Privacy Regulation and Its Unintended Consequence on Consumption Behaviors: Evidence From CCPA

Donghwa Bae*[1], Raveesh Mayya[2], and Tingting Nian[1]

[1]Paul Merage School of Business,
University of California, Irvine

[2]Leonard N. Stern School of Business,
New York University

April 22, 2023

Abstract

This study explores the unintended consequences of data protection regulations on consumer purchasing behavior and satisfaction. Specifically, we examine the California Consumer Privacy Act, which restricts companies from collecting, buying, or selling the personal information of California residents. Regulations that increase the liability of firms for data collection and storage may impede firms from uncovering latent consumer preferences that underlie the data, potentially altering consumers' subsequent consumption behavior. These regulations especially pose a challenge to digital platforms since a platform strategy's success hinges on facilitating the smooth flow of data between participants, generating value for both producers and consumers. Drawing on a unique panel dataset compiled from billions of individual monetary transactions on a payment gateway, we employ a difference-in-differences approach to contrast changes in the shopping behavior between Californians and non-Californians. Our analyses reveal that, post policy enactment, Californians reduce purchases by 4.3%, increase returns by 3.0%, resulting in a $96 drop in discretionary spending. Moreover, we employ a proprietary browsing behavior dataset and find that Californians spend more time online and view more pages per website, potentially indicating more search efforts. Mechanism analysis suggests that firms covered by CCPA proactively alter their data collection strategy to reduce the liability under the law. These results reveal the complex interplay between privacy regulation and consumer behavior, highlighting the need for a nuanced understanding of the trade-offs between privacy protection and economic outcomes. The results have important implications for affected businesses as well as policymakers involved in designing and implementing future privacy regulations.

Keywords: Privacy Regulation, Consumer Behaviors, Unintended Consequences

*Corresponding Author
Donghwa Bae: donghwab@uci.edu, Raveesh Mayya: raveesh@stern.nyu.edu, Tingting Nian: tnian@uci.edu.

1 Introduction

In the digital era, personal data has become a valuable asset for businesses as they collect, store, and share it to improve consumer satisfaction, personalize recommendations, and enhance their marketing campaigns (Montgomery et al., 2004; Johnson et al., 2005; Tambe et al., 2012). However, the vast amounts of personal information also pose a significant risk that the firms could misuse it in ways that violate consumers' privacy rights or even harm the individuals. For instance, companies may use personal data to target individuals with unwanted marketing, manipulate their behavior (Fleder and Hosanagar, 2009), or use data for reasons that consumers did not consent for (Cimpanu, 2018). Additionally, companies may mishandle personal data, leading to data breaches and identity theft. Such incidents have been observed in recent years, with notable examples including the Facebook-Cambridge Analytica scandal (Cadwalladr and Graham-Harrison, 2018) and data breaches at Yahoo! (McMillan and Knutson, 2017), Marriott (Perlroth et al., 2018) and Equifax (Lee, 2017). Millions of Facebook users' data were harvested without consent and have been shown to be used for political advertising (Cadwalladr and Graham-Harrison, 2018). Three billion Yahoo user accounts, 500 million Marriott guests, and 143 million Equifax users were affected in data breaches. Furthermore, companies may also sell personal data to third parties, which can then exploit them for malicious purposes (Cimpanu, 2018).

The heightened risk of companies misusing and mishandling personal data has prompted various governments to intervene and regulate data collection practices. This process has resulted in the implementation of robust privacy regulations, such as the European Union's General Data Protection Regulation (GDPR), California's California Consumer Privacy Act (CCPA), and several country-specific regulations worldwide. These regulations serve to protect individuals' rights to their data privacy and ensure that their personal data is handled responsibly and ethically. For instance, under the CCPA, companies must inform Californian consumers of their rights regarding their personal data and provide them with a simple and accessible means of asserting these rights, which include accessing and obtaining a copy of their data, deleting it, and opting out of its sale.

The governance of data used by firms presents a complex trade-off. On the one hand, privacy regulations can enhance consumer agency by improving transparency and choice over how personal data is leveraged by companies. The credible threat of auditing and enforcement can also nudge firms to prioritize better data practices (Cusumano et al., 2021). Yet, the very prospect of liability can also constrain companies' willingness to learn from the valuable behavioral signals embedded in personal data. The reduced willingness to proactively anticipate consumer needs from data can diminish their capacity to tailor their offerings to evolving consumer preferences, potentially curbing innovation, and value creation in the digital economy.

