NET Institute*

www.NETinst.org

Working Paper #10-07

September 2010

Social Networks, Personalized Advertising, and Privacy Controls

Catherine Tucker
Sloan School of Business, MIT

Social Networks, Personalized Advertising, and Privacy Controls

Catherine Tucker*

September 29, 2010

*Catherine Tucker is Douglas Drane Career Development Professor in IT and Management and Assistant Professor of Marketing at MIT Sloan School of Business, MIT, Cambridge, MA; 617 252-1499; cetucker@MIT.EDU. I thank the NET Institute (www.NETinst.org) for financial support. I also thank Alessandro Acquisti, Avi Goldfarb, Alex Marthews, Martin Peitz, Ken Wilbur and seminar participants at the University of Mannheim for useful suggestions. All mistakes are mine alone.

Social Networks, Personalized Advertising, and Privacy Controls

Abstract

This paper investigates how internet users' perception of control over their personal information affects how likely they are to click on online advertising. The paper uses data from a randomized field experiment that examined the relative effectiveness of personalizing ad copy to mesh with existing personal information on a social networking website. The website gave users more control over their personally identifiable information in the middle of the field test. The website did not change how advertisers used anonymous data to target ads. After this policy change, users were twice as likely to click on personalized ads. There was no comparable change in the effectiveness of ads that did not signal that they used private information when targeting. The increase in effectiveness was larger for ads that used less commonly available private information to personalize their message. This suggests that giving users the perception of more control over their private information can be an effective strategy for advertising-supported websites.

Keywords: Privacy, Online Advertising, Social Networks

Introduction

Firms that try to use consumers' data for marketing purposes face a trade-off between consumers' expressed preferences for control over how their private information is used and shared, and the need to use this information to increase interest in the firm's product offering. This tension has been particularly pronounced on social networking websites like Facebook and MySpace. These websites have collated a huge amount of personal data from their users and offer advertisers unique, proprietary ad networks that push the boundaries of tailored advertising. To reassure customers about such uses of customer data, social-networking sites have pioneered new technologies that allow consumers explicit control over how much information about them is publicly available. In this paper, we assess how such strategies for giving customers control over their personally identifiable information affects the ability of such sites to support themselves through advertising.

This matters because social networks allow advertisers to go beyond simple targeting techniques and actually personalize ad content. If a nonprofit selects fans of Oprah Winfrey on Facebook as a fruitful pool of potential supporters, it can then include a reference to Oprah Winfrey in its personalized ad copy. However, research has shown that consumer privacy concerns affect the performance of targeted ads (Goldfarb and Tucker, 2010a). Consumers might see personalized ad content as more appealing and tied to their interests, but they might conversely see it as as 'not only creepy but off-putting' if they felt the firm had violated their privacy (Stone, 2010).[1] The paper studies the extent to which consumer perceptions of control over their private information affect the performance of this type of personalized advertising.

This research uses data from a randomized field experiment conducted by a US-based nonprofit to optimize their advertising campaigns on Facebook, a social networking website. These campaigns were shown to 1.2 million Facebook users. The nonprofit's aim was to raise awareness of its work improving education for women in East Africa. The nonprofit randomized whether they explicitly personalized the ad copy to match the user's profile. For example, sometimes the text of the ad made an appeal based on the consumer's self-expression of whether they "Liked" celebrities and writers who in the past had made statements supporting the education of girls in Africa or African female empowerment in general. On other occasions, the nonprofit used that information only to target the ad: i.e. they targeted the ad to people who liked a celebrity, but showed them only a generic ad that was not personalized.

In the middle of the field experiment, after Facebook had encountered much criticism for not allowing users enough control over their privacy, Facebook announced a large and well-publicized shift in their privacy policy. The aim was to reassure users about how their data were used, by giving them more control over their privacy settings. This change did not, however, affect how advertising was displayed and targeted, since the advertising platform used anonymous data. The nonprofit had not anticipated there would be such a change when it launched its field test of the ads. However, the fact that the change occurred mid-way through the field experiment is valuable for measuring the effect on advertising effectiveness of such a change in privacy policies, while circumventing the usual endogeneity issues.

We compare how well the personalized ads performed before and after this change in policy relative to targeted and also untargeted ads. Targeted but non-personalized ads provide a particularly useful comparison, because they were served to the same individuals as the personalized ads. The only difference was that the ad itself did not make clear what information (if any) was being used to target the ad. It is not clear how strengthening users' perceived control over their personal information would affect attempts by advertisers to use users' information to enhance their ad content. On the one hand, the shift might sensitize users to privacy concerns, and make advertisers who try to use such information more unpopular, even if the use of such information were effectively anonymous. On the other hand, behavioral research has emphasized the importance of consumer perceptions of control for mediating privacy concerns (Xu, 2007; Brandimarte et al., 2010). If users feel reassured that they have control over how the information about them is being used and that their personal information is secure, they may find such ads less off-putting.

We use data on how many users clicked through on each ad for a five-week period spanning the policy change. Empirical analysis of both campaign-level and individual-level click-through data suggests that personalized advertising was almost twice as effective at attracting users to the nonprofit's Facebook page than was targeted but non-personalized advertising, after the shift in Facebook policy which alleviated many users' privacy concerns. Before the policy change, when privacy concerns were highly salient due to very heavy press coverage, personalized advertising performed slightly less well than merely targeted ads. There was little change in the effectiveness of targeted advertising and non-targeted advertising over the period. This is to be expected, because such ads do not make clear to consumers whether their private information is being used to target, and the privacy settings policy change did not change how ads were served to users.

Identification comes from the assumption that there were no underlying changes in user behavior that coincided with the policy change but were not directly attributable to the policy change. To check robustness to this assumption, we both check that our results are robust to a shorter time frame and run a falsification check for an earlier period where there was no such policy change. We find no significant change in user behavior. We also check that there was no significant change in the user composition of Facebook or advertiser behavior during the period we study.

We then investigated the mechanism underlying our finding that consumers responded more positively to personalized advertising after the policy change. In particular, we evaluated whether consumers were concerned that advertisers, when using this information, were tracking the user's behavior in a potentially intrusive way. To evaluate this, we used data on the number of users who fell into each targeting category. There are almost two million fans of Oprah Winfrey on Facebook, but some targeting variables were unusual enough that their potential reach was only in the thousands. We found that personalization was relatively more effective for personalized ads in ad-targeting categories which used unusual information after privacy controls were enhanced. This provides suggestive evidence that indeed consumers were concerned that the information being used in the ads was simply too personally identifiable without a corresponding sense of control over their data.

These findings contribute at four levels.

To our knowledge, this is the first paper that studies how giving web users better control over how their private information is shared, affects advertising outcomes. The finding that there are positive effects for advertisers of addressing users' privacy concerns is potentially useful to advertising-supported websites, as currently there is much debate in the industry about the effects of addressing users' concerns. In a well-publicized survey, Turow et al. (2009) found that 66 percent of Americans do not want marketers to tailor advertisements to their interests. Fear of such resistance has led advertisers to limit their tailoring of ads (Lohr, 2010). However, our results suggest that there are benefits to the advertising-supported internet of reassuring users explicitly about how their private information is shared and used. Our results emphasize that the design of privacy settings can be a crucial strategic marketing variable.

