Cookies and Shopping

Bo Bian[*], Michaela Pagel[†], Huan Tang[‡], and Emily Williams[§]

Abstract

This study examines the impact of enhanced data privacy on online shopping. Specifically, we analyze the staggered adoption of cookie permission and compliance systems by US retailers in response to the European Union's data protection standards and the California Privacy Act. We combine information on when specific retailers implemented a cookie compliance system with individual-level bank and credit card transactions data from a US data aggregation and analytics provider. We find that online spending at retailers decreases significantly after they introduce cookie compliance systems. To address potential selection into who and when individuals shop with enhanced privacy, we also use exposure to cookie compliance systems, based on pre-policy shopping baskets, in a reduced form IV specification. Our IV specification results confirm the initial findings with respect to spending and show that individuals incur less overdraft and late fees, roll over less credit card debt, and borrow less in other high-interest unsecured credit such as payday loans when they are treated with enhanced privacy. We discuss targeted advertising, third-degree price discrimination, and shopping convenience as the three main channels behind these results. Our findings provide insights into the relationship between data privacy regulations and consumer actions, informing policy considerations at the state and federal levels.

Keywords: Data security, privacy regulation, web cookies, online shopping, financial well-being

JEL Codes: G5, G28, L86

[*] University of British Columbia, Sauder School of Business. E-mail: bo.bian@sauder.ubc.ca

[†] Columbia Graduate School of Business, NBER, and CEPR. E-mail: mpagel@columbia.edu

[‡] London School of Economics & CEPR. E-mail: huan.ht.tang@gmail.com

[§] Harvard Business School & CEPR. E-mail: ewilliams@hbs.edu

We thank Hunt Allcott, Simona Abis, Tania Babina, Markus Mobius, David Rothschild, Steffen Meyer, Charline Uhr, and Matt Backus for valuable comments as well as seminar participants at Barnard College, Aarhus University, UT Austin, OSU, UIUC, WashU Olin, and Stockholm School of Economics. We also thank Anton Tarasenko for providing outstanding research assistance.

1 Introduction

Over the past few decades, privacy has undergone profound changes driven by technological advances and evolving societal norms. The ubiquity of digital devices and online services has led to unprecedented data collection. At the same time, data storage capacities as well as analyzing capabilities have evolved at a similar pace. Increased individual identification and tracking have eroded anonymity and contribute to the complex interplay between convenience and privacy in the digital age.

At the same time, there has been a significant acceleration in electronic commerce and retail sales, which was further amplified during the Covid pandemic. Retail companies profit from private information by leveraging consumer data for targeted advertising, personalized pricing, and operational optimization. These practices, while beneficial for business, necessitate a delicate balance to address privacy concerns and comply with regulations. After all, targeted advertising can harm consumers by fostering manipulation, exploit vulnerabilities, and promote overspending. Price discrimination leads to a redistribution of consumer to producer surplus and can be perceived as unfair.

The most prominent data privacy regulations have specifically targeted web cookies to protect user privacy and ensure transparent and secure online experiences. The European Union's strengthened data protection standards in May 2018 (EU GDPR) required that websites must obtain clear and informed consent from users before storing or accessing cookies on their devices. The California Privacy Act (CAPRA) was enacted in January 2020 and established privacy standards for the collection of personal information, which includes data gathered through cookies.

In this paper, we analyze the staggered adoption of cookie compliance systems by US retailers in the years after these two regulations targeting cookies. Our study holds direct policy relevance, particularly at the federal level in the United States, where regulatory bodies such as the Federal Trade Commission (FTC) and the Consumer Financial Protection Bureau (CFPB) are currently considering the implementation of new rules to safeguard people's privacy and enhance data security.

Our objective is to evaluate and quantify the impact of strengthened data security standards on online shopping and individual financial health. To that end, we first explore how the introduction of cookie compliance systems by retailers, via cookie permission banners or pop-up windows, impact online shopping at these retailers compared to offline shopping. In turn, we assess whether the induced changes in spending translate into other measures of individual financial well-being.

To obtain information on the adoption of cookie compliance systems, we obtain data from the company BuiltWith, a web technology information profiler tool. BuiltWith gathered all information about the technologies used by websites through automated web crawling and code analysis since it was launched in August 2007. To become compliant with cookie privacy regulations, retailers typically implement a cookie consent mechanism on their website through third-party providers, who ensure that they comply with relevant regulations. In turn, BuiltWith detected the most commonly used of these third-party providers: OneTrust, Cookiebot, and Osano. These companies' cookie compliance solutions are designed to align with the EU GDPR and CAPRA. They offer customizable banners, consent management platforms, and granular consent options to ensure compliance. They also automate cookie scanning, record user consent, and provide preference centers for users to manage privacy settings.

The cookie compliance systems are a shock that enhances data privacy of customers because some consumers decline cookies, which hinders online retailers' ability to gather information and track them. According to survey data conducted by NordVPN, around 50% of customers do not consistently accept cookie permission requests. Additionally, the cookie compliance frameworks explicitly encourage data minimization which enhances data security even if individuals accepts all cookies.

We combine the information on the adoption of cookie compliance systems with data on online and offline shopping obtained from a transactions dataset of a prominent US data aggregation and analytics provider. This provider utilizes advanced data analytics to clean and categorize transactions data, which is then offered as a product to institutional investors and investment managers in aggregated and disaggregated forms. We use the de-identified transactions dataset, which includes bank and credit card transactions, as well as demographics data (income and geographical location) for an unbalanced panel of approximately 50 million active consumers spanning from January 2010 to August 2023.

We rely on merchant identifications provided by the data aggregator, which allows us to know the exact merchant. Additionally, we know whether a transaction was physical, which refers to an in-person payment with a card at a physical location such as in-store purchases at a point-of-sale terminal, or non-physical, which occur through digital or electronic means facilitated by technology and conducted remotely.

By examining the introduction of cookie permission policies at the retailer-day level and comparing online shopping with offline shopping as the control, we can analyze the effects of these policies on consumer behavior. With our large sample size and the substantial number of online retailers that implemented cookie permission policies at different times, we possess the capability to detect effects of cookie compliance systems even when they are modest.

We find that cookie permission banners reduce online spending at the retailer in question. To address the concern that certain types of people shop with enhanced data privacy at certain times, and that this selection drives our results, we perform a reduced form IV analysis. Our IV instrument is the exposure of each individual at each point in time to cookie compliance policies based on their pre-policy basket shares of any given retailer. We confirm our initial results and also estimate a positive effect on financial health, as measured by overdraft and late fees, credit card interest payments, and other high-interest unsecured loans such as payday loans, from exposure to enhanced privacy through cookie compliance systems.

2 Conceptual Framework and Related Literature

This study investigates the impact of data privacy, specifically the use of cookies for tracking users across the internet, on online shopping behavior. Potential channels are targeted advertising, facilitated by the collection and analysis of user data, and price discrimination. The analysis considers the broader implications of data privacy regulations, such as the GDPR and state-level privacy acts, as well as potential federal regulations for enhancing data security.

In the digital age, websites use cookies to track users' browsing activities, enabling the collection of valuable data on consumer behavior. These cookies, particularly third-party cookies, allow websites to target users with personalized ads based on their interests, preferences, and past online activities. Consumer privacy may affect spending through targeted advertising or price discrimination.

Targeted advertising, also known as personalized advertising or interest-based advertising, is a marketing strategy that aims to deliver relevant and tailored advertisements to specific groups of individuals or individual consumers. This approach uses data and information about a user's interests, behaviors, demographics, and online activities to serve them with ads that are more likely to be of interest to them. Targeted advertising relies on data collection and tracking, which can raise privacy concerns. Many websites and platforms ask for user consent to use cookies and similar tracking technologies for ad targeting, in compliance with data protection regulations such as the GDPR or the CCPA. Users typically have the option to opt-out of targeted advertising or manage their ad preferences through browser settings or privacy preference centers.

Third-degree price discrimination is referred to as charging different prices to different groups of customers based on various factors like age, location, or income, is generally not illegal. In many jurisdictions, it is considered a common business practice and not explicitly prohibited by law. However, some forms of price discrimination may be subject to regulations and laws in certain contexts. For example:

- Antitrust Laws: If a company engages in price discrimination with the intent to create a monopoly or to restrict competition, it may be in violation of antitrust laws. Antitrust laws aim to promote fair competition and prevent practices that harm consumers or stifle market competition.