Extant literature has recognized this dichotomy by exploring the influence of privacy regulations on firms' strategies and outcomes, including both their advantages and disadvantages (Goldfarb and Tucker, 2011; Campbell et al., 2015; Aridor et al., 2020; Johnson et al., 2021; Janssen et al., 2022; Lefrere et al., 2022; Peukert et al., 2022). However, few studies, except for Zhao et al. (2021), have analyzed the impact of privacy regulations by observing consumer behaviors, such as search, purchase, and return.

To address this research gap, we examine how data protection regulations influence consumers' consumption behaviors. We document the unintended consequence of such regulations by leveraging a natural experiment arising from the CCPA, which grants Californians rights over their personal data collected by firms and took effect on January 1, 2020. The CCPA clearly empowers consumers to protect their data, but how it impacts their consumption behaviors remains unclear. The bill's adoption creates a natural experiment that allows us to study the unintended consequence of data protection regulation for Californians, who are protected by the CCPA rules, compared to non-Californians.

We posit that the CCPA affects Californians' purchase and return behaviors through firms' data collection and targeting strategies. Firms may collect fewer data from Californians and provide less tailored advertising to comply with the regulation and avoid potential liability. Consequently, this reduction hurts their precision in recommendations and product-consumer matching, which alters consumers' purchase and return behaviors. Specifically, in this study we examine the following question empirically: *How do purchase and return behaviors change following the CCPA?*

To answer the question, we leverage a unique dataset from a payment processing gateway that has billions of individual transactions of U.S. consumers. The dataset includes consumer information, such as city-level location, merchant details, credit or debit, and etc. We compile a monthly panel to contrast the purchases and returns for Californians with non-Californians in four neighboring states: Arizona, Oregon, Nevada, and Washington. Empirically, we follow the literature on natural experiments (Smith and Todd, 2005; Liu and Lynch, 2011): we first pre-process the data using propensity score matching (Dehejia and Wahba, 2002) and then employ a difference-in-differences (DiD) identification strategy (Meyer, 1995).

Our results indicate that Californians decrease their purchases by about \$94 per period after the CCPA, a 4.3% drop relative to their matched counterparts in neighboring states. They also increase their returns by \$2, a 3.0% increase in return amount relative to non-Californians. Overall, the loss of commerce in California per period is \$96. Our results are robust to various alternative model specifications and account for macro changes such as the pandemic lockdown effect.

The change in the purchase and return patterns may stem from the challenge of finding a suitable product online. We use another proprietary dataset of browsing behaviors to test our hypothesis about the privacy regulation's effect on consumption patterns. We find that Californians increase their time spent browsing for information on the web, with longer sessions and more page visits than residents of the other four states. This implies that Californians may need more time to search for products that match their preferences. This indirectly supports our hypothesis that Californians were less satisfied with firms' recommendations after the privacy regulation.

We delve into the mechanisms by examining how firms adapt to the CCPA and how this affects consumer behavior. We argue that firms' compliance efforts may influence consumers' online activity. For this test, we contrast firms' ad-related web technologies before and after the CCPA. We exploit a natural experiment based on the CCPA's enforcement criteria: only firms with annual revenues above \$25 million are subject to the CCPA[1]. We split the firms into two groups according to their revenues and measure their usage of ad technologies. We find that firms affected by the CCPA reduce their use of ad technologies, which are often employed for personalized advertising. This suggests that firms limit their targeting practices to comply with the new privacy law.

This study contributes to the extant theory on privacy regulations. We use several proprietary datasets to be among the first to show how data protection laws affect consumer consumption behavior. We also reveal the unintended effects of privacy regulations on consumer welfare. The benefits of data protection regulations are uncertain, but the costs for firms and consumers are clear, as shown by investigations on data protection regulations (Aridor et al., 2020; Johnson et al., 2021; Peukert et al., 2022). Our finding also has practical implications for platforms. As regulatory interventions increase on data collection and processing, platforms may benefit from self-regulating and voluntarily limiting data collection by complementors, to avoid the need for public regulations.

[1] Am I Subject to the CCPA? - Higgs Law

2 Related Literature

Our research integrates two different streams of literature: the value of data in the digital economy and the impact of data protection regulations on platforms. To provide context and define our research questions, we first review the literature and identify the research gaps.