On the academic side, this paper's focus on advertising complements research that has focused on more general questions of information sharing and privacy in social networks (Caverlee and Webb, 2008; Golder et al., 2007; Acquisti and Gross, 2006). Early research on privacy tended to simply describe privacy as a matter of giving users control over their data (Miller, 1971; Culnan, 1993; Smith et al., 1996). However, more recent research has challenged this and shown how individual-level control over information can mediate privacy concerns, even if access to the data remains unchanged. Xu (2007) shows in a lab-experiment that people have fewer privacy concerns when given explicit control over publication of location data. Brandimarte et al. (2010) show in three lab experiments that giving users control over how much of their private information is published paradoxically encourages them to reveal more personal and potentially sensitive information than they may do otherwise, even though people's actual access to the data in question did not change. This is consistent with research such as Spiekermann et al. (2001), which suggests that the very existence of privacy protection, even if not actually pertinent, can lull users into a perhaps false sense of security. This paradox is also clear in our own data. The mere fact of giving users control over personally identifiable information appears to reassure users about the use of anonymized information in advertising and makes them more amenable to it. More generally, our research also reflects the importance of trust in mediating privacy concerns from ad personalization, which has been documented in survey-based research by Chellappa and Sin (2005).

The paper also contributes to the online advertising literature. It appears that personalizing ads using user-disclosed information in the ad copy increases their appeal if accompanied by appropriate privacy controls. This was studied from a theoretical perspective by Anand and Shachar (2009), who pointed out that the signaling power of a targeted ad in the traditional ad-signaling framework (as laid out by Kihlstrom and Riordan (1984); Milgrom and Roberts (1986)), could be strengthened by personalizing the ad, making consumers more likely to assume there is a match between them and the product. This paper is also one of the first studies of advertising by an external firm on a social networking site. The majority of the empirical work on targeting and social networks has studied offline methods. For example, Manchanda et al. (2008) has studied the role of social networks and targeting offline. Previous studies in marketing about social networking sites, have questioned how such sites can use advertising to obtain members (Trusov et al., 2009), and also how makers of applications designed to be used on social networking sites can best advertise their products (Aral and Walker, 2010) through viral marketing.

There have been no studies, however, to the author's knowledge, that examine social network advertising by firms not directly related to the site who simply wish to exploit the unusual features of social networking sites as an attractive advertising channel. This is important, because social networking sites are attractive media venues that are growing rapidly in importance. They have a youthful and passionate following: The average Facebook user in the United States spent 6.5 hours on Facebook over the course of December 2009, which was more than twice as long as the next leading web brand (Nielsen, 2010). Facebook doubled its U.S. audience from 54.5 million visitors in December 2008 to 111.9 million visitors in December 2009, and now accounts for 7% of all time spent online in the U.S (Lipsman, 2010); worldwide, its membership passed 500 million in July 2010. However, social-networking websited have previously been perceived as being problematic venues for advertising because of extremely low click-through rates (Holahan, 2007). This research suggests that if consumers are reassured about their privacy, firms can use personalization of ads to generate higher click-through rates.

Finally, these findings are also important from a marketing policy perspective. The results shed light on the benefits of websites giving consumers clear control over how their information is used and accessed by advertisers. Currently proposed regulations governing behavioral advertising in the US are focused around the mechanics of how firms implement opt-in and opt-out use of cookies and other tracking devices (Corbin, 2010). Previous empirical research suggests that this approach, by limiting the use of data by firms, reduces ad effectiveness (Goldfarb and Tucker (2010b)). By contrast, the results in this paper suggest that when privacy regimes are primarily focused on letting customers choose how personally identifiable information about them is shared and used, there is no negative effect on advertising performance. The current staff-discussion draft of US privacy legislation proposed by Representatives Boucher and Stearns exempts individually managed preference profiles (P.17, Sec. 3(e)). This provision may be an important way of ensuring the advertising-supported internet can continue to thrive.

Institutions and Data

The Nonprofit

The nonprofit running the experiment provides educational scholarships in East Africa that enable bright girls from poor families to go to or stay in high school. Part of the nonprofit's mission involves explaining its work in Africa to US residents and also engaging their enthusiasm and support for its programs. In order to do this, the nonprofit set up a Facebook 'page' which explained its mission, and also allowed people who were interested to see photos, read stories and watch videos about the girls who had been helped by the program.

To attract people to become fans of its Facebook page, the nonprofit started advertising using Facebook's own advertising platform. Initially, they ran an untargeted ad campaign which displayed an ad in April 2010 to all users of Facebook that live in the US, and are 18 years and older. This campaign experienced a very low click-through rate. The disappointing nature of this campaign led them to want to see if they could engage further with their potential supporters by both targeting and personalizing ad content.

Randomized Campaign

The nonprofit designed two separate campaigns with two separate target populations. The aim of the campaign was to encourage users to click on the ad and become a fan of the nonprofit's website. The first target population were college graduates from small liberal arts colleges that had a reputation of emphasizing the benefits of education for the community. Facebook started as a college-based social network, so it explicitly facilitates the identification of such graduates, and most users indicate what educational institutions they have attended and whether they are a current student or a graduate.

The second target population were Facebook users who had expressed appreciation for celebrities and writers who in the past had made statements supporting the education of girls in Africa or African female empowerment in general. Examples could be Oprah Winfrey, who has set up a girls' school in South Africa, or Serena Williams, who was a supporter of 'Build African Schools.' There were 19 such celebrities in total.[2] Such Facebook users are identified by whether they mention they 'like' such a person in their likes or interests section on their Facebook profile.

However, it was unclear to the nonprofit whether they should also personalize the ad content that these users saw. They thought that personalization might improve their ad's appeal, but they also did not want their ad to be unattractively intrusive or make potential supporters feel that their privacy had been violated. In order to establish whether Facebook user data should be used merely to target ads, or should in addition be used to personalize the content of the advertising appeal, they decided to experiment with two different ad formats. Table 1 summarizes the different conditions used. In the personalized condition, the ad explicitly mentioned the targeting variable. In the targeted but non-personalized case, the ad was similar in content but did not explicitly mention the targeting variable. In each case the ad was accompanied by the same picture of a girl who had been helped by the program. Based on the work of Small and Verrochi (2009), this girl had a slightly mournful expression.

[Table 1 about here.]

In addition to these two campaigns, the nonprofit also continued to use as its baseline, a non-targeted campaign which reached out to all adult US Facebook users simultaneously. This provided an additional baseline control for advertising effectiveness over the course of the study. The text of this baseline and non-targeted ad read "Support [Charity Name]. Help girls in East Africa change their lives through education." This ad and the two targeted campaigns were restricted to Facebook users who live in the US, and were 18 years and older. The charity set a daily maximum spending cap on advertising campaigns that was significantly below the $250 a day maximum spending cap mandated by Facebook. It also agreed to pay at most $0.50 for each click produced by the different advertising campaigns.