- Discrimination Laws: Price discrimination based on certain protected characteristics, such as race, gender, religion, or nationality, could be considered discriminatory and may be illegal under civil rights or anti-discrimination laws in some jurisdictions.

- Consumer Protection Laws: Some countries have consumer protection laws that govern pricing practices to ensure transparency and fairness for consumers. Price discrimination that involves deceptive practices or unfair treatment of consumers may be subject to legal scrutiny.

- Price Gouging Laws: During certain emergency situations, such as natural disasters or public health crises, some jurisdictions may have laws against price gouging, which is an extreme form of price discrimination that takes advantage of consumers in vulnerable situations.

Overall, the legality of third-degree price discrimination depends on the specific context and the laws of the country or region in which the business operates.

Johnson (2013) shows that online retailers are willing to pay 52% more for ads with third-party cookies which is the same estimate as one provided by a google blog (Bindra, 2019) in 2019. Wernerfelt et al. (2022) study the extent to which advertisers benefit from data that are shared across applications. They focus on one of the most common ways advertisers use offsite data and run a large-scale study with hundreds of thousands of advertisers on Meta. Within campaigns, we experimentally estimate both the effectiveness of advertising under business as usual, which uses offsite data, as well as how that would change under a loss of offsite data. They find a median cost per incremental customer using business as usual targeting techniques of $43.88 that under the median loss in effectiveness would rise to $60.19, a 37% increase. Todri (2022) show that ad-blockers have a significant effect on online purchasing behavior: online consumer spending decreases due to ad-blockers by approximately $14.2 billion per year in total. Aridor et al. (2021) show that the GDPR resulted in a 12.5% drop in the online-travel-intermediary-observed consumers, whereas the remaining consumers were trackable over a longer time. The average value of the remaining consumers to advertisers increased which offset the losses from the consumer opt-outs. Zhao et al. (2021) investigates the impact of the GDPR on consumers' online browsing and search behavior using consumer panels from four countries, finding evidence consistent with higher frictions in online search. Berman and Israeli (2021) show that the adoption of data analytics impact retailers' performance. The authors exploit the staggered adoption of a retail analytics service by more than 1,000 e-commerce websites and find an average increase of 8-29% in monthly revenues post adoption.

Dubé and Misra (2023) study third-degree price discrimination implemented with machine learning for a large, digital firm. Their results reveal unexercised market power that allows the firm to raise its price optimally, generating a 55% increase in profits. Personalized pricing improves the firm's expected posterior profits by an additional 19%, relative to the optimized uniform price, and by 86%, relative to the firm's unoptimized status quo price. Turning to welfare effects on the demand side, total consumer surplus declines 23% under personalized pricing relative to uniform pricing, and 47% relative to the firm's unoptimized status quo price.

We also situate our work within the context of recent research on the effects of data privacy regulations. In a related paper, Bian et al. (2023) raise concerns about consumer privacy and the potential for financial fraud due to data breaches.

Our paper adds to recent work that examine the effects of data privacy regulation. Lukic et al. (2023) determine whether GDPR's enforcement increased consumers' online privacy by decreasing the amount of online tracking. Other existing research has linked GDPR to European web traffic (Goldberg et al., 2019), the entry and exit of apps (Janssen et al., 2021), VC financing (Jia et al., 2021), and the ability of firms to collect and monetize consumer data (Aridor et al., 2020; Bessen et al., 2020; Peukert et al., 2021). Babina et al. (2022) show that open banking policies spur investments into FinTech startups. A couple of recent studies focus on Apple's privacy initiatives, including the privacy label policy and the App Tracking Transparency (ATT) policy. Bian et al. (2021) show that Apple's privacy labels lead to a 14% weekly download reduction and a 15% decline in revenue from user subscriptions and in-app purchases for iPhone users (using Android users as the control group). In addition, the ATT policy leads to an immediate negative stock market reaction for public firms with apps. Kesler (2022) show that the ATT framework implemented by Apple leads more apps to become paid apps and turn to in-app purchases as an alternative revenue source.

3 Empirical Analysis

3.1 Privacy Policy Changes

Companies implement privacy policies and cookie permissions to comply with changing regulations and to address evolving privacy concerns. To comply with the EU GDPR and CAPRA in the domain of web cookies, companies typically partner with a cookie compliance system provider, such as OneTrust, to help them set up the cookie compliance infrastructure. We obtained our data on cookie permission systems using data from the company BuiltWith. BuiltWith is a web technology profiler company. It provides a web service that helps businesses and individuals discover the technology stack used by various websites, such as the content management system (CMS), web hosting provider, analytics tools, advertising platforms, e-commerce platforms, and more. BuiltWith uses automated data analysis and crawling techniques to gather information about websites and their underlying technologies. We collected the dates that BuiltWith first detected the most common cookie compliance system providers: OneTrust, Osano, and Cookie Bot. These companies provide solutions related to data privacy and compliance, particularly concerning cookies and online tracking.

As an example, McDonald partnered with OneTrust since August 9, 2019 and has a privacy policy in place that outlines how they collect, use, and protect user data. When a user visits the McDonald's website, they would ask for your consent to use cookies and similar tracking technologies. The specific cookie permissions include:

- Necessary Cookies: These are essential for the website to function correctly, and they do not require explicit consent.

- Analytics Cookies: These cookies collect anonymous data about how users interact with the website, helping McDonald's improve its site's performance and user experience.

- Targeting Cookies: These cookies track users' browsing habits to provide personalized ads, promotions, and offers based on their interests and behavior.

Users are typically presented with a cookie banner or pop-up when they first visit the website, and they can manage their cookie preferences by accepting or declining specific cookie categories. Figure 1 provides screenshots of McDonald's privacy center.

Barnes & Noble has a very similar cookie policy, also sponsored by OneTrust, as McDonalds, as can be seen in Figure 2.

3.2 Transaction-Level Bank Account Data

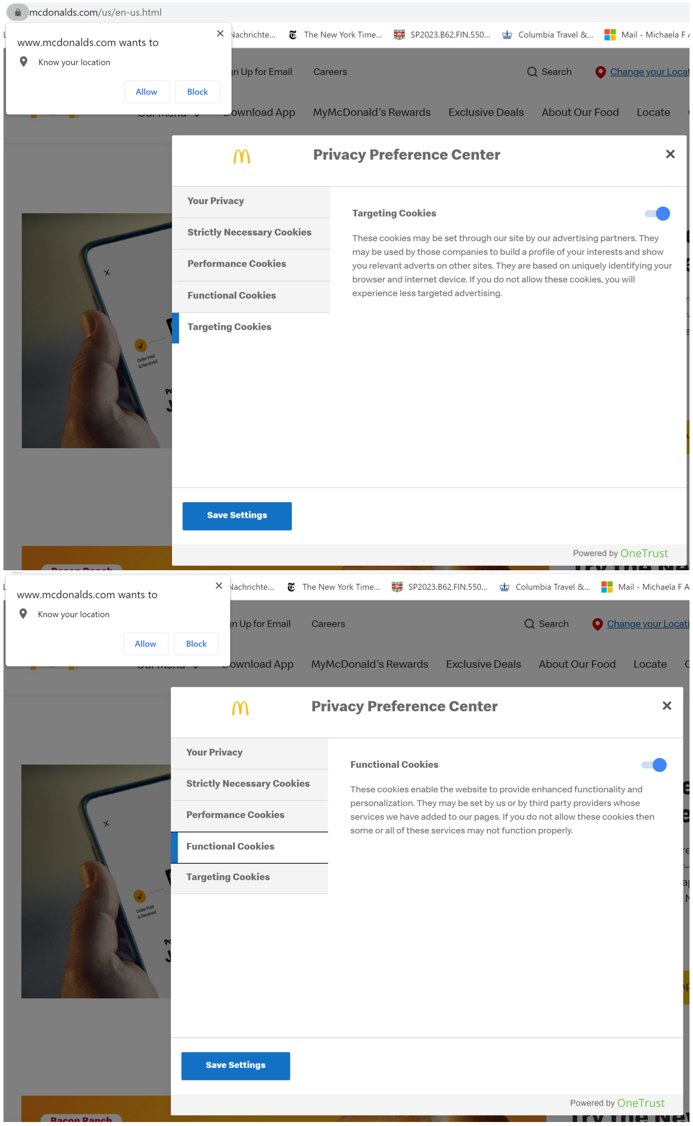
We combine our information on cookie compliance introduction dates at the retailer level with transaction-level data of individual bank and credit card accounts provided by a financial aggregation and analytics firm. The dataset contains all income and spending transactions from a number of bank and credit card accounts per individual, it also includes information on overdraft fees, savings income, as well as other measures of financial health.

