009). Nothing But Net: 2009 Internet Investment Guide. *Technical Report, JP Morgan.*

Kirmani, Amna and Rui (Juliet) Zhu (2007). Vigilant Against Manipulation: The Effect of Regulatory Focus on the Use of Persuasion Knowledge," *Journal of Marketing Research,* 44(4), 688-701.

Kline, Theresa J.B. 2005. *Psychological Testing: A Practical Approach to Design and Evaluation*. SAGE Publications, Thousand Oaks, CA.

Krauss, Cheryl, Marla Nitke and Frank Pinto (2008) Bain/IAB Digital Pricing Research. *Technical Report*, August.

Kumaraguru, Ponnurangam, and Lorrie Faith Cranor (2005). Privacy Indexes: A Survey of Westin's Studies. *Working Paper CMU-ISRI-5-138, Carnegie Mellon University.*

Levinsohn, James, and Amil Petrin (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *Review of Economic Studies* 70(2), 317-341.

Lewis, R., and Reiley, D. (2009). Retail Advertising Works! Measuring the Effects of Advertising on Sales via a Controlled Experiment on Yahoo! *Yahoo Research Technical Report.*

Malhotra, Naresh K. (2007). *Marketing Research: An Applied Orientation, Fifth Edition*. Pearson Education Inc.: Upper Saddle River, NJ.

Morwitz, Vicki G. Joel H. Steckel, and Alok Gupta, (2007). When do purchase intentions predict sales? *International Journal of Forecasting*, 23(3), 347-364.

Palumbo, Paul (2009). "Online Video Media Spend: 2003 - 2010 ". AccuStream iMedia Research technical report, Jan 2009.

Russell, Cristel Antonia (2002). Investigating the Effectiveness of Product Placements in Television Shows: The Role of Modality and Plot Connection Congruence on Brand Memory and Attitude. *Journal of Consumer Research* 29 (December), 306-318.

Rutz, O., and Bucklin, R. (2009). Does Banner Advertising Affect Browsing Paths: Clickstream Model says Yes, for some. *Working Paper, Yale University*.

Shatz, Amy (2009). Regulators Rethink Approach to Online Privacy. *Wall Street Journal*, August 5.

Tsai, Janice Serge Egelman, Lorrie Cranor, and Alessandro Acquisti (2010). The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study, forthcoming *Information Systems Research.*

Turow, Joseph, Jennifer King, Chris Jay Hoofnagle, Amy Bleakley, and Michael Hennessy (2009). Contrary to what marketers say, Americans reject tailored advertising and three activities that enable it. *Working paper, Annenberg School for Communication, University of Pennsylvania.* Accessed at http://graphics8.nytimes.com/packages/pdf/business/20090929-Tailored_ Advertising.pdf on October 1, 2009.

Turrell, Gavin (2000). Income non-reporting: Implications for health inequalities research. *Journal of Epidemiology and Community Health* 54, 207-214.

Van Noort, Guda, Peter Kerkhof, and Bob M. Fennis (2008). The Persuasiveness of Online Safety Cues: The Impact of Prevention Focus Compatibility of Web Content on Consumers' Risk Perceptions, Attitudes, and Intentions. *Journal of Interactive Marketing* 22(4), 58-72.

Wang, K., Chen, S.-H., and Chang, H.-L. (2008). The effects of forced ad exposure on the web. *Journal of Informatics & Electronics*, 27-38.

Wathieu, Luc, and Allan Friedman (2009). An Empirical Approach to Understanding Privacy Concerns. ESMT Working Paper # 09-001.

Wilbur, Kenneth C., Michelle Sovinsky Goeree, and Geert Ridder (2009). Effects of Advertising and Product Placement on Television Audiences. *Marshall Research Paper Series, Working Paper MKT 09-09.*

Appendix: Further Robustness Checks

Table A1: Replicates Table 2 with continuous scale
Dependent Variable: Purchase intent (scale of 1-5 where 5 is highest)