Figure 1 is a screenshot which shows the Facebook interface used to design and target an ad. To preserve the anonymity of the nonprofit, it shows a mock ad for a marketing journal. On the right-hand side, there is a sample ad which is similar in format to the ad used in the tests, and gives an accurate representation of the relative size of text and photo in the actual ad. The lower panel shows how an advertiser would theoretically target people who are interested in online marketing and who also had a college degree in marketing. On the right-hand panel, Facebook offers an estimate of the potential ad-reach of such targeting - in this case just 380 people in the US. We use such ad-reach data in our subsequent regressions to explore the behavioral mechanism driving our results.

[Figure 1 about here.]

Facebook's ad server encourages the comparative testing of different types of ads within a campaign, but there are no guarantees that each ad will be shown to users in exactly the same proportions. For example, if an advertiser was concerned primarily with retaining equivalent statistical power, they might choose to show each of the advertising campaigns being tested 10,000 times every day. However, Facebook does not guarantee this because it is constrained by who is using Facebook at any one time and what advertising inventory it has to sell. For example, it is easier for Facebook's ad servers to identify 10,000 people who are fans of Oprah Winfrey than 10,000 people who are fans of an obscure 1960s feminist poet. We control for error this may have introduced in our regression analysis by controlling for both the type of targeting variable and the date. However, ultimately what is important for our key findings is that there was no change in how the nonprofit managed the field experiment before and after the policy change, in particular in how it handled personalized relative to targeted ads, and also that there was no change in the way that Facebook's ad servers delivered ads before and after the policy change (This is documented in the appendix).

Policy Change

What was unique and potentially valuable about this field experiment was that on May 24 2010 (after the field experiment was planned and initiated and the first data collected), Mark Zuckerberg, the CEO of Facebook, announced that the company would be simplifying and clarifying their privacy settings as well as rolling back some previous changes that had made Facebook users' information more public. Studying this change was not the purpose of the randomized field experiment, but it fortuitously presented a unique opportunity to study how a change in user privacy controls in social networking sites can change consumer responses to advertising, since the nonprofit tested the ads using the same randomization technique before and after the policy change.

[Table 2 about here.]

The background to this policy change was that Facebook had been heavily criticized because its privacy settings were very granular and difficult to access. For example, Bilton (2010) pointed out that the 5,850 words of Facebook's privacy policy were longer than the United States Constitution, and that users wanting to manage their privacy settings had to navigate through 50 settings with more than 170 options. As detailed by Table 2, Facebook had previously acted to reduce the amount of control users had over their data and had attracted bad publicity for doing so. As well as bad press, Facebook faced legal challenges. In December 2009, ten privacy groups filed a complaint with the Federal Trade Commission[3] over changes to Facebook's privacy policy which included default settings that made users' status updates available potentially to all Internet users, as well making users' friend lists publicly available.

There were three major components to Facebook's policy change. The first was that all privacy settings were aggregated into one simple control. Users no longer had to deal with 170 granular options. As depicted in Figure 2, this interface was far more approachable and easily adjustable than before. Second, Facebook no longer required users' friends and connections to be visible to everyone. Third, Facebook made it easier to opt out with a single click from third-party applications from accessing users' personal information. Generally, these changes were met favorably. For example, the chairman of the American Civil Liberties Union, Chris Conley, wrote 'The addition of simplified options (combined with the continued ability to fine-tune your settings if you wish) and user control over Facebook's 'connections' are significant improvements to Facebook's privacy.'

[Figure 2 about here.]

This change in privacy settings did not change how the banner ads that were served on a Facebook's website were targeted, or whether advertisers could use user information to personalize ads. Display advertising was treated separately because, as Facebook states, 'Facebook's ad targeting is done entirely anonymously. If advertisers select demographic targeting for their ads, Facebook automatically matches those ads to the appropriate audience. Advertisers only receive anonymous data reports.'[4]

To reassure advertisers that the change would not adversely affect them, Facebook sent out an email to its advertisers saying that 'this change will not affect your advertising campaigns' (The full letter is reproduced in the appendix.) This means that though users were given control over how much information was being shared and the extent to which they were being tracked by third parties, the actual mechanism by which the ads tested were targeted and served was not changed. This means that the policy change should be viewed as the effect of a policy which was designed to let users control what information was published and whether they were being tracked by third parties in a personally identifiable way, rather than a policy change that limited how well advertisers could use information to advertise with.

We compare the relative performance of personalized and targeted ads before and after this policy change to see whether the ability of users to control their personal data, makes them more or less responsive to different forms of advertising. In particular, we evaluate whether explicit control over their privacy settings makes Facebook users more or less responsive to ads that are personalized. It is not clear, *ex ante*, how the policy change will affect the performance of personalized ads. On the one hand, more control over data may make users more likely to respond to advertising favorably, as they know that their data have not and will not be misused. Further, behavioral research has emphasized the importance of consumer perceptions of control for mitigating privacy concerns (Xu, 2007; Brandimarte et al., 2010).On the other hand, when users have explicit control over their privacy settings, that may simply increase the salience of privacy concerns, making personalized advertising more problematic, since users may already feel that they have rejected it. Targeted ads provide a particularly useful comparison because they were randomly served to the same type of individuals as the personalized ads, and the only difference was that the ad itself did not make clear what information (if any) was being used to target the ad. A further comparison to untargeted ads allows us to make sure there was no underlying shift in advertising effectiveness that was general to Facebook and not related to the salience of privacy concerns.

Data

We obtained daily data from the nonprofit on how well each of the ads performed for the duration of the experiment. There were 79 different ad campaigns for which we obtained daily data on the number of times they were shown and the number of clicks. In total these ads were shown to 1.2 million users and they received 1,995 clicks. When a user clicked on the ad, they were taken to the nonprofit's Facebook page.

There were 39 different targeting variables, which each had a personalized and a targeted variant. 19 of these targeting variables were based on where the user had expressed liking of various celebrities, 20 on whether the user had attended a particular liberal arts college. There was also an untargeted campaign. These data spanned 2.5 weeks on either side of the introduction of privacy controls on May 28, 2010. We also check robustness to this time-span in Table 4.

This data included the number of unique impressions (that is the number of users the ad was shown to) and the number of clicks each ad received. It contains information on the date that click was received but does not provide time of day information. It also includes data on the cost to the nonprofit per click and the imputed cost per thousand impressions. There is also information collected separately by the nonprofit while designing its ads on the potential 'ad-reach' of each of its targeting variables. This reflects the number of Facebook users whom Facebook estimated could be in the target segment for any targeted ad-campaign. To protect the privacy of the nonprofit's supporters, we did not receive information about the backgrounds or identities of those who chose to like it, or on any of their actions after they made that choice. We also do not have information about whether these users did indeed change their privacy settings.

Table 3 reports the summary statistics. The average number of clicks relative to ad impressions is relatively small, at two-tenths of one percent. This is even smaller when looking at the daily level, since many campaigns received no clicks on a given day, inflating the appearance of low click-through rates. We use both aggregate and daily measures of click-through rates in our regressions, and find qualitatively similar results. However, this is similar to rates reported by other advertisers for Facebook ads. In their provocatively-titled piece 'Facebook Ad Click-Through Rates Are Really Pitiful', Barefoot and Szabo (2008) reported average click-through rates between .01% and 0.06%. As a side note, the nonprofit considers the campaign to have been an immense success, especially given the relatively small cost of the trial (less than $1,000). From having a small social media presence, it was able to achieve a vibrant and active set of fans who respond and interact to news updates delivered about the nonprofit's mission and activities. Compared to its peer nonprofits, it now has a far broader and deeper social media presence.

