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 and financial well-being. 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 transaction data from a US data aggregation and analytics provider. We first show that online spending at retailers decreases 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, late, and cash-advance 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 offer insights into the impact of data privacy regulations on consumers, informing policy considerations at both 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

[†]Washington University's Olin Business School, NBER, and CEPR. E-mail: mpagel@wustl.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, Stockholm School of Economics, Federal Reserve Bank of Boston, HEC Paris, University of Heidelberg, DICE, Copenhagen Business School, University of Melbourne, Monash University, and the University of Missouri. We also thank Anton Tarasenko for providing outstanding research assistance as well as Ansgar Walther, Antoine Uettwiller, and Tarun Ramadorai for sharing their data on cookies and web trackers.

1 Introduction

Over the past few decades, privacy has undergone significant changes driven by technological advancements and shifting societal norms. The ubiquity of digital devices and online services has led to unprecedented data collection. At the same time, data storage capacities and analyzing capabilities have evolved at a similar pace. Increased individual identification and tracking have eroded anonymity, contributing 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 may benefit businesses but harm customers by fostering manipulation, exploiting vulnerabilities, and promoting overspending. Additionally, third-degree price discrimination may lead 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, introduced in May 2018 (EU GDPR), required websites to 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. 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 examine how the introduction of cookie compliance systems by retailers, via cookie permission banners or pop-up windows, affects 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 collect data from 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 comply with cookie privacy regulations, retailers typically implement a cookie consent mechanism on their websites through third-party providers, which ensure that they adhere to 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 automate cookie scanning, record user consent, and offer preference centers that enable users to manage their privacy settings. They also offer customizable banners, consent management platforms, and granular consent options while ensuring compliance.

The introduction of cookie compliance systems is a shock that enhances customer data privacy, as 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 accept all cookies.[1]

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

[1]We provide suggestive evidence for that using actual cookie placement data obtained from Ramadorai et al. (2020). We matched their data with our data on firms' cookie compliance infrastructure dates by firms' URLs to show significant and negative effects for all cookie and tracking measures post cookie compliance frameworks.

data analytics to clean and categorize transaction data, which is then offered to institutional investors and investment managers in both aggregated and disaggregated forms. We utilize the de-identified transactions dataset, which includes all bank and credit card transactions, encompassing spending, income, and geographical information, 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. 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 occurs through digital or electronic means and is 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 the effects of cookie compliance systems even when they are modest.

We find that cookie permission banners lead to a reduction in online spending at the retailer in question. To address the concern that certain types of individuals shop with enhanced data privacy at specific times and that this selection bias 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 at any given retailer. We confirm our initial results and also estimate a positive effect on financial well-being, as measured by overdraft, late, and cash-advance 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 and financial health. Our analysis sheds light on 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 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. Additionally, price discrimination and shopping convenience are likely to be facilitated by the collection and analysis of user data.

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 and is generally not illegal. 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 violate 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 emergencies, 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.

The existing literature suggests that targeted advertisements increase company profits. 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, they 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 friction in online searches. Berman and Israeli (2021) show that the adoption of data analytics impacts 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.

Close to our paper, Bae et al. (2023) use a difference-in-differences approach to contrast changes in the shopping behavior between Californians and non-Californians after the CCPA. They find that post-policy enactment, Californians reduce purchases by 4.3% and increase returns by 3.0%, which results in a $96 drop in discretionary spending. Moreover, using browsing data, they find that Californians spend more time online and view more pages per website, potentially indicating more search efforts. Additionally, Doerr et al. (2023) shows that the CCPA increased loan applicants' willingness to share data and improved the screening process of FinTech lenders, which allows them to offer lower loan rates, in particular, to traditionally under-served groups.

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.

Our paper also adds to recent work that examines 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. And Bian et al. (2023) raise concerns about consumer privacy and the potential for financial fraud due to data breaches.

3 Empirical Analysis

3.1 Privacy Policy Changes

Companies implement privacy policies and cookie permission systems 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.

For example, McDonald's partnered with OneTrust. Since August 9, 2019, the

company has had a privacy policy in place that outlines how it collects, uses, and protects consumer data. When a user visits the McDonald's website, they are asked for consent to place 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.

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

The premise of this paper is that when a retailer introduces a framework that makes its website compliant with privacy laws, this is a step that indeed enhances the privacy of its customers. There are two ways in which privacy would be enhanced. First, some customers, anecdotally around 20% of people, reject cookies. The cookie compliance providers promise that no unnecessary cookies will be used for these customers and that their privacy will be enhanced. Second, even if customers say yes to cookies, if retailers take steps to protect the privacy of their customers, we would expect their cookie placements to be less aggressive.

We provide suggestive evidence for that using actual cookie placement data obtained from Ramadorai et al. (2020), who examine the firms' data extraction practices and how it is influenced by privacy regulation. The authors obtained data on each firm's data extraction using the "OpenWPM" crawler developed in Englehardt

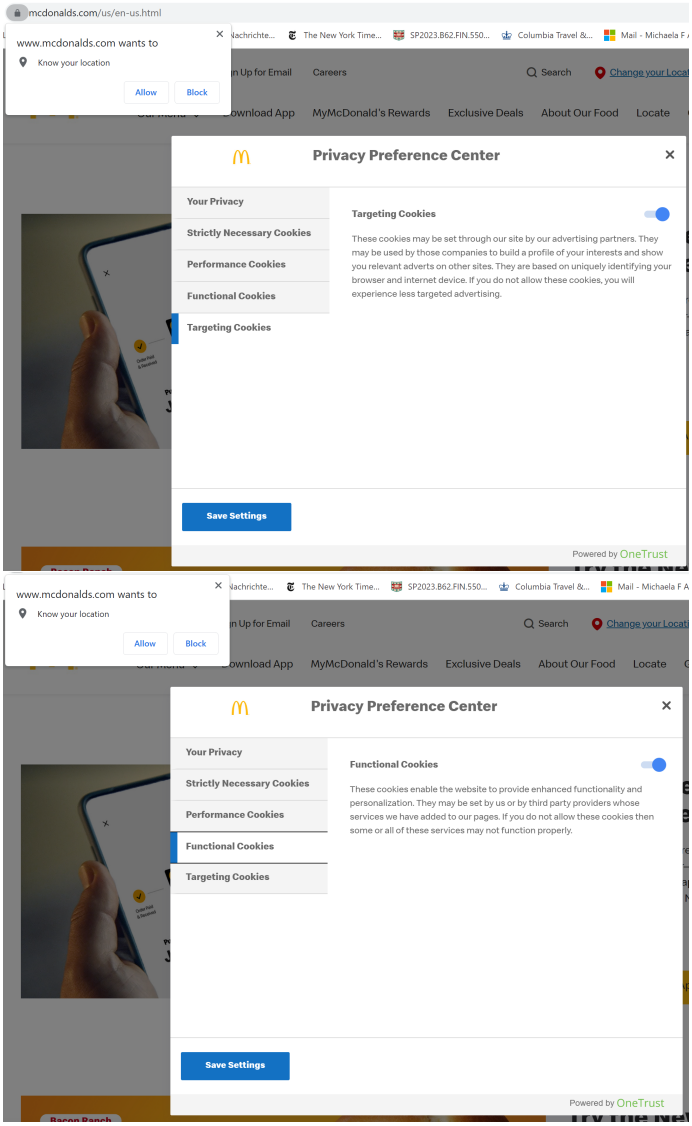


Figure 1: McDonald's privacy center.