3.3 Main Empirical Analyses

We use the date of the policy implementation as staggered treatments and online versus offline shopping as treatment intensity. Our main regression specification is:

$$SpendPerTransaction_{i,m,t} = \alpha_i + \alpha_t + \beta PostPolicy_{m,t} \times Online_{i,m,t} + \varepsilon_{i,m,t}$$

where i stands for individual, m for merchant, and t for time (day-of-sample). We use the raw data at the transaction-level in this regression, i.e., some individual i may have none, one, or multiple transactions at retailer m on day t. When and We control for individual and day-of-sample fixed effects. We cluster standard errors at individual, merchant, and monthly levels.



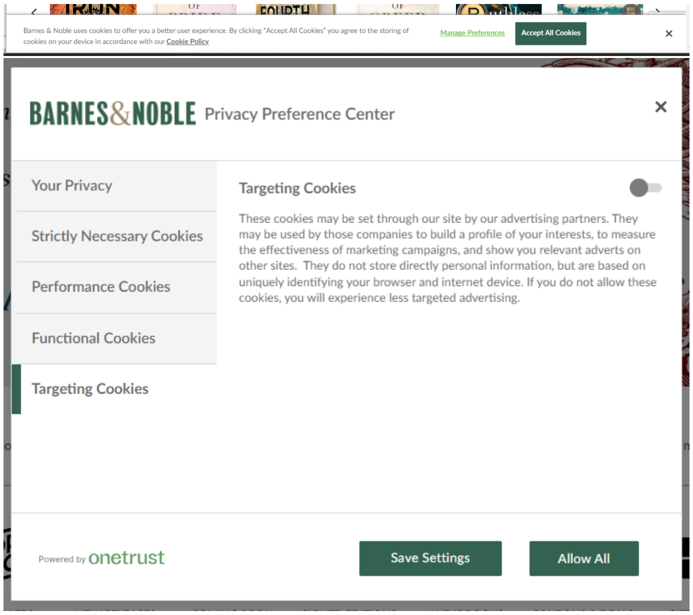
Targeting Cookies

These cookies may be set through our site by our advertising partners. They may be used by those companies to build a profile of your interests and show you relevant adverts on other sites. They are based on uniquely identifying your browser and internet device. If you do not allow these cookies, you will experience less targeted advertising.

Functional Cookies

These cookies enable the website to provide enhanced functionality and personalization. They may be set by us or by third party providers whose services we have added to our pages. If you do not allow these cookies then some or all of these services may not function properly.

Figure 1: McDonald's privacy center.



Barnes & Noble uses cookies to offer you a better user experience. By clicking “Accept All Cookies” you agree to the storing of cookies on your device in accordance with our Cookie Policy

BARNES&NOBLE Privacy Preference Center

Targeting Cookies

These cookies may be set through our site by our advertising partners. They may be used by those companies to build a profile of your interests, to measure the effectiveness of marketing campaigns, and show you relevant adverts on other sites. They do not store directly personal information, but are based on uniquely identifying your browser and internet device. If you do not allow these cookies, you will experience less targeted advertising.

Figure 2: Barnes & Noble’s privacy center.

3.4 Selection and Privacy Exposure

The prevailing endogeneity concern in this analysis is that certain types of individuals might select into shopping at retailers with or without privacy policies (at certain points in time). To overcome these selection problems, we use the pre-privacy-policy basket shares of individuals to calculate exposure-to-enhanced-privacy-policies variable. We then use the exposure variable in a reduced-form IV regression.

$$SpendPerTransaction_{i,m,t} = \alpha_i + \alpha_t + \beta EnhPrivacy_{i,t} \times Online_{i,m,t} + \varepsilon_{i,m,t}$$

where i, m, and t denote individual, merchant, and day-of-sample as above. Again we control for individual and day-of-sample fixed effects and cluster standard errors at individual, merchant, and monthly levels.

We then aggregate the data and obtain a balanced panel in which we can fill in zeros, on days when individuals do not spend. We aggregate the data to the individual-day, pre-post cookie compliance, and online/offline levels. I.e., for each individual i on day t, we sum up all transactions pre-cookie compliance policy that are either online or offline as well as all transactions post-cookie compliance that are either online and offline. We then fill in zeros to obtain a balanced panel and run the following specification:

$$Spend_{i,t} = \alpha_i + \alpha_t + \beta PostPolicy_{i,t} \times Online_{i,t} + \varepsilon_{i,t}$$

where i and t denote individual and day-of-sample as above. Again we control for individual and day-of-sample fixed effects and cluster standard errors at individual and monthly levels.

3.5 Summary Statistics

Table 1 provides summary statistics of the transaction-level data.

	Mean	Standard deviation	Median	5th percentile	25th percentile	75th percentile	95th percentile
All retail spending	66.50	723.9	20.65	2.560	8.980	50	203.1
Online spending	127.3	1299.3	29.97	2.180	10.77	95	416.7
Restaurant spending	24.56	263.7	12.58	2.940	7.080	25	65.89
Grocery spending	36.08	108.2	19.10	1.750	7.580	40.83	119.3
Regular income	1509.3	2101.2	1119.8	79	476.0	1931.8	3872.0
Other income	1097.3	4108.0	200	3.880	41	832	3617.4
Interest income	11.92	402.5	0.520	0.01000	0.0500	1.380	4.700
Overdraft fees	95.61	514.5	36	7	35	36	297.5
Late fees	27.55	7.653	25	15	25	35	39
NSF fees	60.97	204.9	34	5	13.87	42.39	200
Interest charges	46.95	62.08	24.60	0.660	7.500	66.80	159.2
Payday/EWA/FinTech loans	1025.0	3036.2	300	8.630	75	1000	3535.2
Transactions per user per month	129.6	96.17	115	31	76	164	268
Number of users per month	2102.3	1220.7	2105	228	1007	3178	4002
Observations (total)	20,835,241						

Table 1: Full sample, individual-transaction level, January 2014 to August 2023.

Table 2 displays shopper characteristics, their income as well as overall spending at all retailers, in the two weeks before and after any retailer introduced a cookie compliance framework. The p-values of the differences in means are adjusted for multiple hypothesis testing using a Bonferroni correction. We can see that the differences in means of all characteristics are economically small.

	Pre privacy policy		Post privacy policy			
	Mean	Standard error	Mean	Standard error	Difference	P-value
All retail spending	70.56	0.73	66.17	0.22	4.39	0.00
Online spending	127.14	2.13	127.32	0.76	-0.18	16.83
Restaurant spending	24.44	0.21	24.57	0.16	-0.14	10.93
Grocery spending	36.26	0.33	36.06	0.07	0.20	10.05
Regular income	1585.24	11.43	1502.72	3.18	82.52	0.00
Other income	1121.76	42.36	1094.85	13.25	26.91	9.80
Interest income	19.42	2.56	11.46	0.53	7.96	0.04
Savings income	107.16	14.82	94.73	2.80	12.43	7.38
Overdraft fees	28.26	0.35	27.50	0.10	0.76	0.70
NSF fees	56.53	3.77	61.38	1.63	-4.86	4.27
Late fees	54.60	0.93	46.36	0.23	8.24	0.00
Interest charges	1194.19	92.98	1008.61	27.57	185.58	1.01

Table 2: Full sample, individual-transaction level, January 2014 to August 2023, comparison between shopper characteristics two weeks before and after retailers' privacy policies. The p-values of the differences are adjusted for multiple hypotheses testing using a Bonferroni adjustment. The spending data consists of all identified companies including the one with the cookie policy change.

Figure 3 shows the share of transactions done at retailers that have privacy policies and cookie compliance systems in place.

