Privacy Regulation and Online Advertising

Avi Goldfarb and Catherine Tucker

University of Toronto and MIT Sloan

Privacy Regulation is a new but important question.

- The Internet allows firms to collect large amounts of customer data
  - Increase in ability to target, tailor and optimize advertising
- Consumers are concerned about threats to their privacy
- Pressure for regulation in US

Do privacy regulations influence online ad effectiveness?

- Which websites and ads are most affected?
- Speculatively, how might this affect evolution of advertising-supported internet?

Setting

- Regulatory Setting
  - European Law become stricter 2003-4
  - Privacy regulation elsewhere has not changed since advent of commercial internet
  - Compare change in ad effectiveness in Europe relative to elsewhere.
- Data
  - Field tests of 9596 different online display ad campaigns across multiple countries
  - For each campaign, $\tilde{3}47$ web users surveyed on purchase intention and ad recall. Half had seen the ad and half were in a control group
- Method: Diff-in-Diff-in-Diff
  - Difference between treatment and control groups in field studies
  - Difference before and after the regulation in Europe
  - Difference between Europe and elsewhere

Privacy Regulation affects performance of online ads

- Advertising effectiveness dropped 65% in the EU relative to the rest of the world
  - Drop is specific to European websites rather than European consumers
  - When EU consumers visited US websites they behaved like US visitors
- Not all websites were affected equally
  - Ads on general interest websites (e.g. yahoo.com, nytimes.com) were affected more than ads for targeted websites (e.g. cars.com, babycenter.com)
  - Ads on health websites (which were more strictly regulated) were especially affected
- Not all ads were affected equally
  - Unobtrusive ads were affected more than larger ads and multimedia ads

Implications

Regulation may affect the direction of innovation on the advertising-supported internet

- If ads are less effective, it will limit the scope of the ad-supported internet.
- If ads on general internet websites are particularly affected, such sites will be less able to support themselves through advertising.
  - They may become less prevalent or they may begin to support themselves by other means
  - If unobtrusive ads become less effective, advertisers may increase obtrusive multimedia advertising at the expense of subtle, well targeted ads

More Generally

We do not want to disrupt targeted advertising

Representative Boucher

- Currently, debate is conducted in empirical vacuum.
  - Empirical research is not shameful.
- There may be good reasons to regulate privacy but there are trade-offs
  - The potential reduction in the size of the ad-supported internet
  - The potential change in content on the ad-supported internet
  - The potential increase in the obtrusiveness of ads.

Outline

Collected data on privacy laws in Europe

- 2002/58/EC: European Directive on Privacy and Electronic Communications 'E-Privacy Directive'
- Designed to "particularize and complement" the Data Protection Framework Directive (95/46/EC) for electronic communications
  - Enacted end of 2003 to mid 2004 in UK, Germany, Italy, Holland and France
- The scope of the regulations continued to evolve after 2004 through judicial precedent and some new laws

Legal Disclaimer

Focus on the discrete before/after interpretation

- Not all companies have interpreted the laws in the same way and the European courts are still deciding exactly what is allowed.
- Article 29 working party is currently clarifying their application to behavioral targeting.
- It is clear that the European 'prosecutors' view the EU law as more restrictive that their counterparts in the US. It is also clear that many EU firms view the EU law as stricter.

E-Privacy Directive affects data advertisers can use.

1. Web bugs (or 'beacons', 'action tags', 'clear GIFS', etc.)
   - Widely used, 1x1 pixel pieces of code that allow advertisers to track customers as they move within and across websites
   - Unlike cookies, they are invisible to the user and difficult to block.
   - (Rec 24) Web bugs (beacons) 'may seriously intrude upon the privacy of these users' 'only for legitimate purposes, with the knowledge of the users
   - User 'Consent' is necessary and consent means 'a freely given specific and informed indication of the users wishes, including by ticking a box' concerned.'
2. Cookies
   - (Rec 25) Need notification/opt-out for cookies
3. Click stream data retention is problematic if personal.
   - Health, sexuality, religion, trade-unions
   - Google investigated over retention of IP addresses in Germany