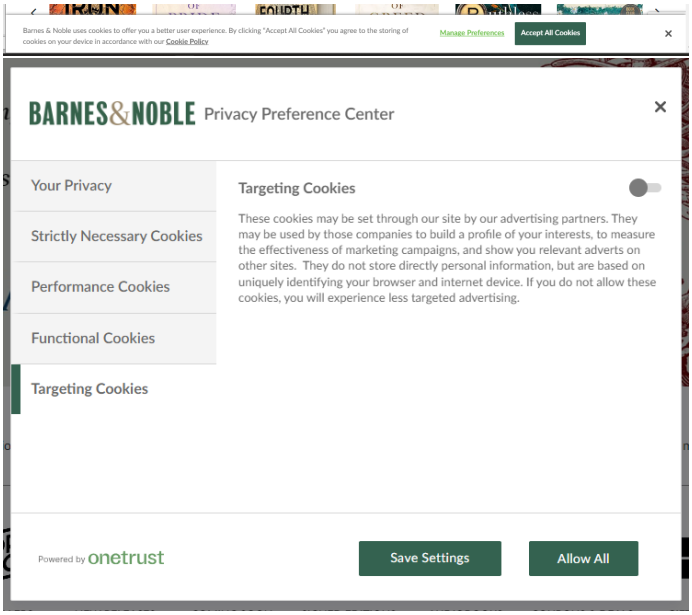


Figure 2: Barnes & Noble's privacy center.

and Narayanan (2016). The authors shared with us the data from their two runs of the crawler in February 2019 and in May 2023. We then matched their data with our data on firms' cookie compliance infrastructure dates by firms' URLs. We ran a regression of each cookie and tracking measure on firm fixed effects, crawling date fixed effects, and an indicator for whether the firm had a cookie compliance framework in place (see Table 1). We find significant and negative results for all cookie and tracking measures which reassures us that the cookie compliance frameworks actually enhanced customer privacy. Our cookie and tracking measures include first-party (FP) cookies, which are mostly used for website functionality and show less of a reduction in response to the cookie compliance framework, and third-party (TP) cookies and tracker requests, which are more affected.

	(1) FP cookies	(2) TP tracking cookies	(3) Tracker requests	(4) TP requests	(5) Cookies	(6) Total cookies
Post cookie policy	-1.283	-4.232***	-8.574***	-24.169***	-12.924***	-5.515***
	(0.954)	(0.717)	(1.741)	(6.230)	(2.559)	(1.496)
Mean of dependent variable	7.76	2.94	9.48	36.69	15.83	10.70
URL/firm FEs	✓	✓	✓	✓	✓	✓
Crawling date FEs	✓	✓	✓	✓	✓	✓
Observations	7,734	7,734	7,734	7,734	7,734	7,734
Adjusted R^2	0.63	0.54	0.56	0.51	0.52	0.59

Standard errors in parentheses

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

Table 1: Matched sample of cookie compliance dates from Builtwith with the crawling cookie and tracking data from Ramadorai et al. (2020). Each cookie and tracking outcome is regressed on URL/company fixed effects, crawling date fixed effects, and an indicator for whether a cookie compliance framework was in place at the time of the data crawl.

Figures 3 and 4 further illustrate the changes in all cookie and tracking measures before versus after firms introduced cookie compliance infrastructures. Figure 3 shows histograms before versus after for each tracking measure, and Figure 4 shows the differences for each cookie measure for each firm that introduced a cookie compliance framework.



Figure 3: Matched sample of cookie compliance dates from Builtwith with the crawling cookie and tracking data from Ramadorai et al. (2020). Histograms showing the frequency of each cookie and tracking measure before versus after firms' introduced cookie compliance frameworks.

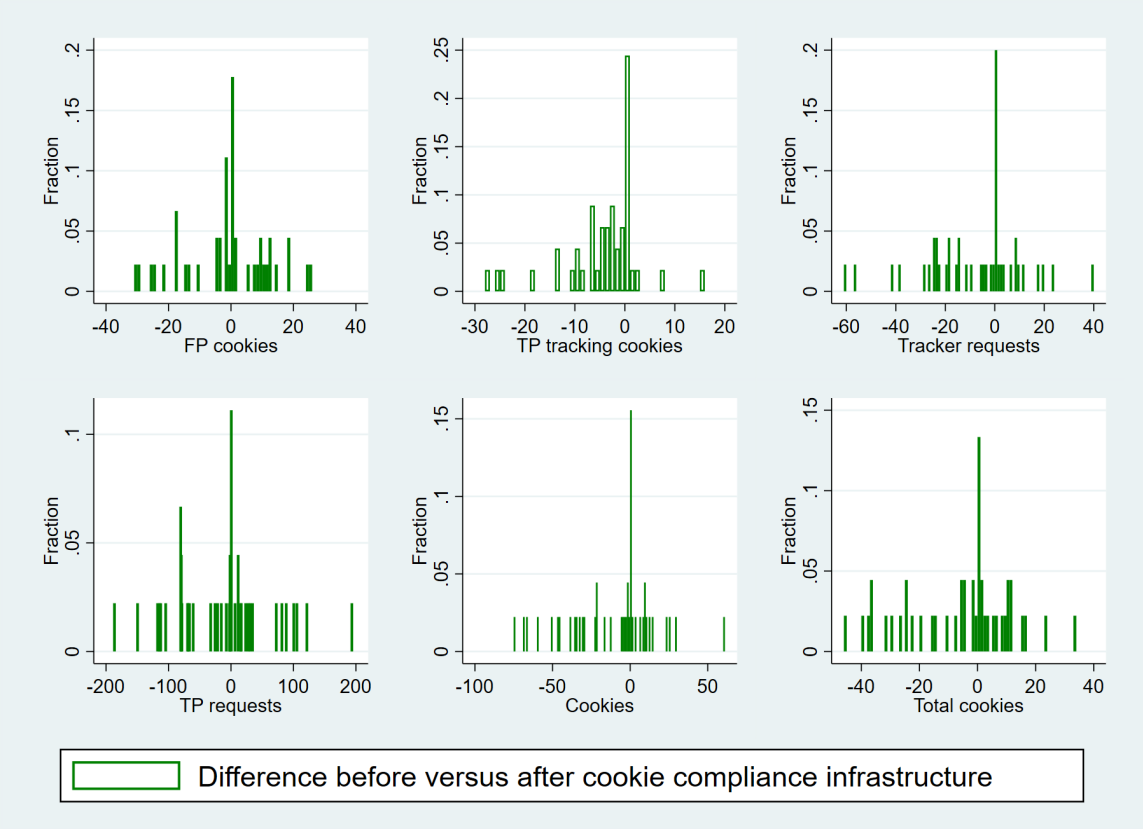


Figure 4: Matched sample of cookie compliance dates from Builtwith with the crawling cookie and tracking data from Ramadorai et al. (2020). Histogram of the differences of each cookie and tracking measure before versus after firms' introduced cookie compliance frameworks.