[Table 3 about here.]

Analysis

Figure 3 displays the average click-through rate for each campaign before and after the policy change. Ads that personalized their content appeared to greatly increase in effectiveness after the policy change. This change was highly significant (p-value=0.0047). The effects of targeting ads without personalizing their content before and after the policy change were not significantly different (p-value=0.407). There appears to be little change in the effectiveness of the un-targeted campaign, though of course with only one campaign it is impossible to assess statistical significance.

[Figure 3 about here.]

Figure 4 examines whether there were any differences for the two types of targeting variable (by school and by celebrity). It is evident that on average the celebrity-focused campaign was more successful on average at attracting clicks. However, it appears clear that there was a similar incremental jump in the effectiveness of personalized ads after the policy change for both kinds of targeting variable.

[Figure 4 about here.]

Figure 3 suggests that the personalization of ads was more effective after Facebook reassured users about privacy concerns by facilitating users' taking control of their personal information. To check the robustness of this result, we also performed regression analysis. This allows us to assess the statistical significance of our results in various robust ways. It also allows us to make sure the result is not an accidental result of a Facebook ad server's particular randomized allocation of one campaign to a particular day.

We model the click-through rate $ClickRate_{jt}$ for ad j on day t in the following manner:

$$ClickRate_{jt} = Personalized_j \times PostPolicy_t \gamma_k + \delta_t + \epsilon_j \tag{1}$$

$Personalize_j$ is an indicator variable which is equal to one if the ad contained personalized content matched to the variable on which it was targeted, and zero if there was no personalized content. $PostPolicy_t$ is an indicator variable equal to one if the date was after the privacy-settings policy change took place, and zero otherwise. γ_k is a vector of fixed effects for the 20 different undergraduate institutions targeted and each of the 19 celebrities targeted. These control for underlying systematic differences in how likely people within that target segment were to respond to this charity. We include a vector of date dummies δ_t. These are collinear with the $PostPolicy_t$ which means $PostPolicy_t$ is dropped from the specification. Because the ads are randomized, both of these fixed effects should primarily improve efficiency. We estimate the regression using ordinary least squares. Following evidence presented by Bertrand et al. (2004) we cluster standard errors at the ad-campaign level to avoid artificially understating our standard errors due to the fact we have panel data.

[Table 4 about here.]

Table 4 presents our results. Column (1) is our main specification, as suggested by Equation 1. The crucial variable of interest is $Personalize \times PostPolicy$. This captures how an individual exposed to a personalized ad responds differently to a personalized ad compared to a targeted ad with generic wording after Facebook's change in privacy policy. It suggests a positive and significant increase in the performance of personalized ads relative to merely targeted ads after the policy change enhanced user privacy controls. The magnitude of our estimates suggest that the percentage click-through rate increased by 0.024, relative to a baseline percentage click-through rate of .0232 for personalized ads. In other words, the click-through rate doubled. The negative coefficient $Personalized$ which is marginally significant suggests that prior to the change in privacy settings personalized ads were less effective than ads that did not use personalized ad copy.

This empirical analysis uses a short time window of 5 weeks. This has a flavor of regression discontinuity in its approach, in that it looks for changes in consumer behavior around a short time window (Busse et al., 2006). This means that it is unlikely that there was some long-run trend, for example, in terms of increasing user acceptance of personalization of ads, that drove the results. It also helps to rule out an explanation that people were becoming 'habituated' to privacy concerns because of the relentless media coverage and consequently less reluctant to click on personalized ads. To provide further evidence that the change was connected with the policy change, we repeated our estimation using a shorter time window of 10 days around the policy change. We use a specification similar to that in column (2) of Table 4. The results, reported in Column (2) of Table 4, were similar.

We also did a falsification check, where we examined a similar 10 day window that occurred before the policy change at the end of May, to see whether we see any similar time trend in the second week compared to the first week even when there was no policy change in effect. By examining such a 'placebo' policy change that occurred midway through this 10-day window, we are able to provide some suggestive evidence or not about whether such dramatic swings as we observe in our data are observed even when there was no actual policy change. Column (3) of Table 4 reports the results. It is clear that there was no statistically significant change in the performance of ads and that in particular there was no difference in the time-trend of the performance of personalized ads versus merely targeted ads.

The policy change can be tied to a date more specifically than the introduction of many policy changes, but there are still a few open questions about the effective timing of the change, given that, as shown in Table 2, there was a series of discrete events surrounding the policy change. The remainder of Table 4 checks robustness to various timing assumptions we made about the actual date of the policy change. There were two days in between Facebook announcing the policy change on May 24 and announcing the details of the policy change on May 26. Since this may have introduced uncertainty, we checked robustness to omitting these two days in column (4). The results are similar. The privacy changes were heralded in the press on May 27th, but there was also a rollout period where Facebook introduced the first of the three changes in its privacy policy - the new privacy settings page - gradually, presumably to minimize the risk of servers becoming overloaded. It is not clear how long this rollout period lasted or how long the majority of US subscribers (who we study) took to obtain access. To check robustness to this rollout, we excluded the first five days of data after the policy change.[5] The results are reported in column (5). Again, the results are similar.

The results in Table 4 confirm the insight from Figure 3 that the policy change improved the performance of personalized ads in terms of the average campaign click-through rate. To try to measure how the policy change affected an individual's likelihood of clicking on an ad, we also estimate an individual-level logit model. One advantage of an individual-level model is that we can include the untargeted campaign in our regressions as an additional baseline, as rather than one observation of a click-through rate of the untargeted campaign which is collinear with the fixed effects, there are hundreds of thousands of observations of how individuals responded to that campaign.

We model the probability that an individual i clicks on ad j on day t as:

$$\begin{aligned} Clicked_{ijt} = I(&Personalized_j \times PostPolicy_t + Targeted_j \times PostPolicy_t \\ &+ Personalized_j + Targeted_j + Postpolicy_t + \gamma_k + \delta_t + \epsilon_j) \end{aligned} \tag{2}$$

Equation (2) is similar to Equation (1), except for the inclusion of a new indicator variable $Targeted_j$. $Targeted_j$ is an indicator variable for whether the ad was targeted, but had no attempts at personalization - in other words, it would have been difficult for the consumer to know why they received that ad. As explained by Ai and Norton (2003), in a logit model, compared to a linear probability model, interpretation of interaction terms is not straightforward, as they are a cross-derivative of the expected value of the dependent variable. This is a particular issue for three-way interactions. To address this, we estimated a logit model and used these logit estimates to predict average probabilities while taking into account the fact that there were cross-derivatives in the specification. Table 5 reports the results of these logit model predicted probabilities. The results are reassuringly similar to those in Table 4, even though we are now studying behavior at an individual level. Personalized ads were significantly worse than non-personalized ads before the policy, but performed almost twice as well after the policy. There was no significant shift in the efficacy of untargeted or targeted ads before and after the policy.

[Table 5 about here.]