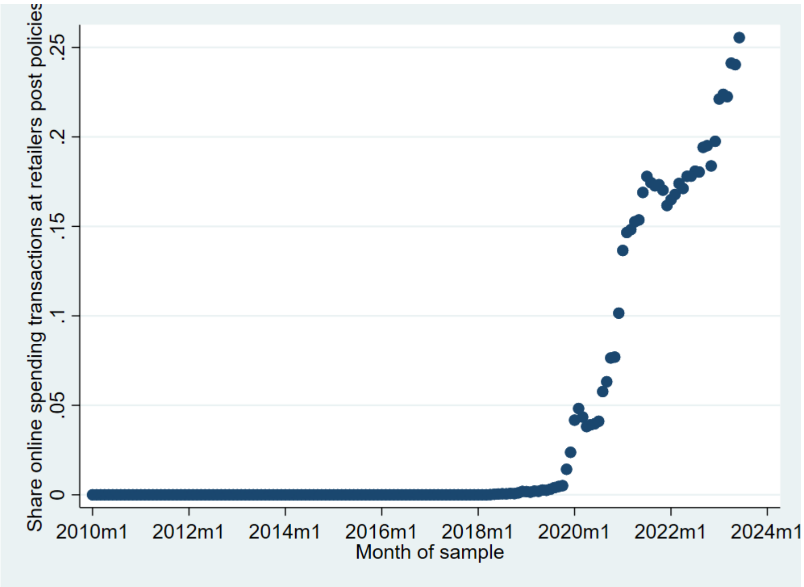


Figure 3: Share of online retail transactions at retailers post their privacy policy.

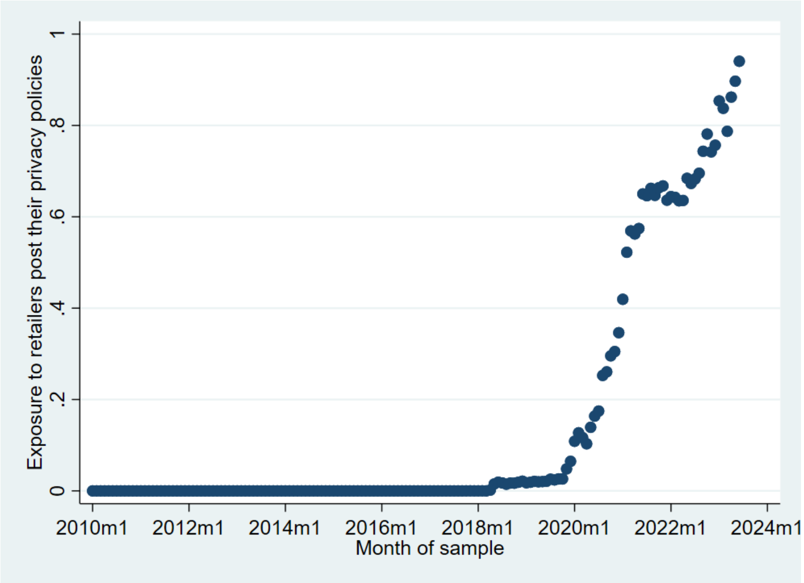


Figure 4: Exposure to online retailers' privacy policies (basket share times indicator of privacy policy implementation).

4 Results

4.1 Main Regression Results

The analysis reveals that the restriction of cookies, targeted advertising, and consumer profiling significantly decreases consumer spending. This is consistent with targeted advertising or price discrimination, where retailers leverage user data to charge higher prices to specific segments of consumers based on their perceived willingness to pay.

We find decreases in spending in both the OLS and IV analyses, using exposure to privacy as the (treatment) instrument.

Table 3 uses the raw transactions sample without any aggregation steps.

Tables 4 and 5 break down the main regression by merchant categories. We can see that the negative effects are concentrated in more retail spending such as electronics, merchandise, home improvement, and travel. In contrast, in categories such as restaurants or groceries, we do not see effects. This is consistent with our prior that price discrimination and targeted advertising are more prevalent in more discretionary rather than necessary spending.

	(1) Spending winsorized	(2) Spending winsorized	(3) Spending log	(4) Spending winsorized	(5) Spending log
Post cookie policy =1	-10.960***	0.123	0.078***		
	(1.089)	(0.629)	(0.008)		
Online =1	61.894***	47.594***	0.355***	47.722***	0.360***
	(1.523)	(2.005)	(0.013)	(2.006)	(0.012)
Post cookie policy =1 × Online=1	-21.639***	-13.930***	-0.104***		
	(3.985)	(2.178)	(0.017)		
Enhanced privacy				-0.497	0.026***
				(0.377)	(0.004)
Online =1 × Enhanced privacy				-4.529***	-0.055***
				(0.798)	(0.007)
Individual FEs	✓	✓	✓	✓	✓
Day-of-sample FEs	✓	✓	✓	✓	✓
Merchant FEs		✓	✓	✓	✓
Observations	11,841,351	11,841,264	11,841,264	11,841,264	11,841,264
Adjusted R^2	0.10	0.41	0.48	0.41	0.48

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Full sample, main regression, individual, day-of-sample, and merchant fixed effects, cluster at individual, month, and merchant levels.

	(1)	(2)	(3)	(4)	(5)	(6)
	Spending winsorized	Spending winsorized	Spending winsorized	Spending winsorized	Spending winsorized	Spending winsorized
	Electronics general merchandise	Personal and family	Home improvement	Travel	Restaurants	Groceries
Post cookie policy =1	3.783**	-3.118**	8.233	7.472***	-1.466***	4.454**
	(1.725)	(1.378)	(11.175)	(1.434)	(0.332)	(1.719)
Online =1	12.032***	6.510***	65.662***	14.943***	13.601***	25.359***
	(2.043)	(1.792)	(7.220)	(4.351)	(2.068)	(7.428)
Post cookie policy=1 × Online=1	-29.960***	-2.427	-3.697	-5.928***	-4.358**	6.371
	(4.885)	(3.388)	(11.579)	(1.533)	(2.096)	(5.751)
Individual FEs	✓	✓	✓	✓	✓	✓
Day-of-sample FEs	✓	✓	✓	✓	✓	✓
Merchant FEs	✓	✓	✓	✓	✓	✓
Observations	1,967,379	556,298	311,438	766,902	2,588,443	1,953,712
Adjusted R^2	0.13	0.34	0.30	0.25	0.29	0.31

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Full sample, main regression, individual and day-of-sample fixed effects, cluster at individual, month, and merchant levels, broken down by spending in categories.

	(1)	(2)	(3)	(4)	(5)	(6)
	Spending	Spending	Spending	Spending	Spending	Spending
	winsorized	winsorized	winsorized	winsorized	winsorized	winsorized
	Electronics general merchandise	Personal and family	Home improvement	Travel	Restaurants	Groceries
Enhanced privacy	3.149***	0.001	-2.183	0.914	-0.434**	0.167
	(0.592)	(0.766)	(2.311)	(0.717)	(0.202)	(0.307)
Online =1	11.723***	7.279***	65.181***	14.580***	13.988***	25.494***
	(1.915)	(1.811)	(7.281)	(4.343)	(2.190)	(7.818)
Online=1 × Enhanced privacy	-8.528***	-5.369***	2.327	-1.334	-2.468**	-0.450
	(1.043)	(1.102)	(2.627)	(0.870)	(1.131)	(2.498)
Individual FEs	✓	✓	✓	✓	✓	✓
Day-of-sample FEs	✓	✓	✓	✓	✓	✓
Merchant FEs	✓	✓	✓	✓	✓	✓
Observations	1,967,379	556,298	311,438	766,902	2,588,443	1,953,712
Adjusted R^2	0.13	0.34	0.30	0.25	0.29	0.31

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Full sample, main regression, individual and day-of-sample fixed effects, cluster at individual, month, and merchant levels, broken down by spending in categories.

In order to understand whether the decrease in spending is associated with improvements in financial health, we look at other outcomes, using exposure to privacy as the (treatment) instrument. We find that the likelihood of overdraft fees, the payments made to credit cards, as well as the rolled-over credit card debt, as measured by interest charges, is reduced. This is consistent with lax privacy standard leading to overspending, because of targeted advertising or price discrimination. Again, Table 6 uses the raw transactions sample without any aggregation steps.