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. For each individual, the dataset contains all income and spending transactions from a number of their bank and credit card accounts. The transactions also include 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 cookie compliance policy implementation as staggered treatments and online versus offline shopping as treatment intensity. Our main regression specification is:

$$SpendPerTransaction_{i,m,t} = \alpha_i + \alpha_m + \alpha_t + \beta PostPolicy_{m,t} \times Online_{i,m,t} + \varepsilon_{i,m,t} \quad (1)$$

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. We then control for individual, merchant, and day-of-sample fixed effects. We cluster standard errors at individual, merchant, and monthly levels.

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 an exposure-to-enhanced-privacy-policies instrument. We then use the instrument in a reduced-form IV regression.

$$SpendPerTransaction_{i,m,t} = \alpha_i + \alpha_t + \alpha_m + \beta EnhPrivacy_{i,t} \times Online_{i,m,t} + \varepsilon_{i,m,t} \quad (2)$$

where i, m, and t denote individual, merchant, and day-of-sample as above. Again, we control for individual, merchant, 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 at retailers pre cookie-compliance policy that are either online or offline as well as all transactions at retailers post cookie-compliance that are either online or offline. We then fill in zeros to obtain a panel and run the following specification:

$$Spend_{i,t} = \alpha_i + \alpha_t + \beta PostPolicy_{i,t} \times Online_{i,t} + \varepsilon_{i,t} \quad (3)$$

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

We conduct all analyses on a random subsample of the overall dataset, keeping 5,000 individuals. Table 2 provides summary statistics of the transaction-level data.

Figure 5 shows the share of transactions done at retailers that have privacy policies and cookie compliance systems in place over the sample period.

Figure 6 shows the average value of our enhanced privacy instrument over the sample period.

	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	3.870	41	832.0	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
Cash advance fees	26.91	149.8	6.440	0.0900	1.500	21.17	100
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 2: Full sample, individual-transaction level, January 2014 to August 2023.

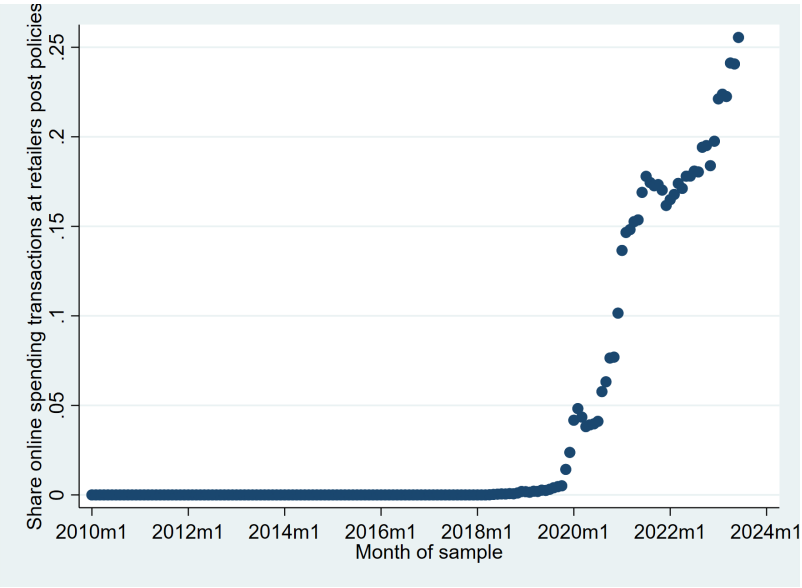


Figure 5: Share of online retail transactions at retailers post their privacy policy.

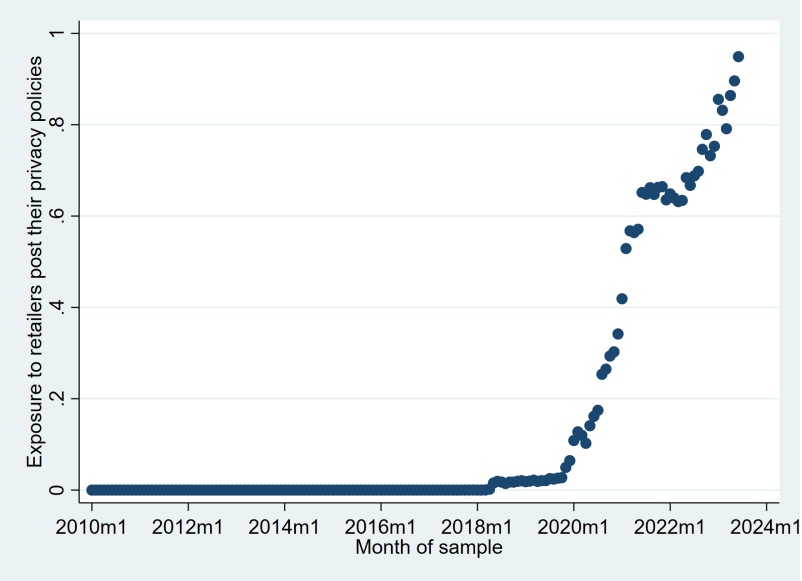


Figure 6: Exposure to online retailers' privacy policies (basket share times indicator of privacy policy implementation).

4 Results

4.1 Main Regression Results

Our analysis reveals that the introduction of cookie compliance systems significantly decreases consumer spending. This is consistent with our proposed channels: restricting targeted advertising, consumer profiling, price discrimination, and shopping convenience should negatively impact spending.

As discussed, Table 3 uses the raw transactions sample running Specifications 1 and 2. We find decreases in spending in both the OLS and IV analyses.