	(1)	(2)	(3)	(4)
Exposed	0.0220***	0.0265***	0.0332***	0.0353***
	(0.00512)	(0.00475)	(0.00355)	(0.00328)
Exposed × Context Ad	0.0455***		0.0204**	
	(0.0109)		(0.00816)	
Exposed × High Visibility Ad	0.0263***	0.0206***		
	(0.00679)	(0.00631)		
Exposed × Context × High Visibility	-0.0560***			
	(0.0163)			
Female	0.0146***	0.0146***	0.0147***	0.0147***
	(0.00519)	(0.00519)	(0.00519)	(0.00519)
Hours on Internet (standardized)	0.0414***	0.0414***	0.0414***	0.0414***
	(0.00123)	(0.00123)	(0.00123)	(0.00123)
Income (standardized)	-0.0366***	-0.0366***	-0.0366***	-0.0366***
	(0.00169)	(0.00169)	(0.00169)	(0.00169)
Age (standardized)	-0.0880***	-0.0880***	-0.0880***	-0.0880***
	(0.00259)	(0.00259)	(0.00260)	(0.00260)
Observations	2464812	2464812	2464812	2464812
Log-Likelihood	-4167674.1	-4167687.8	-4167696.0	-4167700.9
R-Squared	0.200	0.200	0.200	0.200

Fixed effects at ad-site level; Robust standard errors, clustered at ad-site level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2: Replicates Table 3 with continuous scale

Dependent Variable: Purchase intent (scale of 1-5 where 5 is highest)

	(1) No Demographic Controls	(2) Non-Parametric Controls	(3) Favorable Opinion	(4) No Pop-Ups	(5) Multiply Treated	(6) Only ads that appear on both matched and unmatched context websites
Exposed	0.0225***	0.0236***	0.0229***	0.0220***	0.0253***	0.0251***
	(0.00507)	(0.00299)	(0.00444)	(0.00509)	(0.00446)	(0.00509)
Exposed×Context Ad	0.0461***	0.0208***	0.0433***	0.0443***	0.0509***	0.0450***
	(0.0108)	(0.00784)	(0.0110)	(0.0108)	(0.0109)	(0.0118)
Exposed×High VisibilityAd	0.0253***	0.0134***	0.0255***	0.0265***	0.0227***	0.0228***
	(0.00674)	(0.00443)	(0.00602)	(0.00679)	(0.00597)	(0.00692)
Exposed×Context× HighVisibility	-0.0555***	-0.0371***	-0.0425***	-0.0537***	-0.0564***	-0.0577***
	(0.0162)	(0.0118)	(0.0161)	(0.0162)	(0.0152)	(0.0177)
Female		0.0370***	0.0173***	0.0146***	0.0152***	0.0146***
		(0.00305)	(0.00470)	(0.00519)	(0.00489)	(0.00436)
Hours on Internet (standardized)		0.0266***	0.0411***	0.0414***	0.0404***	0.0422***
		(0.00101)	(0.00114)	(0.00123)	(0.00114)	(0.00130)
Income (standardized)		-0.0154***	-0.0364***	-0.0366***	-0.0336***	-0.0363***
		(0.00133)	(0.00156)	(0.00169)	(0.00166)	(0.00177)
Age (standardized)		-0.0108***	-0.0869***	-0.0880***	-0.0864***	-0.0848***
		(0.00182)	(0.00233)	(0.00259)	(0.00250)	(0.00253)
Age, Income, Internet Use Fixed Effects	Yes	No	No	No	No	No
Observations	2464812	2443939	2932278	2464812	3196474	1976180
Log-Likelihood	-4165117.9	-3488010.7	-4958771.8	-4167674.3	-5409175.6	-3348518.8
R-Squared	0.202	0.156	0.202	0.200	0.202	0.200

Fixed effects at ad-site level; Robust standard errors, clustered at ad-site level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3: Replicates Table 5 with continuous scale

Dependent Variable: Purchase intent (scale of 1-5 where 5 is highest)