	(1)	(2)	(3)	(4)	(5)	(6)
	Overdraft	Late	NSF	Interest	Payday/EWA	Checking account
	fees	fees	fees	charges	/FinTech loans	balance log
Enhanced privacy	-0.052***	-0.003***	-0.005	-0.062***	-0.318***	0.020
	(0.015)	(0.000)	(0.008)	(0.008)	(0.083)	(0.016)
Mean of dependent variable	0.15	0.01	0.05	0.18	0.65	8.40
Individual FEs	✓	✓	✓	✓	✓	✓
Day-of-sample FEs	✓	✓	✓	✓	✓	✓
Observations	20,834,520	20,834,520	20,834,520	20,834,520	20,834,520	4,612,816
Adjusted R^2	0.00	0.00	0.01	0.01	0.01	0.83

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Full sample, exposure regression, individual and day-of-sample fixed effects, cluster at individual and month levels, financial health outcomes.

In turn, we aggregate the data to the offline/online pre/post cookie compliance system, date, and individual level. This allows us to fill in zeros when an individual did not spend at any retailer offline/online pre/post their cookie compliance system. Tables **??** and 8 show the same results in the aggregated data sets.

	(1)	(2)	(3)	(4)	(5)	(6)
	Spending	Number of	Spending	Spending	Number of	Spending
	winsorized	transactions	deviation from mean	winsorized	transactions	deviation from mean
Enhanced privacy	47.083***	1.623***	2.575***	48.508***	1.674***	2.595***
	(0.758)	(0.034)	(0.072)	(0.789)	(0.037)	(0.077)
Online =1	-15.560***	-0.501***	-1.176***	-15.605***	-0.503***	-1.175***
	(0.354)	(0.013)	(0.025)	(0.354)	(0.013)	(0.025)
Online=1 × Enhanced privacy	-33.395***	-1.079***	-1.678***	-33.164***	-1.072***	-1.655***
	(0.902)	(0.043)	(0.074)	(0.898)	(0.043)	(0.072)
Mean of dependent variable	13.61	0.44	1.00	13.61	0.44	1.00
Individual FEs	✓	✓	✓			
Day-of-sample FEs	✓	✓	✓			
Individual times month-of-sample FEs				✓	✓	✓
Day-of-week FEs				✓	✓	✓
Week-of-month FEs				✓	✓	✓
Observations	18,632,633	18,632,633	18,612,439	18,632,326	18,632,326	18,612,133
Adjusted R^2	0.18	0.25	0.01	0.20	0.27	0.02

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Full sample, aggregate to online/offline pre/post-privacy policies levels, individual and day-of-sample fixed effects, cluster at individual and month levels, include zeros, spending outcomes.

	(1)	(2)	(3)	(4)	(5)	(6)
	Overdraft	Late	NSF	Interest	Payday/EWA	Checking account
	fees	fees	fees	charges	/FinTech loans	balance log
Enhanced privacy	-0.185**	-0.010***	0.041	-0.158***	-0.300	0.020
	(0.073)	(0.002)	(0.044)	(0.037)	(0.423)	(0.016)
Mean of dependent variable	0.70	0.04	0.24	0.80	2.93	8.40
Individual FEs	✓	✓	✓	✓	✓	✓
Day-of-sample FEs	✓	✓	✓	✓	✓	✓
Observations	4,613,414	4,613,414	4,613,414	4,613,414	4,613,414	4,613,414
Adjusted R^2	0.01	0.01	0.04	0.04	0.01	0.83

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Full sample, aggregate to online/offline pre/post-privacy policies levels, exposure regression, individual and day-of-sample fixed effects, cluster at individual and month levels, financial health outcomes.

4.2 Dynamics

We then look at dynamics to study how online and offline spending evolved around the each-retailer-specific cookie permission introduction dates. We first use the raw transactions data, winsorized or divided-by-individual-means spending transactions and include individual, day-of-sample, and merchant fixed effects.

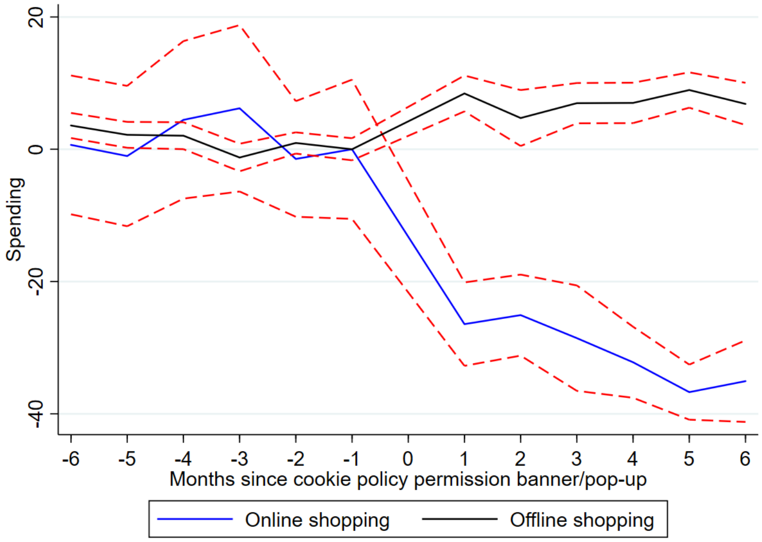


Figure 5: Full sample, main regression, winsorized transactions, individual, day-of-sample, and merchant fixed effects, cluster at individual and month levels.

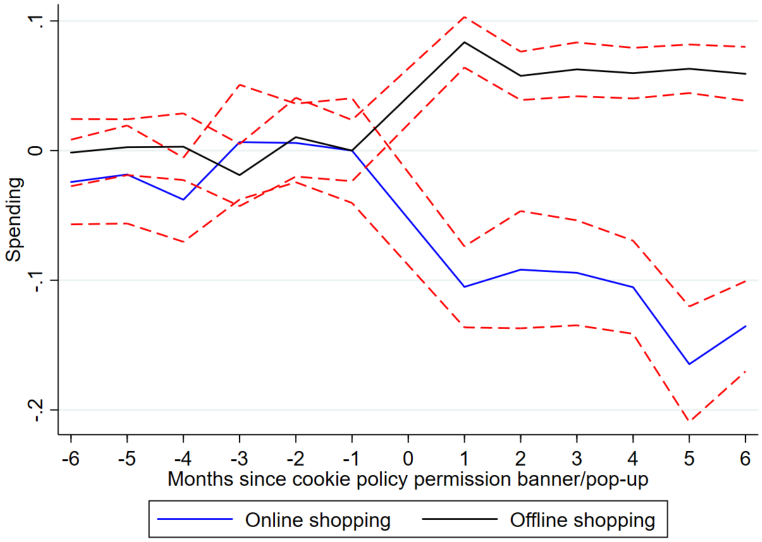


Figure 6: Full sample, main regression, log spending, individual, day-of-sample, and merchant fixed effects, cluster at individual and month levels.

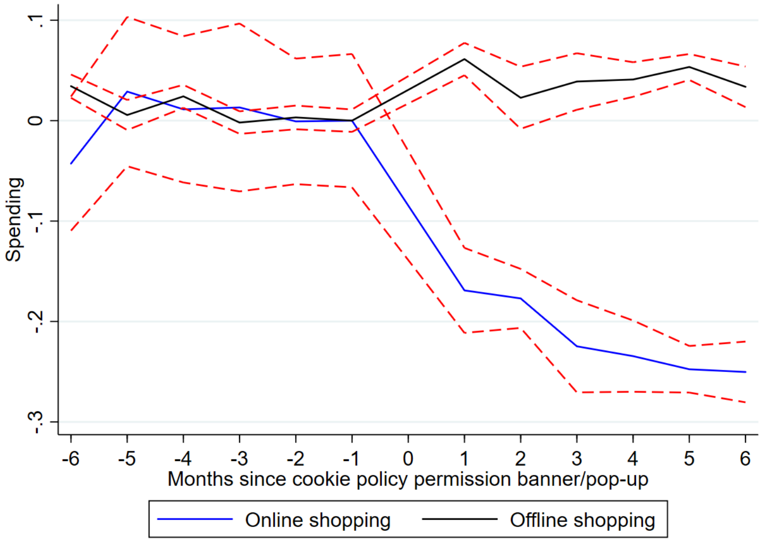


Figure 7: Full sample, main regression, deviations relative to individual-level means, individual, day-of-sample, and merchant fixed effects, cluster at individual and month levels.