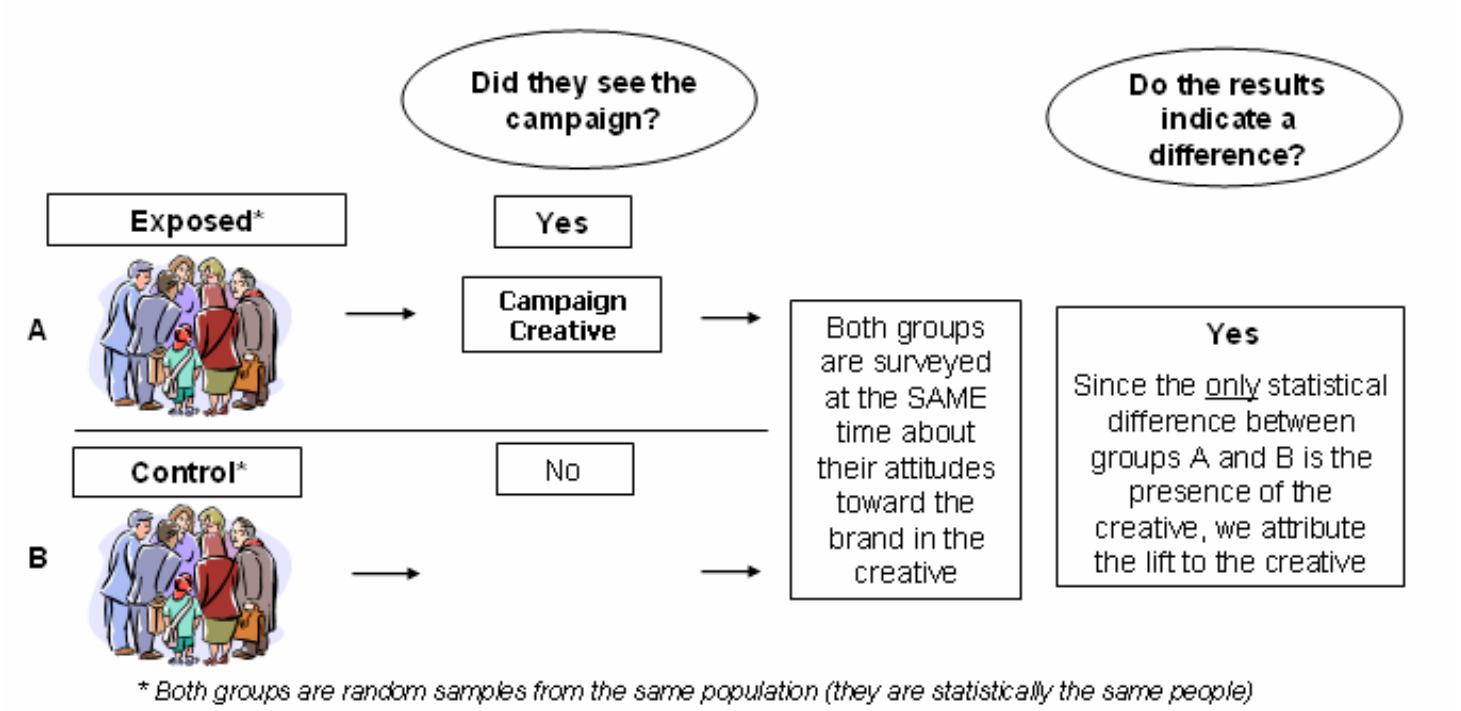
Use survey data from a panel of field tests

- Repeated cross-section survey data from 9,596 field tests of online banner ads in EU, US and rest of world.
- Collected by a media measurement agency to examine the effectiveness of different ad campaigns.
- 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.
  - Both exposed and non-exposed (control) respondents are recruited via an online survey invitation that appears after they have finished browsing the website.
  - Because of the random nature of the advertising allocation, both exposed and control groups have similar unobservable characteristics