	(1) Not Reveal	(2) Reveals Income	(3) Private	(4) Not Private	(5) Health CPG	(6) Other CPG
Exposed	0.0187**	0.0233***	0.0257***	0.00472	-0.00343	0.0444***
	(0.00846)	(0.00510)	(0.00585)	(0.00856)	(0.0202)	(0.0117)
Exposed×Context Ad	0.0533**	0.0440***	0.0446***	0.0563***	0.117***	0.0705
	(0.0236)	(0.0114)	(0.0133)	(0.0174)	(0.0301)	(0.0633)
Exposed× High Visibility Ad	0.0357***	0.0234***	0.0276***	0.0185	0.0449	0.0275
	(0.0126)	(0.00688)	(0.00761)	(0.0138)	(0.0276)	(0.0181)
Exposed×Context× High Visibility	-0.0912***	-0.0473***	-0.0588***	-0.0505	-0.245***	-0.0552
	(0.0341)	(0.0174)	(0.0187)	(0.0350)	(0.0686)	(0.0731)
Female	0.0295***	0.0168***	0.0227***	-0.0230**	0.0755**	0.182***
	(0.00711)	(0.00545)	(0.00581)	(0.0114)	(0.0314)	(0.0131)
Hours on Internet (standardized)	0.0292***	0.0428***	0.0416***	0.0398***	0.0491***	0.0542***
	(0.00265)	(0.00133)	(0.00137)	(0.00284)	(0.00526)	(0.00452)
Income (standardized)	N/A	-0.0368***	-0.0379***	-0.0302***	-0.0452***	-0.0483***
		(0.00169)	(0.00192)	(0.00344)	(0.00653)	(0.00525)
Age (standardized)	-0.107***	-0.0830***	-0.0794***	-0.130***	-0.128***	-0.0458***
	(0.00399)	(0.00252)	(0.00289)	(0.00545)	(0.0120)	(0.00819)
Observations	390608	2074204	2038776	426036	114922	192609
Log-Likelihood	-655863.6	-3504391.6	-3466766.5	-699515.4	-195107.6	-335447.4
R-Squared	0.219	0.202	0.176	0.285	0.177	0.164

Fixed effects at ad-site level; Robust standard errors, clustered at ad-site level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Four-way interaction specification

Dependent Variable: Purchase Intent (Indicator for Very Likely to make a Purchase)

	(1) Does not reveal income	(2) Privacy-related category	(3) CPG only
Exposed	0.00882***	0.00515***	0.00860**
	(0.00113)	(0.00119)	(0.00345)
Exposed×Context Ad	0.00529*	0.00846**	0.0133
	(0.00314)	(0.00343)	(0.0137)
Exposed×High Visibility Ad	0.00112	0.00486***	0.0107**
	(0.00166)	(0.00171)	(0.00518)
Exposed×Context Ad×	-0.00631	-0.00958**	-0.0161
High Visibility Ad	(0.00465)	(0.00483)	(0.0181)
Female	0.0122***	0.0117***	0.0571***
	(0.00123)	(0.00123)	(0.00354)
Std. Weekly Hours on Internet	0.0111***	0.0113***	0.0137***
	(0.000355)	(0.000356)	(0.00103)
Std. Income	-0.00171***	-0.00182***	-0.00632***
	(0.000415)	(0.000418)	(0.00117)
Std. Age	-0.00959***	-0.00953***	-0.00401**
	(0.000588)	(0.000589)	(0.00192)
Exposed×Income Secret	-0.0253***		
	(0.00157)		
Exposed×Context Ad×Income Secret	0.0257***		
	(0.00641)		
Exposed×High Visibility	0.0266***		
Ad×Income Secret	(0.00284)		
Exposed×Context Ad×	**-0.0369*** **		
High Visiblity×Income Secret	**(0.00919)**		
Exposed×Private Site		-0.00353	
		(0.00324)	
Exposed×Context Ad×Private Site		0.00563	
		(0.00667)	
Exposed×High Visibility Ad×		0.00553	
Private Site		(0.00503)	
Exposed×Context Ad×		**-0.0183* **	
High Visiblity×Private Site		**(0.0109)**	
Exposed×Health CPG			-0.0135**
			(0.00661)
Exposed×Context Ad×Health CPG			0.0294*
			(0.0172)
Exposed×High Visibility Ad×			0.0111
Health CPG			(0.00991)
Exposed×Context Ad×			**-0.0677*** **
High Visibility× Health CPG			**(0.0258)**
Observations	2464812	2464812	290982
Log-Likelihood	-1061852.2	-1062304.4	-150500.1
R-Squared	0.141	0.141	0.148

Fixed effects at ad-site level; Robust standard errors, clustered at ad-site level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Numerous lower-order interactions not involving exposure included but not reported.