Further Robustness Checks

We also obtained further information to determine whether there were any other environmental changes that could explain the result. One potential concern is that our results reflect a change in the numbers of users of Facebook. For example, an alternative explanation of our results could be that the negative publicity drove more experienced users away, leaving only users who were likely to react to personalized advertising using Facebook. However, the data in Table 6 suggest that this was not the case. This data set was obtained from comScore and is based on their panel of two million internet users. There was little change in the composition of the user base in June relative to May, compared to the shifts seen from April to May. Further, the number of unique visitors to Facebook appeared to steadily, if slightly, increase across the months.

[Table 6 about here.]

Another concern is that the results could reflect a change in the composition of advertisers. For example, perhaps other advertisers pulled out of Facebook as a result of the negative publicity of privacy, meaning that perhaps there were fewer advertisers competing to personalize advertising, which made the personalized ads relatively more attractive. Though we cannot check for evidence of this directly, we are able to provide some suggestive evidence against this counter-explanation by looking at the pricing data for the ads. If there had been a drop-off in advertisers, we would expect also to see a decrease in the price paid in the auction, as the price should theoretically be a function of the number of bidders (McAfee and McMillan, 1987). However, the small drop in cost per click of 1.5 cents after the policy change was not statistically significant (p-value=0.59).

Mechanism: Rarity of User Information

One question is what privacy concerns consumers had that the policy change resolved, leading customers then to respond more positively to personalized ads. One explanation is that consumers were concerned that advertisers, when using this information, were tracking the user's behavior in a potentially intrusive way. Ad-intrusiveness has been documented by Goldfarb and Tucker (2010a) as potentially off-putting for consumers. Facebook had experienced negative publicity for sharing data with third-party advertisers, so this may have been a reasonable concern. In this section, the paper presents some suggestive evidence that supports this interpretation.

We explore this by exploiting additional data and variation across ad campaigns about how many users were in the target group for that particular campaign such as that provide in Figure 1. If the main effect of giving users transparent control over how their data are used is that they are less concerned they being tracked, then we might expect to see this reflected in differences in consumer behavior across the rarity of information used in the ad. For example, if an ad was personalized around the fact that a Facebook user liked 'cooking,' then Facebook has 3,167,540 users who state they like cooking. The use of such information might be felt to be more anonymous and less likely to be a result of, or result in, privacy-violating tracking behavior by that advertiser. However, if an ad was personalized around the fact that a user liked the Korean delicacy kimchi, then there are only 6,180 Facebook users who say that they do like kimchi; knowing that such a preference is relatively rare might make the user more concerned they were being tracked by the advertiser in a privacy-violating manner.

Table 7 investigates how our effects were moderated by the rarity of the targeting variable or the ad reach. Column (1) of Table 7 reports how the efficacy of personalized ads relative to ads that were targeted to users' interests before and after the policy change was affected by the reach of these targeting variables. The negative coefficient on 'Post-Policy × Personalized × Ad Reach' suggests that the positive effect is smaller for ads that had a larger ad reach than those that had a smaller ad reach. In other words, personalization was relatively more successful after the policy change for celebrities who had smaller fan bases, as can be seen from the larger point estimate for 'Post-Policy × Personalized' relative to Table 4, column (1). Ad-Reach is denominated in millions of users. Therefore, roughly extrapolating from the linear functional form, our estimates suggest that for ads for the 7 percent of campaigns in our sample that have target audiences of greater than 372,000, the effect of the policy was canceled out. However, for the median campaign, which had 7,560 people in the target market, the effect of the policy change actually raised the click-through percentage by 0.03, relative to a mean of 0.02.

Column (2) of Table 7 repeats this exercise for ads that used the shorter ten-day window. Again, the results appear robust, providing evidence against an interpretation that that an unobserved time-trend unrelated to the change in policy drove the results. Column (3) of Table 7 provides further evidence that the change observed was connected with the policy change by conducting a 'falsification' check for a placebo policy-change mid-way through a 10-day window prior to the policy change. The results are again insignificant.

These results suggest that the shift towards giving users control over their personal information, had the largest effect for personalized advertising that attempted to use more unusual pieces of information. This provides suggestive evidence that the change in privacy policy was able to reassure users that despite the advertisers' use of potential unusual information about them, that they should be less worried about their privacy. It also suggests that the gains to personalized advertising with appropriate privacy concerns, may be highest when the advertiser uses information that is relatively unique about that user.

[Table 7 about here.]

Implications

This paper is the first to explore the consequences for advertising-supported websites of giving users more control over how their personal information is shared. The paper uses data from a randomized experiment conducted by a nonprofit that was designed to explore the relative merits of targeting ads, and ads that used user information to personalize content of the ad. During the field experiment, the social networking site on which the experiment was being conducted unexpectedly announced that it would change users' privacy settings. These changes, which were publicly applauded by consumer advocates, gave users greater control over what personally identifiable information was shared and whether third parties could track their movements. They did not however, affect the anonymous use of information by advertisers to target their information. Empirical analysis suggests that after this change in policy, the websites' users were roughly twice as likely to react positively to personalized ad content and click on personalized ads. There was generally no economically significant change in their reactions to un-targeted or merely targeted ads. This suggests that publicly giving users control over their private information can benefit advertising-supported media and advertisers on such sites.

There are obvious limitations to this research that are worth mentioning. First, the randomized experiment was conducted by a nonprofit with an appealing cause. Consumers may be ready to ascribe less pernicious motives to a nonprofit than to a for-profit company when they observe their advertising. Second, this randomized experiment was conducted at a time when privacy concerns were particularly sensitive and salient in consumers' eyes. It is not clear how the results will change when social network users are not being reminded about the control they have (or lack of it) over their private information. Third, the type of privacy control introduced by Facebook that we study was just one of a myriad of potential ways that social networks or other advertising-supported websites could used to give control to their users over their privacy settings. It would be interesting for future research to see whether an explicit 'opt-in' approach to sharing information or changes in privacy policies that explicitly addressed advertising could produce equally striking results. Notwithstanding these limitations this paper does provide initial evidence of the potential role of privacy-settings and privacy control as a useful and important strategic marketing variable.

Notes

[1]As stated in Stone (2010) 'What a marketer might think is endearing, by knowing a little bit about you, actually crosses the line pretty easily.'

[2]The nonprofit is eager to protect the privacy of its supporters, and consequently has asked the authors to not reveal either the names of the celebrities or of the schools that were used in this advertising campaign.

[3]http://epic.org/privacy/inrefacebook/EPIC-FacebookComplaint.pdf.

[4]There is one privacy control that governs the use of information in Facebook ads which can be reached from an entirely different menu than the regular user privacy settings (Account Settings > Facebook Ads rather than Account > Privacy Settings). This enables users to control whether ads shown by Facebook can use their name or photo in ads served to their friends that are linked with social actions (such as becoming a fan of a page). The settings for this control appear to have not been affected by the change in privacy policy. This feature was also not used in the advertising campaigns studied.

[5]Anecdotal evidence from independent forums discussing the introduction of privacy settings suggests that US users had been given access within the five days.

References

Acquisti, A. and R. Gross (2006). Imagined communities: Awareness, information sharing, and privacy on the Facebook. In *6th Workshop on Privacy Enhancing Technologies*, pp. 36–58.

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

Anand, B. and R. Shachar (2009, September). Targeted advertising as a signal. *Quantitative Marketing and Economics 7*(3), 237–266.