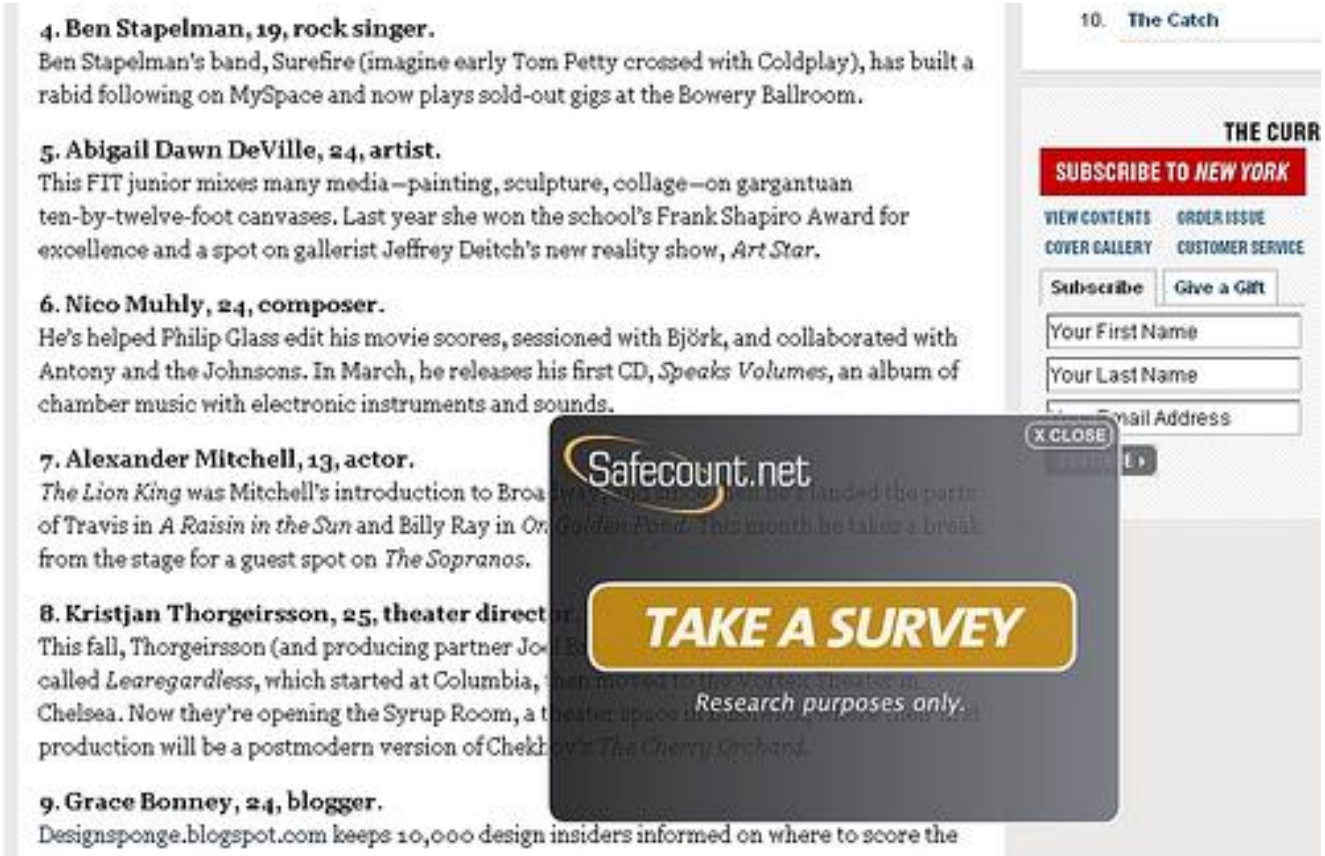
What happens

- Advertiser initiates contract with marketing metrics company to evaluate banner ad performance beyond click-through rates
- Marketing metrics company integrates its services into existing advertising campaign practice
- Anonymized benchmarking data shared with other advertisers (this database is what we were given access to)