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

	(1) Spending winsorized	(2) Spending winsorized	(3) Spending log	(4) Spending winsorized	(5) Spending log
Post cookie policy =1	-10.957***	0.126	0.078***		
	(1.089)	(0.629)	(0.008)		
Online =1	61.896***	47.594***	0.355***	47.707***	0.360***
	(1.523)	(2.005)	(0.013)	(2.007)	(0.012)
Post cookie policy =1 × Online=1	-21.645***	-13.932***	-0.104***		
	(3.985)	(2.178)	(0.017)		
Enhanced privacy				-0.487	0.025***
				(0.380)	(0.004)
Online =1 × Enhanced privacy				-4.443***	-0.054***
				(0.792)	(0.007)
Individual FEs	✓	✓	✓	✓	✓
Day-of-sample FEs	✓	✓	✓	✓	✓
Merchant FEs		✓	✓	✓	✓
Observations	11,841,348	11,841,261	11,841,261	11,841,261	11,841,261
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, Regressions 1 and 2, individual, day-of-sample, and merchant fixed effects, cluster at individual, month, and merchant levels.

		(1) Spending winsorized	(2) Spending winsorized	(3) Spending winsorized	(4) Spending winsorized	(5) Spending winsorized	(6) Spending winsorized
		Electronics general merchandise	Personal and family	Home improvement	Travel	Restaurants	Groceries
Post cookie policy	=1	3.784**	-3.119**	8.304	7.473***	-1.466***	4.455**
		(1.725)	(1.378)	(11.185)	(1.434)	(0.332)	(1.719)
Online	=1	12.033***	6.510***	65.677***	14.943***	13.601***	25.359***
		(2.043)	(1.792)	(7.219)	(4.351)	(2.068)	(7.428)
Post cookie policy	=1 × Online=1	-29.961***	-2.427	-3.742	-5.930***	-4.357**	6.371
		(4.885)	(3.388)	(11.587)	(1.533)	(2.096)	(5.751)
Individual FEs		✓	✓	✓	✓	✓	✓
Day-of-sample FEs		✓	✓	✓	✓	✓	✓
Merchant FEs		✓	✓	✓	✓	✓	✓
Observations		1,967,384	556,298	311,424	766,906	2,588,458	1,953,711
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, Regression 1, individual and day-of-sample fixed effects, cluster at individual, month, and merchant levels, broken down by spending in categories.

		(1) Spending winsorized	(2) Spending winsorized	(3) Spending winsorized	(4) Spending winsorized	(5) Spending winsorized	(6) Spending winsorized
		Electronics general merchandise	Personal and family	Home improvement	Travel	Restaurants	Groceries
Enhanced privacy		3.020***	0.002	-1.475	0.755	-0.429**	0.065
		(0.569)	(0.833)	(2.543)	(0.694)	(0.203)	(0.299)
Online	=1	11.666***	7.247***	65.301***	14.580***	13.930***	25.565***
		(1.919)	(1.814)	(7.280)	(4.342)	(2.177)	(7.819)
Online=1 × Enhanced privacy		-8.303***	-5.180***	1.525	-1.346	-2.333**	-0.770
		(1.031)	(1.103)	(2.748)	(0.854)	(1.125)	(2.453)
Individual FEs		✓	✓	✓	✓	✓	✓
Day-of-sample FEs		✓	✓	✓	✓	✓	✓
Merchant FEs		✓	✓	✓	✓	✓	✓
Observations		1,967,384	556,298	311,424	766,906	2,588,458	1,953,711
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, Regression 2, 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 the enhanced privacy instrument in Table 6. We find that the likelihood of overdraft fees, the payments made to credit cards, the rolled-over credit card debt, as measured by interest charges, as well as the prevalence of other high-interest unsecured loans is reduced. This is consistent with lax privacy standards, which lead to overspending because of targeted advertising or price discrimination.

	(1) Overdraft fees	(2) Late fees	(3) NSF fees	(4) Cash advance fees	(5) Interest charges	(6) Payday/EWA /FinTech loans	(7) Checking account balance log
Enhanced privacy	-0.051***	-0.003***	-0.004	-0.005**	-0.060***	-0.331***	0.020
	(0.014)	(0.000)	(0.008)	(0.002)	(0.008)	(0.083)	(0.016)
Mean of dependent variable	0.15	0.01	0.05	0.02	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	20,834,520	4,612,816
Adjusted R^2	0.00	0.00	0.01	0.00	0.01	0.01	0.83

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Full sample, Regression 2 with financial health outcomes, individual and day-of-sample fixed effects, cluster at individual and month levels.

In turn, we aggregate the data to the offline/online pre/post cookie compliance system, date, and individual levels. 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 7 and 8 show the same results as in the previous regressions in this aggregated data set.

	(1) Spending winsorized	(2) Number of transactions	(3) Spending deviation from mean	(4) Spending winsorized	(5) Number of transactions	(6) Spending 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, Regression 3, aggregate to online/offline pre/post-privacy policies levels, individual and day-of-sample (or individual times month-of-sample as well as day-of-week and week-of-month) fixed effects, cluster at individual and month levels, include zeros, spending outcomes.

	(1) Overdraft fees	(2) Late fees	(3) NSF fees	(4) Cash advance fees	(5) Interest charges	(6) Payday/EWA/ FinTech loans	(7) Checking account balance log
Enhanced privacy	-0.187***	-0.009***	0.053	-0.005	-0.149***	-0.415	0.020
	(0.063)	(0.003)	(0.044)	(0.014)	(0.037)	(0.401)	(0.016)
Mean of dependent variable	0.70	0.04	0.24	0.11	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	4,613,414
Adjusted R^2	0.01	0.01	0.04	0.00	0.04	0.01	0.83

Standard errors in parentheses

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

Table 8: Full sample, Regression 3, aggregate to online/offline pre/post-privacy policies levels, 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 retailer-specific cookie permission introduction dates. We first use the raw transaction data, winsorized, logged, or divided-by-individual-means spending transactions, and include individual, day-of-sample, as well as their interaction fixed effects. We also aggregate to the individual-day online/offline pre/post policy levels and include zeros.

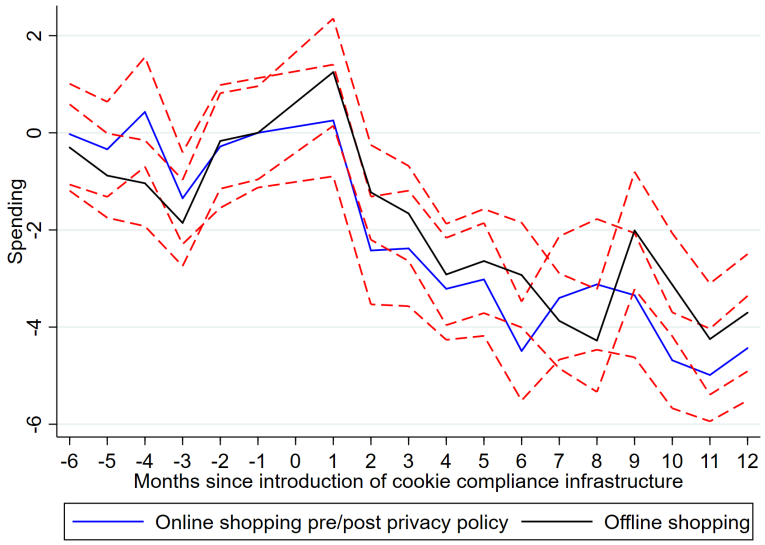


Figure 7: Full sample, Regression 3, aggregated to online/offline pre/post policies levels, winsorized transactions,individual and day-of-sample fixed effects, cluster at individual and month levels.