Aral, S. and D. Walker (2010). Creating social contagion through viral product design: Theory and evidence from a randomized field experiment. *Mimeo, NYU*.

Barefoot, D. and J. Szabo (2008, April). Facebook ad click-through rates are really pitiful. *Friends with Benefits: A social media handbook Blog*.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics 119*(1), 249–275.

Bilton, N. (2010, May 12). Price of Facebook Privacy? Start Clicking. *New York Times*.

Brandimarte, L., A. Acquisti, and G. Loewenstein (2010). Misplaced confidences: Privacy and the control paradox. *Mimeo, Carnegie Mellon University*.

Busse, M., J. Silva-Risso, and F. Zettelmeyer (2006, September). $1,000 cash back: The pass-through of auto manufacturer promotions. *American Economic Review 96*(4), 1253–1270.

Caverlee, J. and S. Webb (2008). A Large-Scale Study of MySpace: Observations and Implications for Online Social Networks. In *Proceedings from the 2nd International Conference on Weblogs and Social Media (AAAI)*.

Chellappa, R. K. and R. G. Sin (2005). Personalization versus privacy: An empirical examination of the online consumer's dilemma. *Information Technology and Management 6*(2-3), 181–202.

Corbin, K. (2010, January 28). Privacy bill nears introduction in house. Technical report, Esecurity Planet.

Culnan, M. J. (1993). How Did They Get My Name? An Exploratory Investigation of Consumer Attitudes Toward Secondary Information Use. *MIS Quarterly 17*(3), 341–363.

Golder, S. A., D. M. Wilkinson, and B. A. Huberman (2007). Rhythms of social interaction: Messaging within a massive online network. In *Proceedings of the Third International Conference on Communities and Technologies*.

Goldfarb, A. and C. Tucker (2010a). Online display advertising: Targeting and intrusiveness. Forthcoming, Marketing Science.

Goldfarb, A. and C. Tucker (2010b). Privacy regulation and online advertising. Forthcoming, Management Science.

Holahan, C. (2007, November). So Many Ads, So Few Clicks: Can more targeted pitches on Facebook and other sites reverse the shrinking response to online ads? *Business Week*.

Kihlstrom, R. E. and M. H. Riordan (1984, June). Advertising as a signal. *Journal of Political Economy 92*(3), 427–50.

Lipsman, A. (2010, January 2010). 2009: Another Strong Year for Facebook. *Comscore Voices*.

Lohr, S. (2010, April 30). Privacy concerns limit online ads, study says. *New York Times*.

Manchanda, P., Y. Xie, and N. Youn (2008). The role of targeted communication and contagion in new product adoption. *Marketing Science 27*(6), 961–976.

McAfee, R. P. and J. McMillan (1987). Auctions and bidding. *Journal of Economic Literature 25*(2), 699–738.

Milgrom, P. and J. Roberts (1986, August). Price and advertising signals of product quality. *Journal of Political Economy 94*(4), 796–821.

Miller, A. (1971). *The assault on privacy: Computers, data banks, and dossiers*. University of Michigan Press.

Nielsen (2010, January 14, 2010). Top U.S. web brands and site usage: December 2009. *Nielsen Wire*.

Small, D. and N. Verrochi (2009). The face of need: Facial emotion expression on charity advertisements. *Journal of Marketing Research 46*, 777–787.

Smith, H. J., S. J. Milberg, and S. J. Burke (1996). Information privacy: Measuring individuals concerns about organizational practices. *MIS Quarterly 20*(2), 167–196.

Spiekermann, S., J. Grossklags, and B. Berendt (2001). E-privacy in 2nd generation E-commerce: privacy preferences versus actual behavior. In *EC '01: Proceedings of the 3rd ACM conference on Electronic Commerce*, New York, NY, USA, pp. 38–47. ACM.

Stone, B. (2010, March 3). Ads Posted on Facebook Strike Some as Off-Key. *New York Times*.

Trusov, M., R. E. Bucklin, and K. Pauwels (2009, September). Effects of word-of-mouth versus traditional marketing: Findings from an internet social networking site. *Journal of Marketing 73*, 90–102.

Turow, J., J. King, C. J. Hoofnagle, A. Bleakley, and M. Hennessy (2009). Americans Reject Tailored Advertising and Three Activities that Enable It. *Mimeo, Berkeley*.

Xu, H. (2007). The effects of self-construal and perceived control on privacy concerns. In *28th Annual International Conference on Information Systems (ICIS 2007),Montreal, Canada.*

Facebook's Notification to Advertisers: May 26, 2010

Facebook will roll out changes today that will make it easier for our users to understand and control their privacy settings. As this change will have an impact on our users, we wanted to let you, a valued advertising partner, know about it. Please note that this change will not affect your advertising campaigns and there is no action required on your part.

Facebook is a company that moves quickly, constantly innovating and launching new products to improve the user experience. The feedback we heard from users was that in our efforts to innovate, some of our privacy settings had become confusing.

We believe in listening to our users and taking their feedback into account whenever possible. We think the following changes address these concerns by providing users with more control over their privacy settings and making them more simple to use.

Starting today, Facebook will:

* Provide an easy-to-use "master" control that enables users to set who can see the content they share through Facebook. This enables users to choose, with just one click, the overall privacy level they're comfortable with for the content they share on Facebook. Of course, users can still use all of the granular controls we've always offered, if they wish.

* Significantly reduce the amount of information that must be visible to everyone on Facebook. Facebook will no longer require that users' friends and connections are visible to everyone. Only Name, Profile Picture, Networks and Gender must be publicly available. Users can opt to make all other connections private.

* Make it simple to control whether other applications and websites access any user information. While a majority of our users love Facebook apps and Facebook-enhanced websites, some may prefer not to share their information outside of Facebook. Users

can now opt out with just one click.

I encourage you to take a moment to read our CEO Mark Zuckerberg's blog post and check out the new Facebook Privacy Page.

Thanks, The Facebook Ads Team

Table 1: Campaigns appeals in different conditions

Targeting Variable	College	Interest
Personalized	As a [undergraduate institution name] graduate you had the benefit of a great education. Help girls in East Africa change their lives through education.	As a fan of [name of celebrity] you know that strong women matter. Help girls in East Africa change their lives through education.
Non-Personalized	You had the benefit of a great education. Help girls in East Africa change their lives through education.	You know that strong women matter. Help girls in East Africa change their lives through education.