Methodology



Methodology



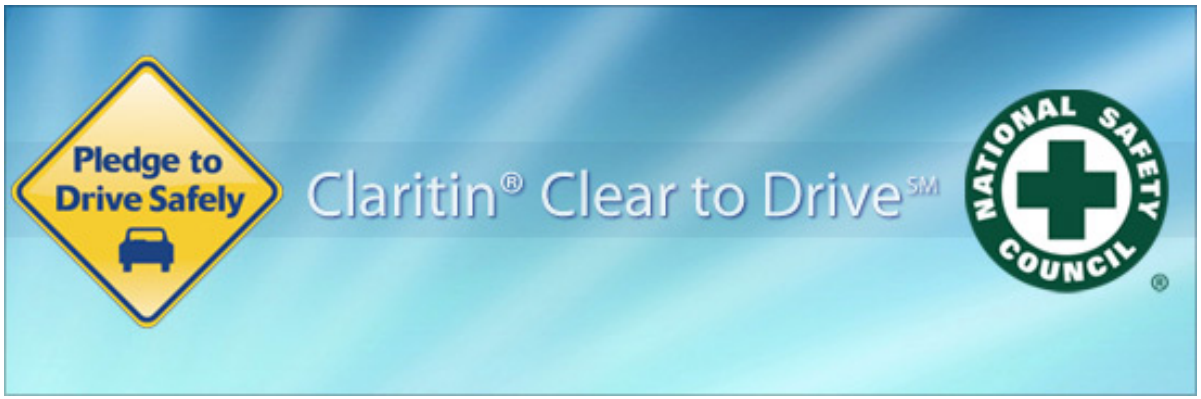
The survey focused on purchase intent

- Survey asked about
  - Purchase intent
  - Demographics and country
  - At end of survey, whether they recalled ad.
- Campaign information
  - Country campaign launched from
  - Media rich features and size of ad
- 400 categories of products, 40 categories of websites, 8 years (2001-2008)

There are two types of selection concerns

- The campaigns are non-random
  - Main data source for the industry
  - Comforting that campaign characteristics change little in the EU relative to elsewhere
  - Acknowledge that little can be said about different types of advertisers
- The respondents might not represent the general internet population
  - Do not know anything about response rates except that they are low.
  - Demographics look representative
  - Cost-effectiveness calculations look right
  - Effects are small (2-4 percentage points) but worthwhile if high purchase intent is worth about 42 cents.
  - Methodology was constant over time

Who should see this ad?



This regulation affects advertisers in our field data via the group they can test ads on

- Prior to field test, the advertisers decide which group of people they wish to show ad to
  - Do not know for certain which tactics where used in the campaigns in the data.
  - Do know that these are large advertisers on the leading edge of technology
- Yahoo! can identify who visited 'allergy relief advice' when deciding whom to serve allergy ads to on Yahoo! News.
- Change in regulation could lead advertisers to do less well at identifying people who might be influenced by the ad

Dependent Variables

- Focus on 'purchase intention'
  - Response to how likely are you to purchase on a five point scale
  - This is weaker than actual purchase data (for example, use in Reiley & Lewis 2009)
  - But a study like Reiley & Lewis is not possible after the EU privacy directive due to the use of web bugs and the requirement of opt-in consent for sharing data across companies
  - Has the advantage of comparability (and scalability) across many categories as discussed in Clark, Doreszelski, and Draganska (2009)
- Robust to favorability and ad recall
  - Response to do you have a favorable opinion of the product on a five point scale
- Discretization of Dependent Variable
  - Focus on whether the respondent reported the highest score on the scale (Likely or Very Likely to Make a Purchase").
  - 37% of respondents are in this category.