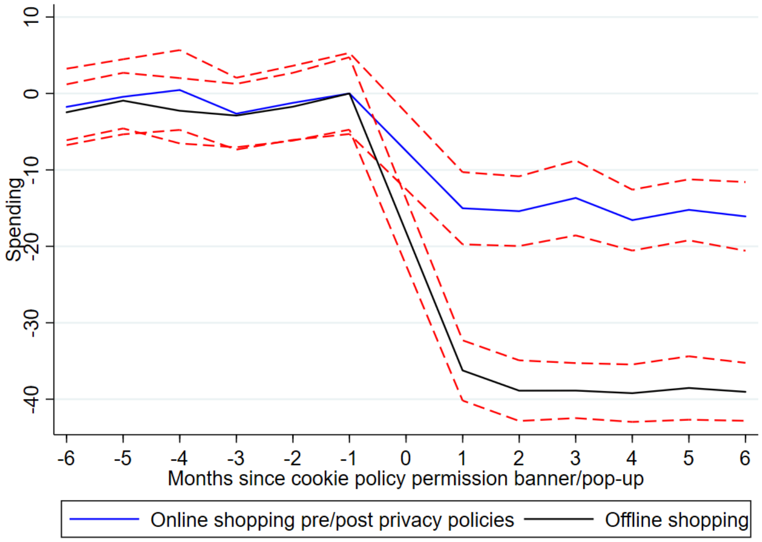


Figure 8: Full sample, aggregated to online/offline pre/post policies levels including zeros, winsorized transactions, individual and day-of-sample fixed effects, cluster at individual and month levels.

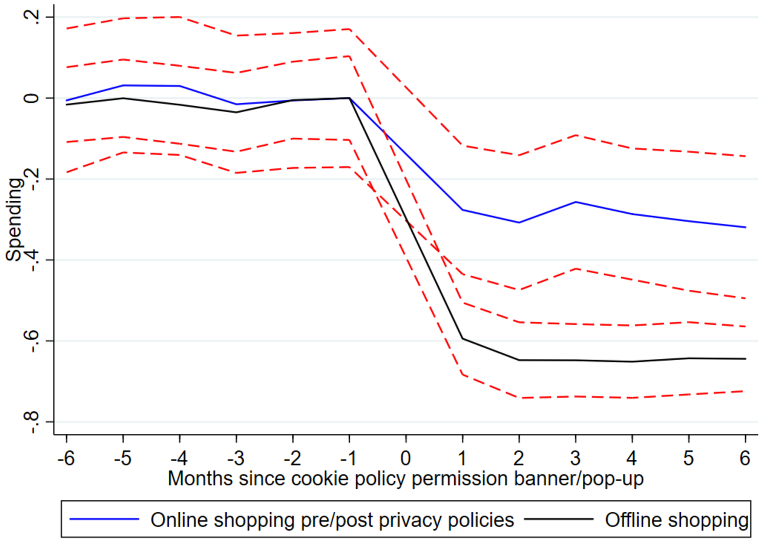


Figure 9: Full sample, aggregated to online/offline pre/post policies levels including zeros, log spending, individual and day-of-sample fixed effects, cluster at individual and month levels.

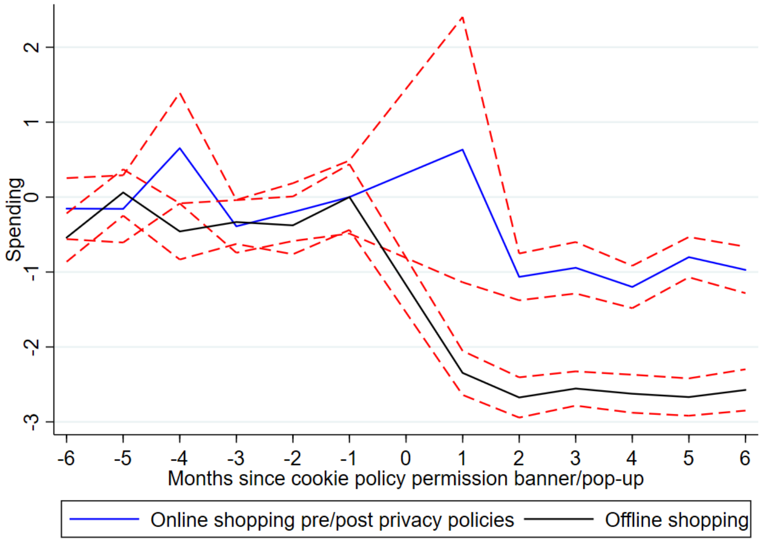


Figure 10: Full sample, aggregated to online/offline pre/post policies levels including zeros, deviations relative to individual-level means, individual and day-of-sample fixed effects, cluster at individual and month levels.

4.3 Robustness

To address the concern that companies introduced cookie compliance systems at the same time as other website features that affect online shopping, we show that common other feature additions are not correlated with the addition of cookie compliance systems in Figure 11, for features that would likely affect online shopping, and Figure 12, for features that are less likely to affect online shopping. All of the subgraphs in these figures have the same range of the x- and y-axes, if the dates of feature additions would be correlated, we would expect a lot on the 45-degree line, which is not what we see.

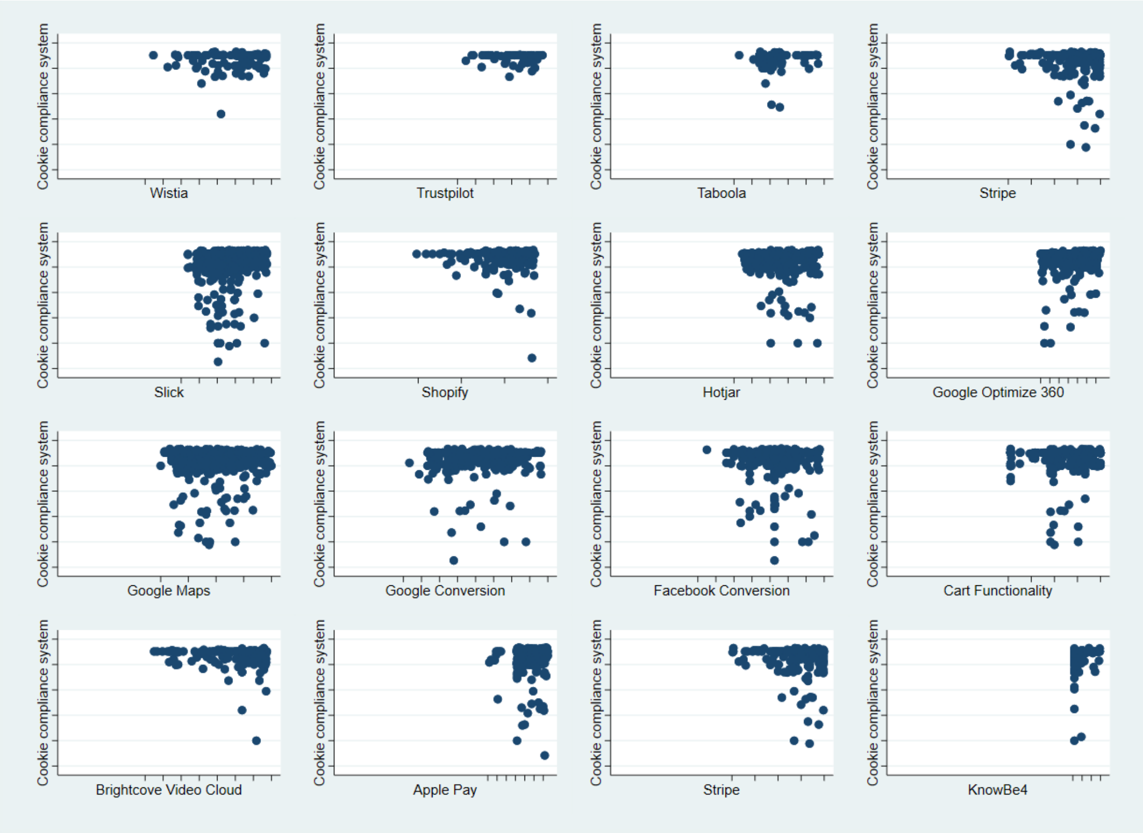


Figure 11: Dates of the additions of a cookie compliance system with the addition of other website features that may also affect online shopping. The date range on the x- and y-axes are the same, from May 2001 June 2023.