Table 2: Timeline for Facebook Growth, Privacy and Advertising

Date	Event
February 2004	Facebook launched from Harvard dorm room.
November 2007	Facebook launches 'Facebook ads'. Advertising pilot involving 'beacons' (small 1x1 pixel web bugs) allows Facebook to track users' movements over other websites for purposes of targeting.
December 2007	Facebook makes Beacon an opt-out service after negative publicity.
September 2009	Beacon ad targeting program shut down amid class-action suit.
November 2009	Facebook changes its default settings to publicly reveal more of its users' information that had previously only been available to Facebook users. This information could now be tracked by third-party search engines.
December 9 2009	Privacy settings are entirely removed from certain categories of users' information. These categories include the user's name, profile photo, list of friends and pages they were a fan of, gender, geographic region, and networks the user was connected to. They are instead labeled as publicly available to everyone, and can only be partially controlled by limiting search privacy settings. Founder Mark Zuckerberg's photos are apparently inadvertently made public by the change in settings.
December 17 2009	Coalition of privacy groups led by the Electronic Frontier Foundation files a complaint with Federal Trade Commission over changes to privacy settings
April 2010	Facebook users' General Information becomes publicly exposed whenever they connect to certain applications or websites such as the online review site Yelp. General Information includes users' name and their friends' names, profile pictures, gender, user IDs, connections, and any content shared using the Everyone privacy setting.
May 12 2010	New York Times publishes article entitled 'Facebook Privacy: A Bewildering Tangle of Options' (Bilton, 2010). This ignites a firestorm of negative press about Facebook and privacy.
Monday May 24 2010	Facebook founder Mark Zuckerberg announces in an editorial in the Washington Post that Facebook will institute new privacy settings
Wednesday May 26 2010	Facebook unveils new privacy settings in press event
Thursday May 27 2010	Facebook starts rollout of privacy settings. New York Times publishes 'A Guide to Facebook's New Privacy Settings'.
Saturday May 29 2010	First reports of new privacy setting controls being seen by users

Additional Sources: Facebook's official public timeline; 'Facebook's Eroding Privacy Policy: A Timeline': Electronic Frontier Foundation April 2010.

Table 3: Summary Statistics

	Mean	Std Dev	Min	Max
Average Impressions	15892.7	63274.2	337	551783
Average Clicks	25.3	53.7	0	374
Average Cost Per Click	0.38	0.096	0.11	0.50
Cost per 1000 views	0.095	0.12	0	0.39
Ad-Reach (000000)	0.095	0.21	0.00098	0.99
Aggregate Click-Through Percentage	0.17	0.23	0	1.37
Daily Click-Through Percentage	0.023	0.14	0	3.13

Campaign level data. 79 Different Campaigns (78 campaigns based on 39 different targeting variables each with personalized and targeted variants. 1 untargeted campaign)

Table 4: Initial Results

	Main Result	10-Day Window	Fake Policy	Unveiling excluded	Rollout excluded
	(1)	(2)	(3)	(4)	(5)
Personalized × PostPolicy	0.0236**	0.0554***		0.0235**	0.0218**
	(0.0102)	(0.0208)		(0.0105)	(0.0110)
Personalized × Fake-PostPolicy			-0.0144		
			(0.0192)		
Personalized	-0.0119*	-0.0112	-0.0116	-0.0118*	-0.0112
	(0.00627)	(0.0115)	(0.0103)	(0.00675)	(0.00715)
Date Fixed Effects	Yes	Yes	Yes	Yes	Yes
Targeting Variable Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	2730	780	780	2574	2340
R^2	0.060	0.118	0.082	0.062	0.068

OLS Estimates. Dependent variable is percentage daily click through rate.
Robust standard errors clustered at ad-level. * $p < 0.10$, ** $p < 0.05$,*** $p < 0.01$
$PostPolicy_t$ is collinear with the date fixed effects and dropped from the specification.

Table 5: Predicted Probabilities based on a logit regression

Ad-Type Time-Period	Predicted Probability	Std. Err.	Z-Statistic	[95% Conf.	Interval]
Untargeted × Pre-Policy	0.0003	0.0001	10.7300	0.0002	0.0005
Untargeted × Post-Policy	0.0006	0.0001	10.0000	0.0005	0.0007
Targeted × Pre-Policy	0.0024	0.0001	25.0700	0.0022	0.0026
Targeted × Post-Policy	0.0025	0.0001	22.4100	0.0023	0.0027
Personalized × Pre-Policy	0.0018	0.0001	14.8800	0.0016	0.0021
Personalized × Post-Policy	0.0035	0.0001	25.6400	0.0032	0.0037
Log-likelihood	13,825				
Observations	1,248,899				

Marginal Effects calculated from logit model where dependent variable is individual click-through probabilities. Date and Targeting fixed effects included.

Table 6: Little Change in Facebook User-Base Composition

Proportion of Group	April 2010	May 2010	June 2010
Age <17	10.4	10.6	11.4
Age 18-24	19.2	19.4	18.6
Age 25-34	20.8	20.7	20.8
Age 35-44	20.4	19.9	19.9
Age 45-54	16.7	16.5	16.5
Age 55-64	8	8.1	8.1
Age 65+	4.6	4.8	4.7
Income <$15k	10.1	10.3	9.7
Income $15-24k	6.2	6.1	5.9
Income $25-39k	12.5	12.7	13.5
Income $40-59k	22.1	22	24.2
Income $60-74k	10.9	11.3	9.6
Income $75-99k	16.8	16.3	15.3
Income $100k+	21.5	21.2	21.8
Male	47.2	47.1	48.2
Female	52.8	52.9	51.8
Total Unique Visitors	121 Million	130 Million	141 Million