Outline

	Mean	Std Dev	Min	Max	Observations
Purchase Intent	0.37	0.48	0	1	3329632
Favorable Opinion	0.42	0.49	0	1	3180804
Ad Recall	0.26	0.44	0	1	3035292
Intent Scale	2.93	1.47	1	5	3329632
Opinion Scale	3.48	1.08	1	6	3180804
Exposed	0.56	0.50	0	1	3329632
EU	0.081	0.27	0	1	3329632
After EU Law	0.81	0.39	0	1	3329632
Female	0.54	0.50	0	1	3329632
Income ($)	64912.4	56342.7	15000	250000	2551263
Age	42.2	15.5	10	100	3283997
Weekly Internet Hours	13.9	10.3	1	31	2606978

Table: Change in ad effectiveness in EU

	EU Difference	T-Test
Before Privacy Law	**-0.030**	-6.994
After European Privacy Law	**-0.002**	-1.188

Table: Change in ad effectiveness outside of EU

	EU Difference	T-Test	**Not-EU Difference**	T-Test
Before Privacy Law	**-0.030**	-6.994	**-0.016**	-11.766
After European Privacy Law	**-0.002**	-1.188	**-0.017**	-27.988

Outline

Move to econometric analysis

- Use simple model to evaluate advertising effectiveness before and after the policy change in EU

For person i exposed to campaign j in country c at time t

$$Intent_{ijct} = \alpha Exposure_i + \beta Exposure_i \times Law_{ct} + \theta X_i + \gamma_{jct} + \epsilon_{ijct}$$

This is a differencein-difference on treatment/control and before/after using the Within Europe data

- Focus on linear probability model.

	EU Data only
	(1)
Fixed Effects	
Exposed × After EU Law × EU	-0.0167**
	(-2.41)
Exposed	0.0256***
	(4.00)
Female	0.0198***
	(5.20)
Std.Internet Hours	0.00936***
	(7.48)
Std. Income	-0.0118***
	(-5.26)
Std. Age	-0.0319***
	(-8.18)
Campaign Fixed Effects	Yes
Observations	271,207
R-Squared	0.160

We control for the time-trend

Use three-way differences in differences to control for time-trend by incorporating rest of the world data

$$Intent_{ijct} = \alpha Exposure_i + \beta_1 Exposure_i \times AfterEULaw_{ct} \times EU_c + \beta_2 Exposure_i \times BeforeEULaw_{ct} + \beta_3 Exposure_i \times NotEU_c + \theta X_i + \gamma_{jct} + \epsilon_{ijct}$$

	EU Data only (1) Fixed Effects	All Data (2) Time Trend
Exposed × After EU Law × EU	-0.0167**	-0.0171**
	(-2.41)	(-2.40)
Exposed	0.0256***	0.0263***
	(4.00)	(4.14)
Female	0.0198***	0.0154***
	(5.20)	(10.32)
Std.Internet Hours	0.00936***	0.0122***
	(7.48)	(35.93)
Std. Income	-0.0118***	-0.00288***
	(-5.26)	(-6.00)
Std. Age	-0.0319***	-0.0185***
	(-8.18)	(-27.05)
Exposed × After EU Law		-0.00109
		(-0.56)
Exposed × Not-EU		-0.00979
		(-1.49)
Campaign Fixed Effects	Yes	Yes
Observations	271207	3329632
R-Squared	0.160	0.172

The results are robust to

- Logit specification (no fixed effects)
- Using the full scale (i.e. treating scale as interval)
- Favorable opinion as the dependent variable
- No controls
- Exclusion of people who saw multiple ads
- Country-specific controls for the timing of the laws
- Dropping Latin America
- Country fixed effects
- The use of different dates to mark the beginning of the regulation

Selection Bias in campaigns in the data

After Law	Mean Non-EU	Mean EU	Difference	T-Test
Interactive	0.030	0.024	0.005	0.385
Video	0.125	0.098	0.027	1.021
Large Format	0.203	0.165	0.038	1.170