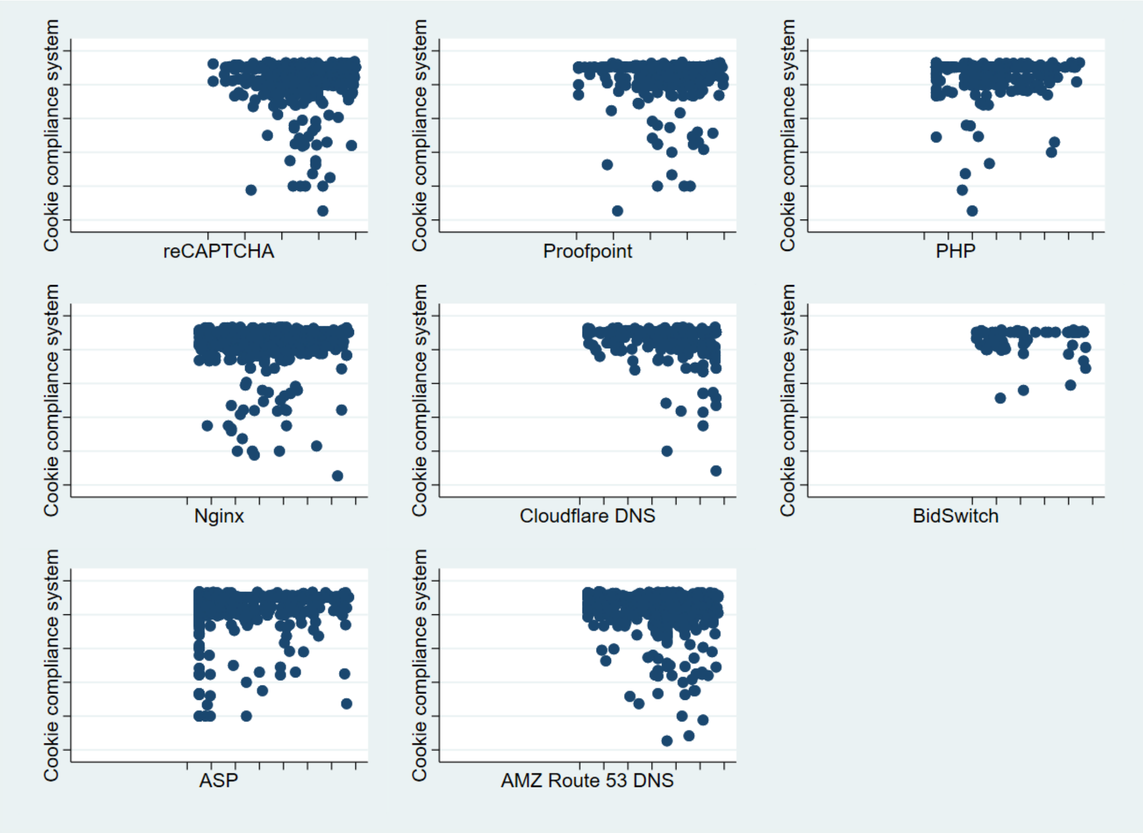


Figure 12: Dates of the additions of a cookie compliance system with the addition of other website features that are unlikely to affect online shopping. The date range on the x- and y-axes are the same, from May 2001 June 2023.

We also split the sample by regular income and time in Tables 9, 10, 11. Finally we go back to the raw transactions data and split by amounts in Table 12. We note here that the effects manifest in medium and larger transactions.

	(1)	(2)	(3)	(4)	(5)	(6)
	Spending	Spending	Spending	Spending	Spending	Spending
	winsorized	deviation from mean	winsorized	deviation from mean	winsorized	deviation from mean
	1st income tercile		2nd income tercile		3rd income tercile	
Post cookie policy=1	-1.317*	-0.000	0.963	0.007	-0.625	-0.018***
	(0.743)	(0.006)	(0.945)	(0.006)	(1.378)	(0.006)
Online=1	35.862***	0.223***	48.784***	0.260***	59.844***	0.216***
	(2.276)	(0.014)	(3.378)	(0.015)	(3.783)	(0.013)
Post cookie policy=1 × Online=1	-8.200***	-0.051**	-14.114***	-0.058**	-22.195***	-0.043
	(2.639)	(0.025)	(2.834)	(0.023)	(4.676)	(0.038)
Median of dependent variable	49.73		57.01		76.20	
Individual FEs	✓	✓	✓	✓	✓	✓
Day-of-sample FEs	✓	✓	✓	✓	✓	✓
Merchant FEs	✓	✓	✓	✓	✓	✓
Observations	3,692,370	3,692,370	3,822,543	3,822,543	3,701,817	3,701,817
Adjusted R^2	0.40	0.21	0.40	0.24	0.47	0.12

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Full sample, main regression, individual, day-of-sample, and merchant fixed effects, cluster at individual, month, and merchant levels, split by regular income, spending outcomes.

	(1) Spending winsorized	(2) Spending deviation from mean	(3) Spending winsorized	(4) Spending deviation from mean	(5) Spending winsorized	(6) Spending deviation from mean
	1st time tercile		2nd time tercile		3rd time tercile	
Post cookie policy =0	0.000	0.000				
	(.)	(.)				
Online =1	58.187***	0.275***	56.784***	0.281***	37.813***	0.189***
	(2.534)	(0.012)	(2.707)	(0.013)	(2.403)	(0.011)
Post cookie policy=0 × Online=1	0.000	0.000				
	(.)	(.)				
Post cookie policy =1			3.052*	0.021	0.690	0.003
			(1.664)	(0.014)	(0.712)	(0.004)
Post cookie policy=1 × Online=1			-20.729	-0.323	-9.994***	-0.034**
			(14.705)	(0.198)	(2.027)	(0.014)
Median of dependent variable	57.10		59.28		67.36	
Individual FEs	✓	✓	✓	✓	✓	✓
Day-of-sample FEs	✓	✓	✓	✓	✓	✓
Merchant FEs	✓	✓	✓	✓	✓	✓
Observations	3,758,563	3,758,563	4,022,221	4,022,221	4,059,590	4,059,590
Adjusted R^2	0.40	0.23	0.44	0.22	0.43	0.14

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Full sample, main regression, individual, day-of-sample, and merchant fixed effects, cluster at individual, month, and merchant levels, split by time of the sample, spending outcomes.

	(1) Spending winsorized	(2) Spending deviation from mean	(3) Spending winsorized	(4) Spending deviation from mean	(5) Spending winsorized	(6) Spending deviation from mean
	1st time tercile		2nd time tercile		3rd time tercile	
Post cookie policy =1	2.404	0.020	2.236^{**}	0.018^{***}	8.401^{***}	0.037^{**}
	(1.370)	(0.011)	(0.777)	(0.005)	(2.044)	(0.014)
Online =1	35.168^{***}	0.163^{***}	36.371^{***}	0.193^{***}	41.006^{***}	0.220^{***}
	(3.456)	(0.014)	(2.874)	(0.015)	(3.503)	(0.022)
Post cookie policy=1 × Online=1	-6.855^{**}	-0.033	-22.456^{***}	-0.096^{***}	-26.756^{***}	-0.125^{***}
	(2.946)	(0.021)	(3.988)	(0.024)	(4.012)	(0.022)
Median of dependent variable	66.05		67.17		73.62	
Individual FEs	✓	✓	✓	✓	✓	✓
Day-of-sample FEs	✓	✓	✓	✓	✓	✓
Merchant FEs	✓	✓	✓	✓	✓	✓
Observations	1,095,864	1,095,864	1,096,747	1,096,747	1,056,951	1,056,951
Adjusted R^2	0.45	0.23	0.43	0.09	0.45	0.17

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Full sample, main regression, individual, day-of-sample, and merchant fixed effects, cluster at individual, month, and merchant levels, split by time of the sample post 2019, spending outcomes.

4.4 Placebo

As our placebo exercise, we reshuffle the cookie permission dates, and then rerun our main dynamic regression analyses.