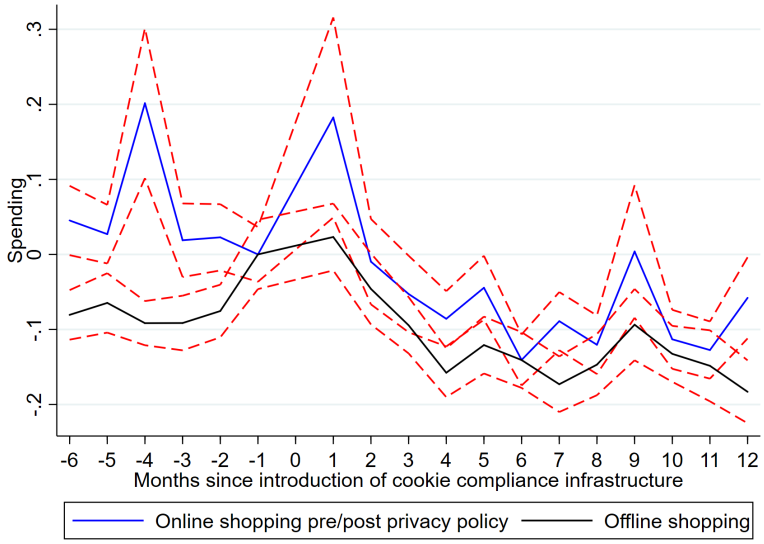


Figure 8: Full sample, Regression 3, aggregated to online/offline pre/post policies levels, deviations relative to individual-level means, individual and day-of-sample fixed effects, cluster at individual and month levels.

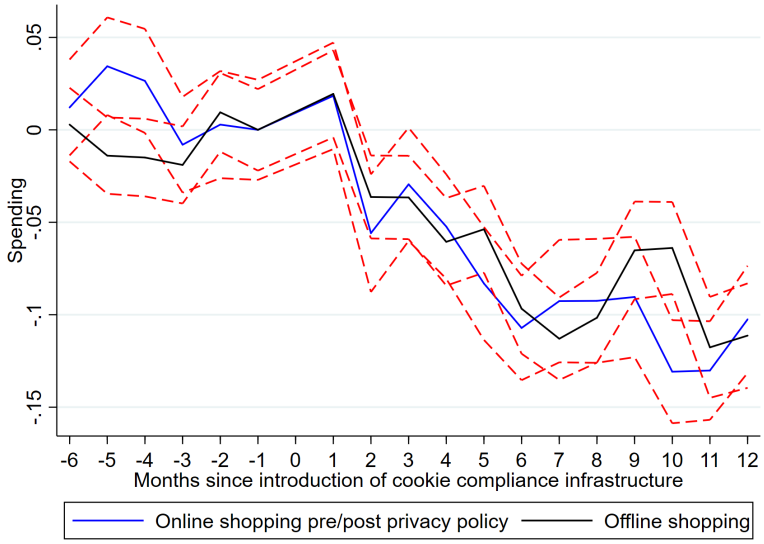


Figure 9: Full sample, Regression 3, aggregated to online/offline pre/post policies levels, log spending, individual and day-of-sample fixed effects, cluster at individual and month levels.

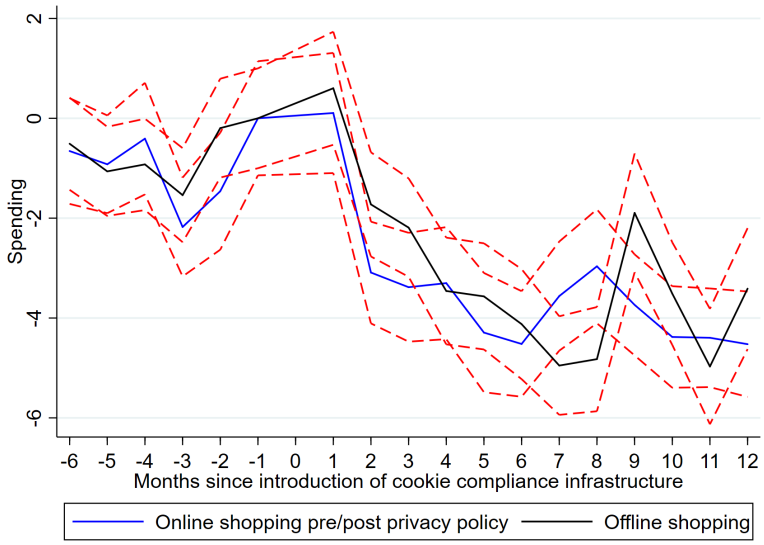


Figure 10: Full sample, Regression 3, aggregated to online/offline pre/post policies levels, winsorized transactions,individual times month-of-sample, day-of-week, and week-in-month fixed effects, cluster at individual and month levels.

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. Figure 16 shows this correlation for features that would likely affect online shopping, and Figure 17 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 were correlated, we would expect a lot of points on the 45-degree line, which is not what we see.

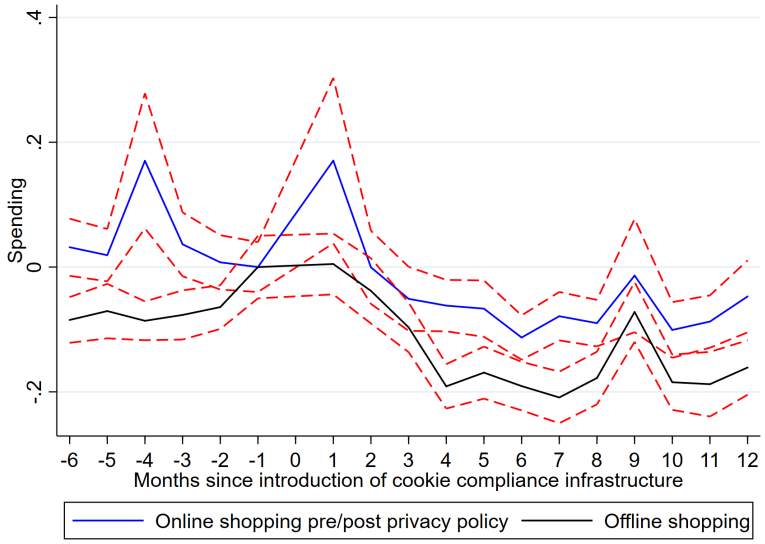


Figure 11: Full sample, Regression 3, aggregated to online/offline pre/post policies levels, deviations relative to individual-level means, individual times month-of-sample, day-of-week, and week-in-month fixed effects, cluster at individual and month levels.