Before Law	Mean Non-EU	Mean EU	Difference	T-Test
Interactive	0.103	0.071	0.032	1.019
Video	0.035	0.009	0.026	1.459
Large Format	0.224	0.212	0.011	0.259

Selection Bias in respondent demographics

- 54% female (slightly more than general internet population of just under 50
- Similar to general internet population in other dimensions (income, time spent online, age), compared to DiMaggio and Bonikowski 2008)
- We do not have the data to say more. Lower bound of interpretation is that that we are measuring accurately the measure that advertisers use to assess advertising.
- We can say that the change is specific to European websites and does not represent a change in preferences of European consumer

Ruling out changing European time-trend

- Next explore whether our results are a result of changing European attitudes to online advertising (rather than a change in what is happening on European websites)
- Exploit the fact that Europeans can browse US (and other) websites and that Americans can browse European websites
- If unexplained heterogeneity in terms of attitudes of EU citizens towards advertising drives the results, then we expect
  - To see a similar collapse in ad effectiveness when they visit websites based in the US
  - To see no collapse in ad effectiveness when Americans visit European websites

Table: EU Survey Takers on non-EU Websites

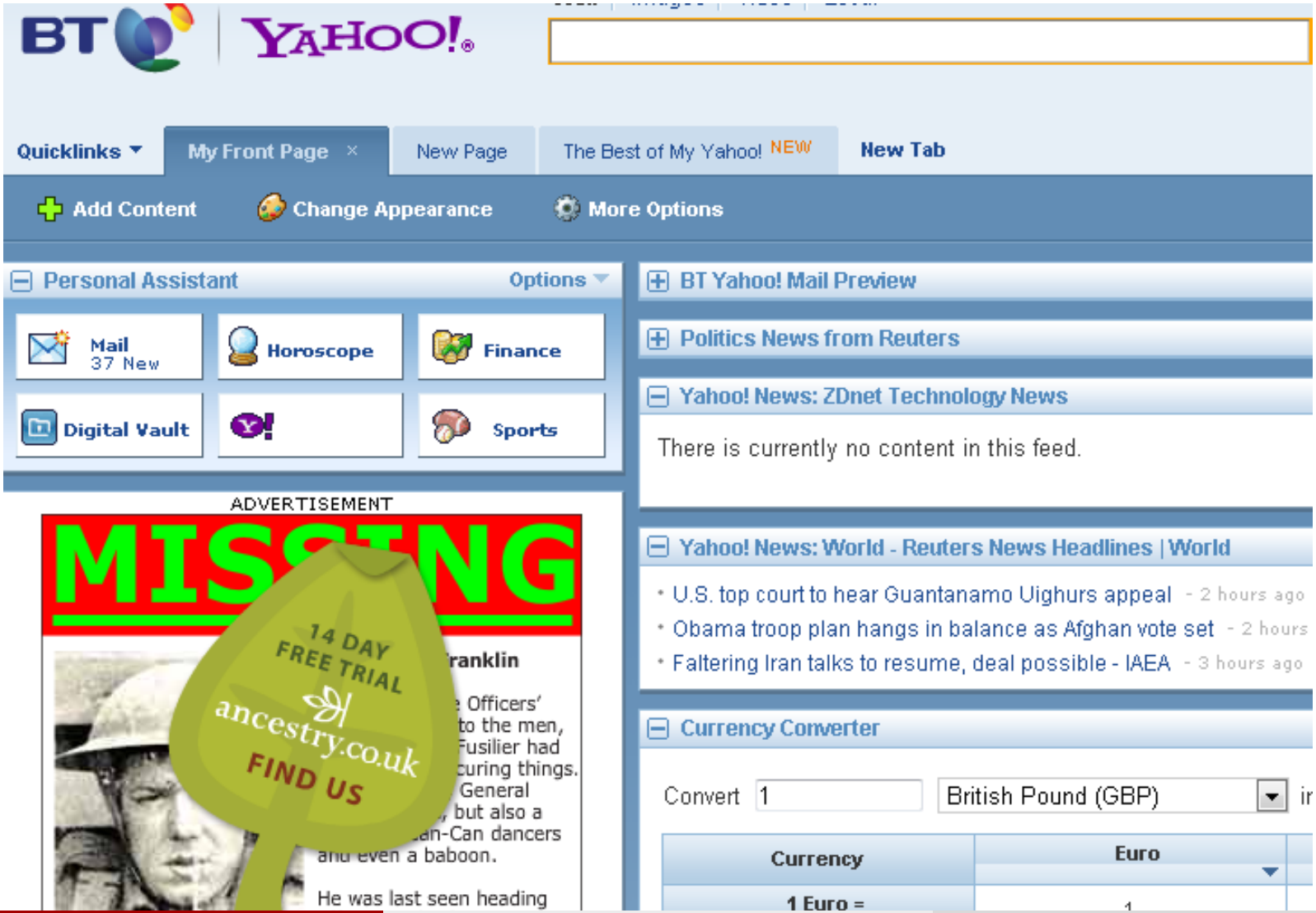
	Difference	T-Test
Before European Privacy Law	**-0.018**	-4.392
After European Privacy Law	**-0.030**	-19.372