	(1)	(2)	(3)	(4)	(5)	(6)
	Spending winsorized	Spending log	Spending winsorized	Spending log	Spending winsorized	Spending log
	1st spent tercile		2nd spent tercile		3rd spent tercile	
Post cookie policy =1	0.237***	0.026***	0.636***	0.027***	-1.297	0.002
	(0.042)	(0.007)	(0.076)	(0.003)	(2.254)	(0.008)
Online =1	-0.381***	-0.077***	1.279***	0.050***	56.009***	0.180***
	(0.048)	(0.008)	(0.072)	(0.003)	(2.875)	(0.007)
Post cookie policy=1 × Online=1	-0.002	0.013	-0.415***	-0.011*	-15.886**	-0.061***
	(0.068)	(0.012)	(0.149)	(0.006)	(6.287)	(0.018)
Median of dependent variable	6.94		26.80		189.16	
Individual FEs	✓	✓	✓	✓	✓	✓
Day-of-sample FEs	✓	✓	✓	✓	✓	✓
Merchant FEs	✓	✓	✓	✓	✓	✓
Observations	4,393,230	4,393,230	4,388,849	4,388,849	3,057,761	3,057,761
Adjusted R^2	0.23	0.28	0.14	0.15	0.41	0.40

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Full sample, main regression, individual, day-of-sample, and merchant fixed effects, cluster at individual, month, and merchant levels, split by amount, spending outcomes.

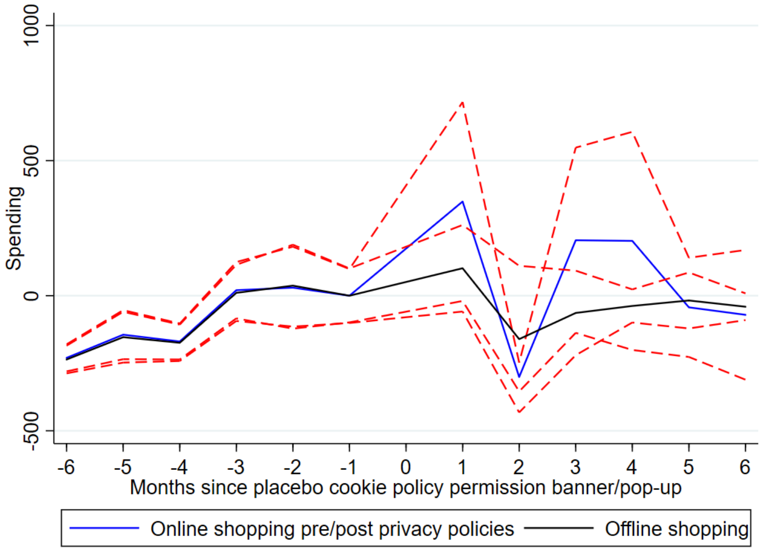


Figure 13: Placebo with reshuffled cookie permission dates, full sample, aggregate to online/offline pre/post-privacy policies levels, winsorized transaction amounts, individual and day-of-sample fixed effects, cluster at individual and month levels.

5 Discussion and Conclusion

Our findings raise important considerations regarding the trade-off between personalized marketing and consumer privacy. While targeted advertising can lead to increased profits for online retailers, it also raises concerns about the economic costs associated with lax data privacy standards.

Our findings indicate that the presence of cookie permission banners leads to a decrease in online spending at the respective retailer. To address concerns regarding potential selection bias and ensure the robustness of our results, we conduct a reduced form instrumental variable (IV) analysis. The IV instrument we employ is the exposure of each individual to cookie compliance policies based on their pre-policy shopping patterns at a specific retailer. Through this analysis, we validate our initial findings and additionally observe a positive impact on financial well-being, as indicated by reduced overdraft fees, higher checking account balances, and decreased credit card interest payments, associated with exposure to enhanced cookie compli-

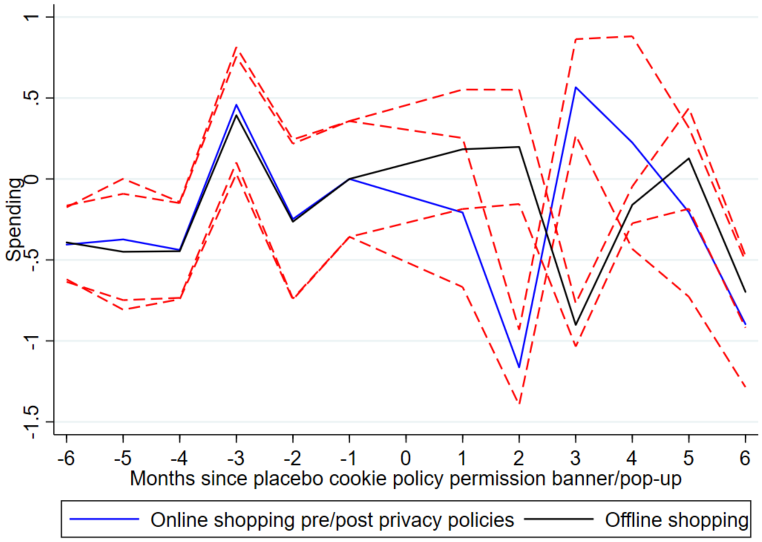


Figure 14: Placebo with reshuffled cookie permission dates, full sample, aggregate to online/offline pre/post-privacy policies levels, deviations relative to individual-level means, individual and day-of-sample fixed effects, cluster at individual and month levels.

ance measures.

Our research thus highlights the significance of data privacy regulations in safeguarding consumer interests. Existing regulations, such as the GDPR and state-level privacy acts, play a vital role in protecting user data. Policymakers should also consider the potential benefits of federal-level regulations aimed at enhancing data security to possibly protect customers from overspending.

References

Aridor, G., Y.-K. Che, and T. Salz (2020). The economic consequences of data privacy regulation: Empirical evidence from GDPR. NBER Working Paper 26900, Columbia University, Massachusetts Institute of Technology.

Aridor, G., Y.-K. Che, and T. Salz (2021). The effect of privacy regulation on the data industry: Empirical evidence from gdpr. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pp. 93–94.

Babina, T., G. Buchak, and W. Gornall (2022). Customer data access and fintech entry: Early evidence from open banking. *Available at SSRN*.

Berman, R. and A. Israeli (2021). The added value of data-analytics: Evidence from online retailers. Technical report, working paper.

Bessen, J. E., S. M. Impink, L. Reichensperger, and R. Seamans (2020). GDPR and the importance of data to AI startups. Working paper, New York University, Boston University.

Bian, B., X. Ma, and H. Tang (2021). The supply and demand for data privacy: Evidence from mobile apps. *Available at SSRN*.

Bian, B., M. Pagel, and H. Tang (2023). Consumer surveillance and financial fraud. Technical report, mimeo.

Bindra, C. (2019). Next steps to ensure transparency, choice and control in digital advertising.

Dubé, J.-P. and S. Misra (2023). Personalized pricing and consumer welfare. *Journal of Political Economy 131*(1), 131–189.

Goldberg, S., G. Johnson, and S. Shriver (2019). Regulating privacy online: The early impact of the gdpr on european web traffic & e-commerce outcomes. *Available at SSRN 3421731*.

Janssen, R., R. Kesler, M. Kummer, and J. Waldfogel (2021). GDPR and the lost generation of innovative apps. NBER Working Paper 146409, University of Zurich, University of Minnesota, University of East Anglia, Georgia Institute of Technology.

Jia, J., G. Z. Jin, and L. Wagman (2021). The short-run effects of the general data protection regulation on technology venture investment. *Marketing Science*, forthcoming.

Johnson, G. (2013). The impact of privacy policy on the auction market for online display advertising.

Kesler, R. (2022). The impact of Apple's app tracking transparency on app monetization. *Available at SSRN 4090786*.

Lukic, K., K. M. Miller, and B. Skiera (2023). The impact of the general data protection regulation (gdpr) on online tracking. *Available at SSRN*.

Peukert, C., S. Bechtold, M. Batikas, and K. Tobias (2021). Regulatory spillovers and data governance: Evidence from the GDPR. *Marketing Science*, forthcoming.

Todri, V. (2022). Frontiers: The impact of ad-blockers on online consumer behavior. *Marketing Science 41*(1), 7–18.

Wernerfelt, N., A. Tuchman, B. Shapiro, and R. Moakler (2022). Estimating the value of offsite data to advertisers on meta. *University of Chicago, Becker Friedman Institute for Economics Working Paper* (114).

Zhao, Y., P. Yildirim, and P. K. Chintagunta (2021). Privacy regulations and online search friction: Evidence from gdpr. *Available at SSRN 3903599*.