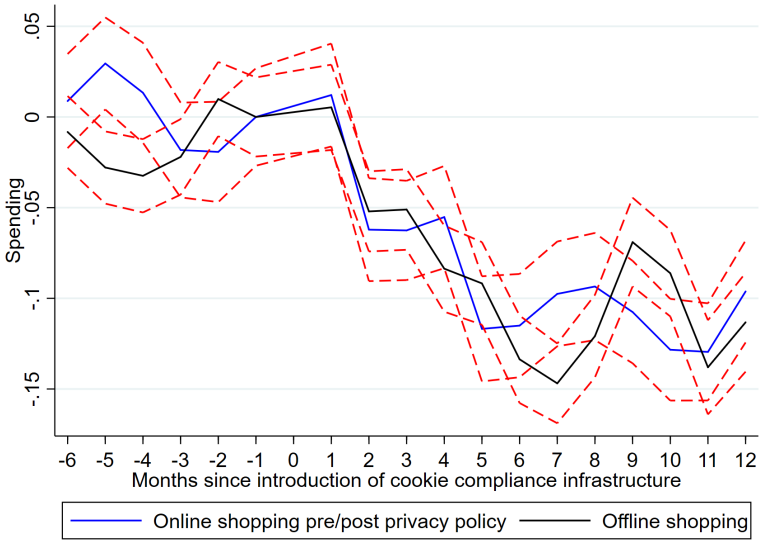


Figure 12: Full sample, Regression 3, aggregated to online/offline pre/post policies levels, log spending, individual times month-of-sample, day-of-week, and week-in-month fixed effects, cluster at individual and month levels.

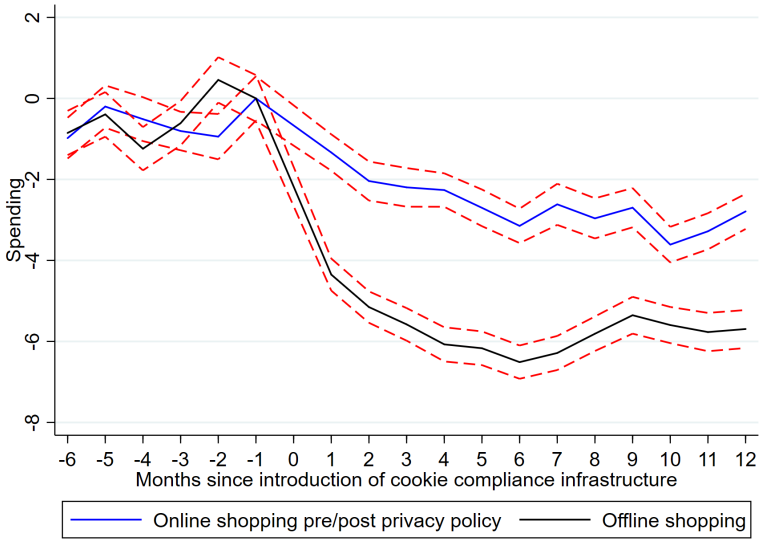


Figure 13: Full sample, Regression 3, aggregated to online/offline pre/post policies levels including zeros, winsorized transactions,individual times month-of-sample, day-of-week, and week-in-month fixed effects, cluster at individual and month levels.

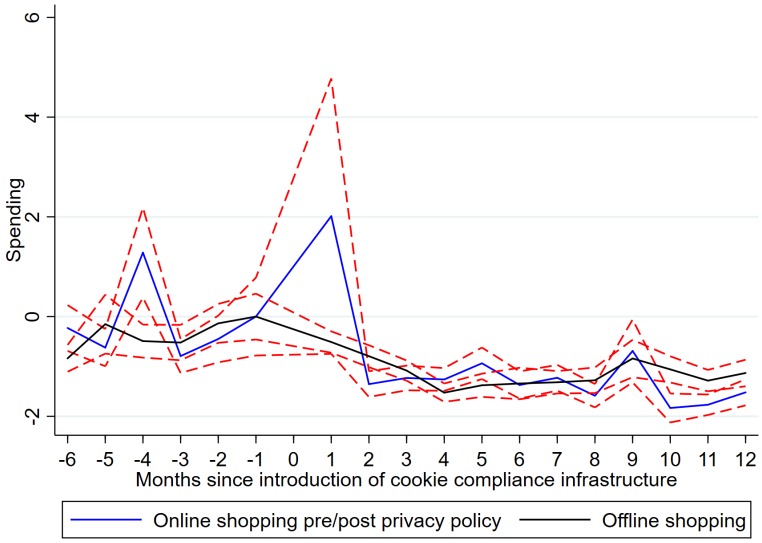


Figure 14: Full sample, Regression 3, aggregated to online/offline pre/post policies levels including zeros, deviations relative to individual-level means, individual times month-of-sample, day-of-week, and week-in-month fixed effects, cluster at individual and month levels.

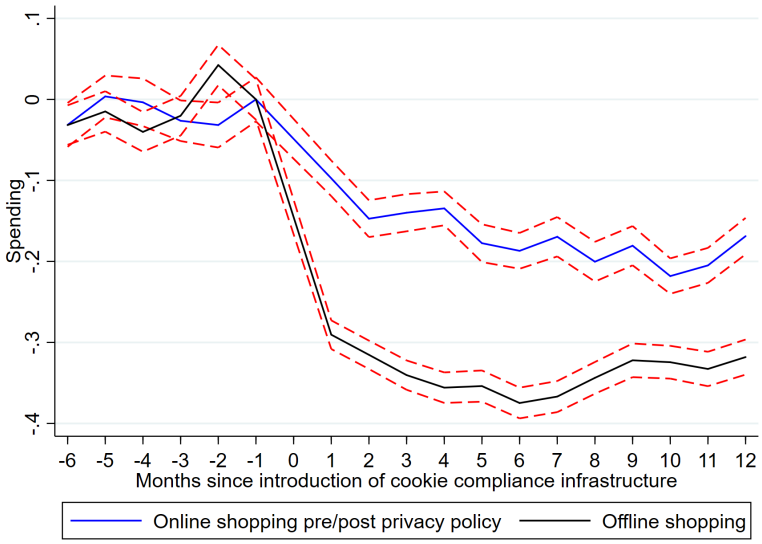


Figure 15: Full sample, Regression 3, aggregated to online/offline pre/post policies levels including zeros, log spending, individual times month-of-sample, day-of-week, and week-in-month fixed effects, cluster at individual and month levels.