Table: Non-EU Survey Takers on EU Websites

	Difference	T-Test
Before European Privacy Law	**-0.032**	-2.942
After European Privacy Law	**0.006**	0.458

Outline

How easy is it to advertise without customer data on this site?



How easy is it to advertise without customer data on this site?

	(1) General Content	(2) Product-Specific Content	(3) Health Site	(4) Parenting Site
Exposed × After EU Law × EU	-0.0589***	-0.00906	-0.281***	-0.00377
	(0.0195)	(0.00772)	(0.0552)	(0.0170)
Exposed	0.0576***	0.0219***	0.116***	0.0199*
	(0.0191)	(0.00697)	(0.00874)	(0.0117)
Exposed × Before EU Law	0.000524	0.00117	-0.00452	-0.00246
	(0.00421)	(0.00221)	(0.00873)	(0.00633)
Exposed × Not-EU	-0.0414**	-0.00683	-0.0905***	-0.00202
	(0.0190)	(0.00690)	(0.00801)	(0.0112)
Campaign Fixed Effects	Yes	Yes	Yes	Yes
Demographic controls	Yes	Yes	Yes	Yes
Observations	1037597	2292035	128956	213894
Log-Likelihood	-619959.7	-1374665.6	-77998.3	-135829.9

We then explore how ad format changed effect of law

	(1) Media Rich Ads	(2) Plain Banners	(3) Large Format Ads	(4) Small Format Ads
Exposed × After EU Law × EU	-0.0148	-0.0184**	0.00109	-0.0235**
	(0.0159)	(0.00771)	(0.0113)	(0.0105)
Exposed	0.0302**	0.0245***	0.0231***	0.0320***
	(0.0146)	(0.00682)	(0.00871)	(0.00973)
Exposed × After EU Law	0.000342	-0.00197	-0.00292	-0.00227
	(0.00338)	(0.00242)	(0.00366)	(0.00278)
Exposed × Not-EU	-0.0144	-0.00775	-0.00703	-0.0142
	(0.0150)	(0.00714)	(0.00907)	(0.0101)
Campaign Fixed Effects	Yes	Yes	Yes	Yes
Demographic controls	Yes	Yes	Yes	Yes
Observations	1098047	2231585	613804	2715828
R-Squared	0.163	0.178	0.156	0.176

There are of course limitations

1. Our data consist of firms that enlisted a particular marketing research company to test their advertising.
2. The ads in our data were all run through the firm rather than through an ad network.
3. Ad networks would likely have been more directly affected but they may also have dedicated more resources to mitigating the effect of the regulation through alternative strategies
4. We do not know if the change in effectiveness led to a change in revenues
5. We rely on stated expressions of purchase intent and not actual purchase data.