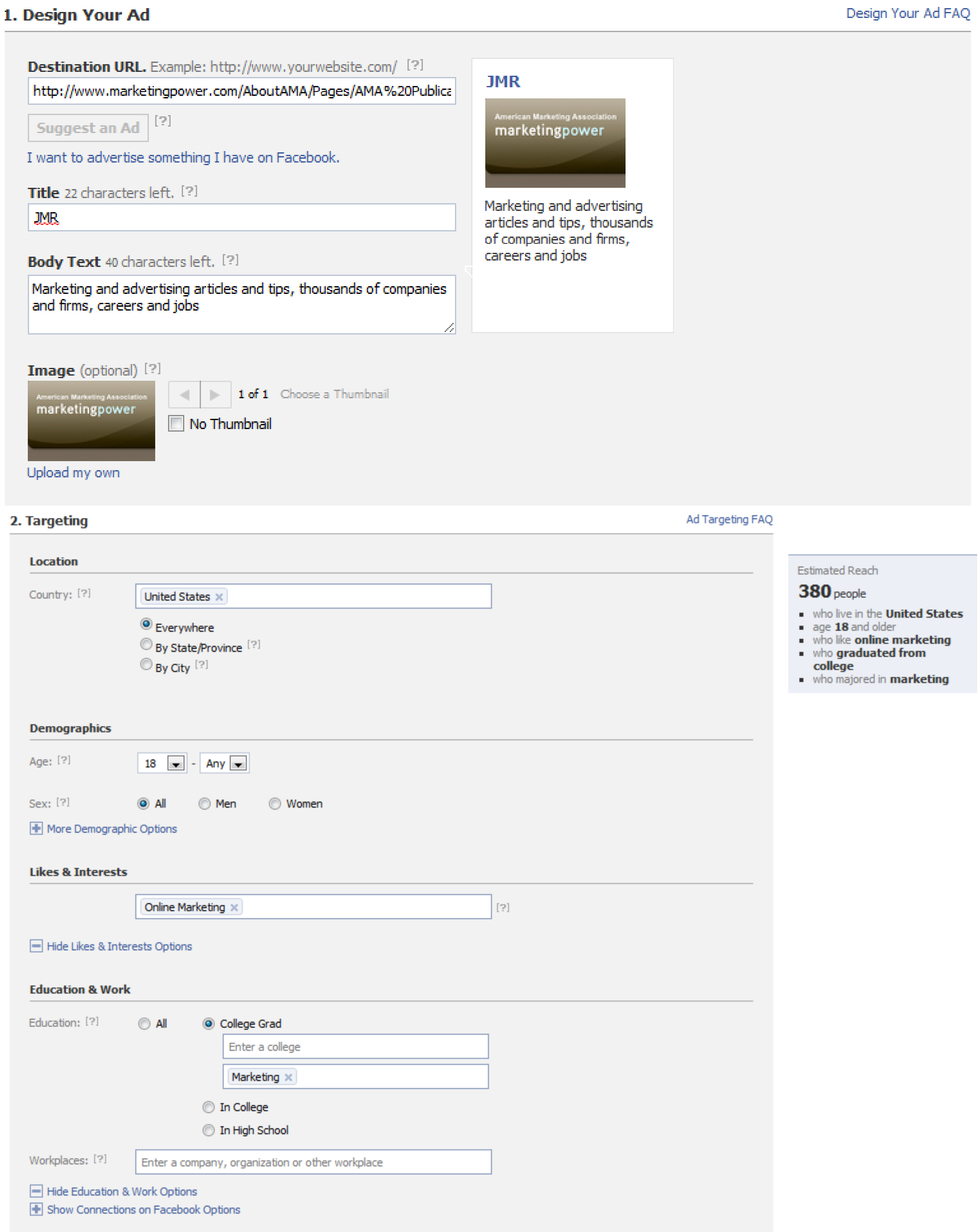
Source: Comscore Marketer Database

Table 7: Stratification

	(1) Main	(2) 10-Day Window	(3) Falsification Check
Post-Policy × Personalized × Ad-Reach	-0.0852**	-0.199**	
	(0.0421)	(0.0966)	
Fake Post-Policy × Personalized × Ad-Reach			0.0615
			(0.0559)
Personalized	-0.0153**	-0.0228	-0.0138
	(0.00670)	(0.0188)	(0.0135)
Post-Policy × Personalized	0.0317**	0.0662***	
	(0.0147)	(0.0251)	
Personalized × Ad-Reach	0.0354	0.125	0.0232
	(0.0214)	(0.0850)	(0.0278)
Post-Policy × Ad-Reach	0.0150	-0.00677	
	(0.0350)	(0.0497)	
Fake Post-Policy × Personalized			-0.0202
			(0.0237)
Fake Post-Policy × Ad-Reach			0.00877
			(0.0493)
Date Fixed Effects	Yes	Yes	Yes
Targeting Variable Fixed Effects	Yes	Yes	Yes
Observations	2730	780	780
R^2	0.062	0.129	0.085

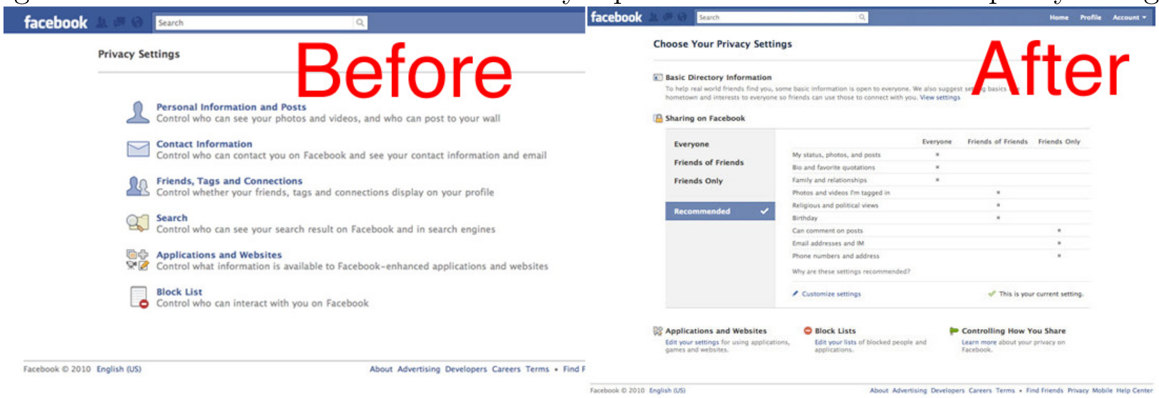
OLS Estimates. Dependent variable is percentage daily click through rate.
Robust standard errors clustered at ad-level. * $p < 0.10$, ** $p < 0.05$,*** $p < 0.01$
$PostPolicy_t$ is collinear with the date fixed effects and dropped from the specification. $Ad\text{-}Reach_k$ is collinear with the targeting variable fixed effects and also dropped from the specification.

Figure 1: Facebook: Screenshot of Ad Targeting Interface



Source: Mock-up ad campaign for marketing journal created by authors

Figure 2: Facebook: Screenshots of Privacy Options before and after the policy change



Source: Gawker Media

Figure 3: Comparison in Click-Through rates before and after

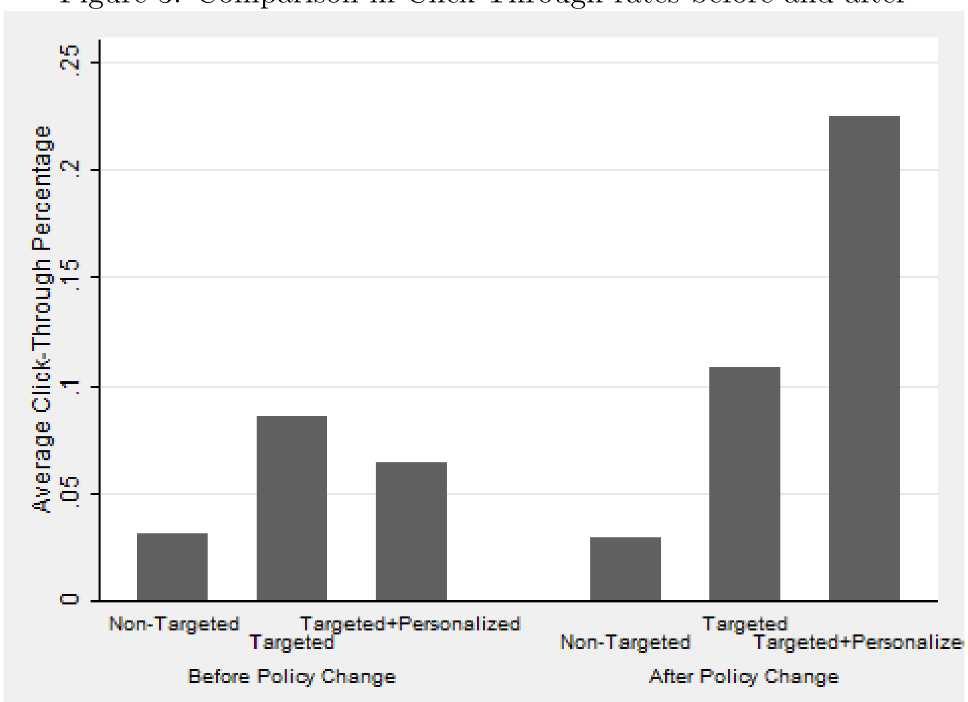


Figure 4: Comparison in Click-Through rates before and after