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

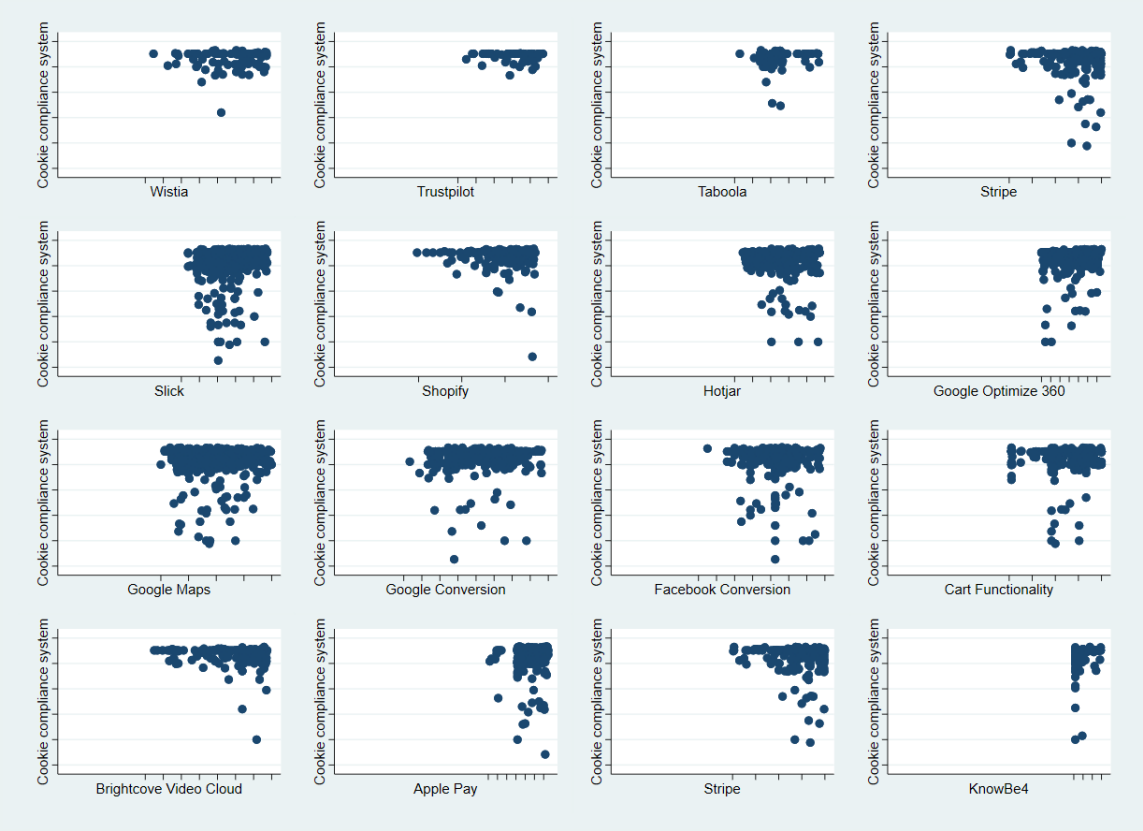


Figure 16: Dates of the additions of a cookie compliance system with the addition of other website features that may also affect online shopping. The date range on the x- and y-axes are the same, from May 2001 to June 2023.

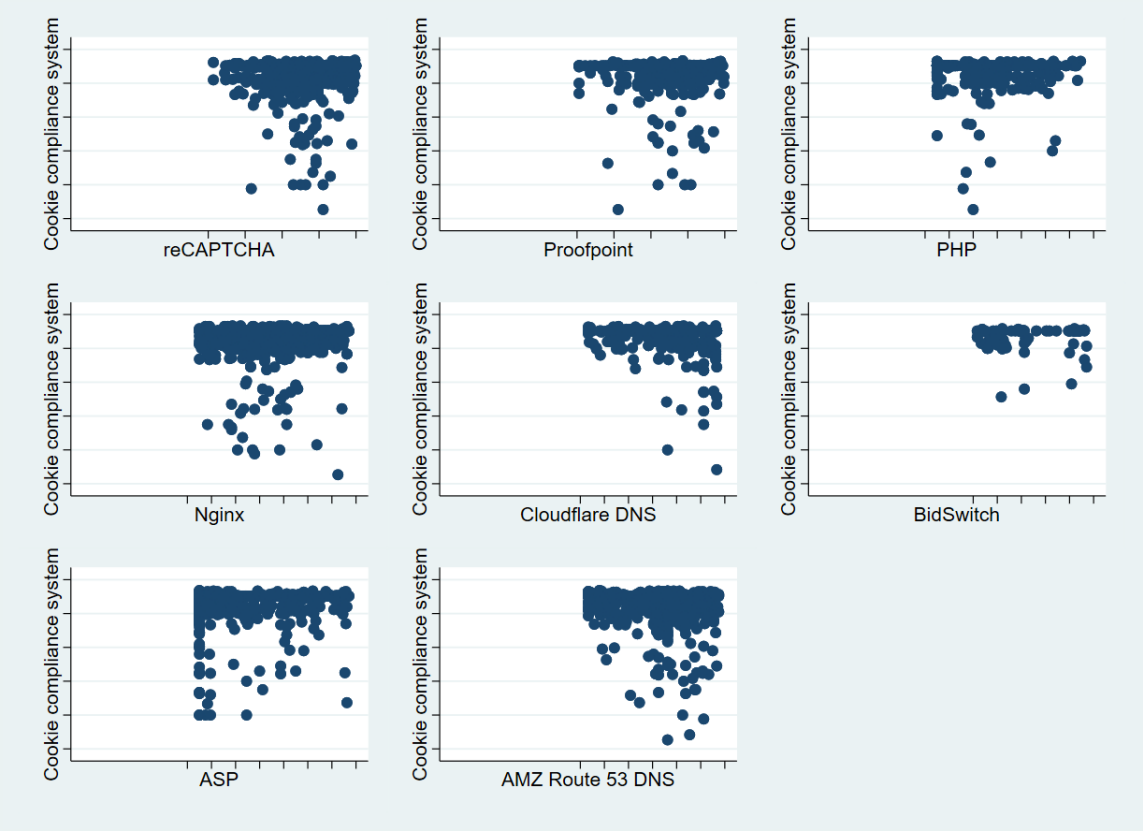


Figure 17: Dates of the additions of a cookie compliance system with the addition of other website features that are unlikely to affect online shopping. The date range on the x- and y-axes are the same, from May 2001 to June 2023.