First study of how privacy regulation affects ad performance

- We find that privacy laws in Europe are associated with reduced ad effectiveness
  - It is not associated with a drop in the effect of ads on American websites to EU-based web surfers
- This drop is not neutral across website types and ad types.
  - May affect ability of 'broad-brush' websites to provide free content
  - May lead to more intrusive advertising
- This suggests that privacy regulation will likely play an important role in shaping future economic activity on the internet.
- Enacting privacy legislation, while potentially worthwhile, does involve trade-offs

Other Findings on Banner Advertising and Privacy

1. Advertising effectiveness (Forthcoming discussion paper in Marketing Science)
   - We show that obtrusive ads work and that targeted ads work, but that obtrusive AND targeted ads are not particularly effective
   - This appears to be related to privacy concerns and might explain the bifurcation of online ads into subtle targeted ads (e.g. AdSense) and obtrusive ads
   - Back-of-the-envelope calculations suggest online advertising is worthwhile (on average) if switching a prospective customer to high purchase intent is worth 42 cents.
2. Social Network and Privacy Control
   - Giving users control over their privacy settings can increase ad-effectiveness

	Logit (1) Purchase Intent	Scale (2) Intent Scale	Opinion (3) Favorable Opinion	Recall (4) Ad Recall
main				
Exposed × After EU Law × EU	-0.117**	-0.0275**	-0.0205***	-0.0312***
	(0.0487)	(0.0136)	(0.00686)	(0.0105)
Exposed	0.128***	0.0547***	0.0257***	0.103***
	(0.0430)	(0.0118)	(0.00592)	(0.00944)
After EU Law × EU	-0.0757			
	(0.110)			
Exposed × Not-EU	-0.0567	-0.00921	-0.00917	-0.0259***
	(0.0418)	(0.0124)	(0.00620)	(0.00997)
Exposed × Before EU Law	-0.00178			
	(0.0266)			
Before EU Law	-0.117***			
	(0.0340)			
Not EU	-0.104			
	(0.0880)			
Constant	-0.559***			
	(0.0902)			
Exposed × After EU Law		-0.00647	0.00187	-0.0267***
		(0.00410)	(0.00199)	(0.00342)
Campaign Fixed Effects	No	Yes	Yes	Yes
Demographic controls	Yes	Yes	Yes	Yes
Observations	3329632	3329632	3180804	3035292
R-Squared		0.200	0.185	0.121
Log-Likelihood	-2190792.7	-5640801.1	-1941938.0	-1624937.8

	Exposed 1x	Date Controls	No Latin-America	Country Controls
	(1)	(2)	(3)	(4)
	Purchase Intent	Purchase Intent	Purchase Intent	Purchase Intent
Exposed × After EU Law × EU	-0.0206***	-0.0166**	-0.0171**	-0.0162**
	(0.00719)	(0.00743)	(0.00714)	(0.00715)
Exposed	0.0268***	0.0249***	0.0263***	0.0263***
	(0.00633)	(0.00686)	(0.00635)	(0.00635)
Exposed × After EU Law	-0.000163		-0.00115	-0.00186
	(0.00209)		(0.00194)	(0.00194)
Exposed × Not-EU	-0.0138**	-0.00902	-0.00975	-0.00982
	(0.00660)	(0.00681)	(0.00658)	(0.00658)
Exposed × Before UK law		0.0141*		
		(0.00784)		
Exposed × Before Italy law		-0.0136*		
		(0.00796)		
Exposed × Before France law		0.00593		
		(0.00603)		
Exposed × Before Germany law		-0.00103		
		(0.00859)		
Exposed × Before Netherlands law		-0.00477		
		(0.00972)		
Exposed × Before Spain law		0.00493		
		(0.00512)		
Campaign Fixed Effects	Yes	Yes	Yes	Yes
Demographic controls	Yes	No	Yes	Yes