	(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 income tercile		2nd income tercile		3rd income tercile	
Post cookie policy =1	-1.312*	-0.000	0.963	0.007	-0.620	-0.018***
	(0.743)	(0.006)	(0.945)	(0.006)	(1.379)	(0.006)
Online =1	35.863***	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	-8.203***	-0.051**	-14.115***	-0.058**	-22.199***	-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	3,692,359	3,692,359	3,822,546	3,822,546	3,701,822	3,701,822
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, Regression 1, individual, day-of-sample, and merchant fixed effects, cluster at individual, month, and merchant levels, split by regular income, spending outcomes.

4.4 Placebo

As our placebo exercise, we reshuffle the cookie permission dates and then rerun our main dynamic regression analyses. Here, we use the specification that aggregates to the individual online/offline pre/post policies levels and use logged spending and deviations relative to the means as the outcome variables, which yielded the strongest effects in the original data.

		(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***	37.815***	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	0.693	0.003
				(1.664)	(0.014)	(0.712)	(0.004)
Post cookie policy	=1 × Online=1			-20.729	-0.323	-9.996***	-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	4,059,587	4,059,587
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, Regression 1, 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***	8.408***	0.037**
		(1.370)	(0.011)	(0.777)	(0.005)	(2.044)	(0.014)
Online	=1	35.168***	0.163***	36.371***	0.193***	41.009***	0.220***
		(3.456)	(0.014)	(2.874)	(0.015)	(3.504)	(0.022)
Post cookie policy	=1 × Online=1	-6.855**	-0.033	-22.456***	-0.096***	-26.761***	-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	1,056,948	1,056,948
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, Regression 1, 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.

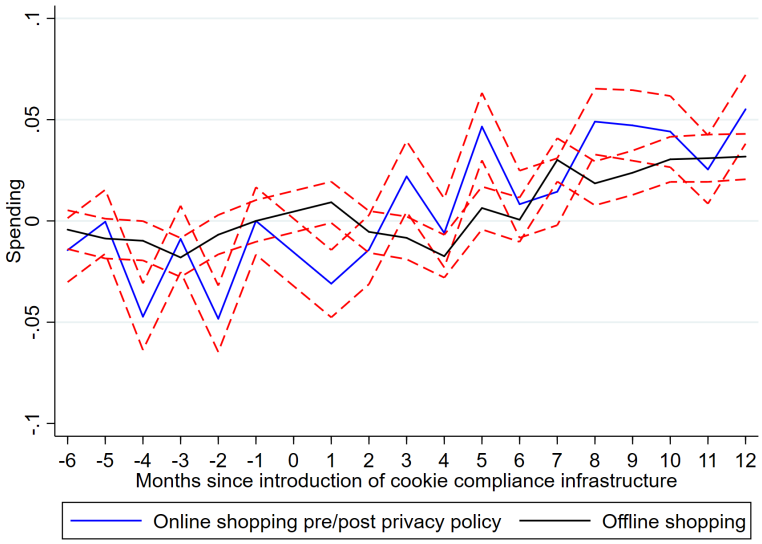


Figure 18: Full sample, Regression 3, aggregated to online/offline pre/post policies levels, logged spending, individual times month-of-sample, day-of-week, and week-in-month fixed effects, cluster at individual and month levels.

		(1) Spending winsorized	(2) Spending log	(3) Spending winsorized	(4) Spending log	(5) Spending winsorized	(6) Spending log
		1st spent tercile		2nd spent tercile		3rd spent tercile	
Post cookie policy	=1	0.237***	0.026***	0.636***	0.027***	-1.291	0.002
		(0.042)	(0.007)	(0.076)	(0.003)	(2.254)	(0.008)
Online	=1	-0.381***	-0.077***	1.279***	0.050***	56.010***	0.180***
		(0.048)	(0.008)	(0.072)	(0.003)	(2.875)	(0.007)
Post cookie policy	=1 × Online=1	-0.002	0.013	-0.416***	-0.011*	-15.893**	-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		4,393,235	4,393,235	4,388,846	4,388,846	3,057,756	3,057,756
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, Regression 1, individual, day-of-sample, and merchant fixed effects, cluster at individual, month, and merchant levels, split by amount, spending outcomes.

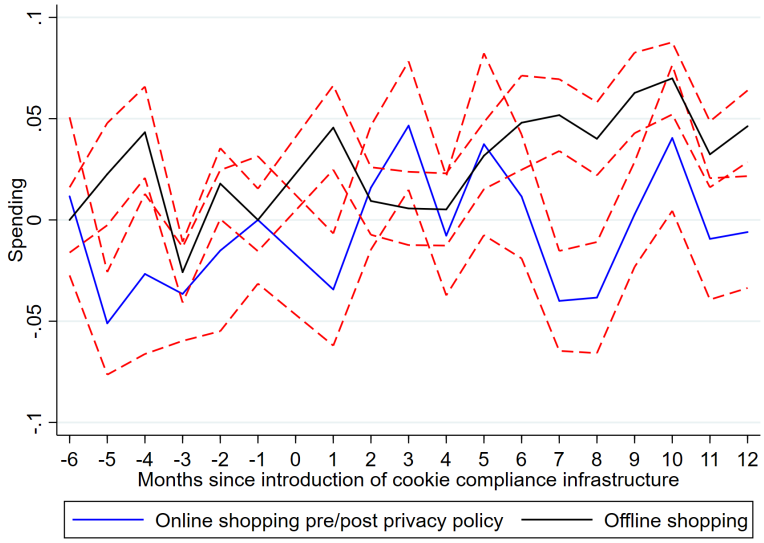


Figure 19: Full sample, Regression 3, aggregated to online/offline pre/post policies levels, deviations relative to individual-level means, individual times month-of-sample, day-of-week, and week-in-month fixed effects, cluster at individual and month levels.

5 Discussion and Conclusion

Our findings raise important considerations regarding the trade-off between personalized marketing and consumer privacy. While targeted advertising and price discrimination 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, decreased credit card interest payments, and decreased high-interest unsecured loans associated with enhanced cookie compliance 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 protect customers from overspending and worse financial health.

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*.

Bae, D., R. Mayya, and T. Nian (2023). Privacy regulation and its unintended consequence on consumption behaviors: Evidence from ccpa.

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.

Doerr, S., L. Gambacorta, L. Guiso, and M. Sanchez del Villar (2023). Privacy regulation and fintech lending. *Available at SSRN 4353798*.

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

Englehardt, S. and A. Narayanan (2016). Online tracking: A 1-million-site measurement and analysis. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, pp. 1388–1401.

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.

Ramadorai, T., A. Uettwiller, and A. Walther (2020). Privacy policies and consumer data extraction: Evidence from u.s. firms. *Available at SSRN 3352175*.

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*.