2.1 Data in Digital Economy

Data has emerged as a critical input in shaping digital societies. Literature has extensively examined the advantages of data collection and analysis across various domains. In the business sector, user-generated content such as online reviews and feedback can help enhance service quality and foster product innovation (Ananthakrishnan et al., 2023; Bertschek and Kesler, 2022). Similarly, Niebel et al. (2019) demonstrate a positive relationship between firms' big data analytics usage and product innovation. In finance, data sharing can enhance efficiency by increasing lending to safe borrowers and decreasing default rates (Jappelli and Pagano, 2002; Doblas-Madrid and Minetti, 2013). In healthcare, medical data utilization can substantially lower neonatal mortality rates (Miller and Tucker, 2011) and improve clinical outcomes (Kuperman and Gibson, 2003).

However, data usage by firms also poses challenges for consumer welfare, especially when firms seek to maximize profits through price discrimination based on consumers' willingness-to-pay (Bonatti and Cisternas, 2020; Bar-Gill, 2021). For instance, personalized pricing based on consumers' web-browsing behaviors can substantially increase profits for companies like Netflix, but it may also raise concerns about discrimination against specific consumer groups (Shiller, 2020). Similarly, product reviews and consumption history can enable dynamic pricing and price discrimination, leading to welfare loss for some consumers (Feng et al., 2019; Bonatti and Cisternas, 2020).

Despite an extensive discussion on the benefits and perils of data collection, the net welfare effects on consumers are unclear (Acquisti et al., 2016). On the one hand, data collection can lower the search cost and increase consumer surplus, which is welfare increasing. On the other hand, data collection can also lead some suppliers to price discriminate and extract consumer surplus, which is welfare reducing. Our study adds to this discussion by investigating change in consumer's purchase, return and web usage patterns upon exogenously restricting data collection.

2.2 Data Protection Regulations and Platforms

Several studies point to the benefits that data protection rights confer on consumers. Van Ooijen and Vrabec (2019), for instance, assert that the GDPR can enhance individual control over personal information by reducing cognitive processing and decision-making threats. Ke and Sudhir (2022) theoretically establish that privacy regulations can increase consumer welfare in a competitive market. Aridor et al. (2020) present empirical evidence that consumers leveraged the opt-out feature to restrict firms' data collection, leading to fewer browsing cookies following the GDPR's implementation. Goldberg et al. (2022) document that the collective page views declined after the GDPR's introduction. Finally, Lefrere et al. (2022) finds that GDPR affected websites to improve their tracking practices, albeit for a short term. Overall, privacy regulations can enable consumers to safeguard their privacy and decrease unauthorized personal data collection.

Data protection regulation can have unintended consequences too, as studies have shown. Early studies found that privacy laws hinder technology diffusion (Miller and Tucker, 2009) and reduce online advertising effectiveness (Goldfarb and Tucker, 2011). Recent studies have investigated the impact of new privacy regulations, such as the GDPR and CCPA. For example, Jia et al. (2021) report fewer venture deals in the E.U. than in the U.S. after the GDPR. Bessen et al. (2020) posit that the GDPR imposed new costs upon AI startups, requiring new positions and resource reallocation to address the GDPR concerns. Canayaz et al. (2022) argue that the CCPA hurts firms with voice-AI products, which are heavily reliant on consumer data compared to firms without voice-AI products. Additionally, Johnson et al. (2021), Janssen et al. (2022) and Peukert et al. (2022) demonstrate that the GDPR increased market concentration toward larger players in web technologies or apps, potentially reducing innovation.

While most investigations related to data protection rights have focused on firms and markets, few studies have focused on consumers, with the exception of Zhao et al. (2021) who focus on browsing behavior and search intensity. Studies have yet to explore the direct impact of data protection rights on consumer purchase and satisfaction outcomes.

2.3 Gap and Research Questions

Firstly, the question on how privacy regulations affect actual purchase behavior has received little attention in extant work. While some studies have shown that privacy regulation can lengthen the process of searching for products and services (Zhao et al., 2021), its impact on purchase quantity remains unanswered.

Purchase intention and search efforts may not always reflect actual purchasing behavior, which is crucial for understanding the full effect of privacy regulations on consumer behavior. By studying the actual purchase behavior, we can assess the efficacy of privacy regulations in promoting or discouraging purchases. Therefore, we ask and answer, *do privacy regulations influence the volume of purchases made?*

We also study how privacy regulation affects post-purchase behavior and satisfaction, which have been largely overlooked. Firms may need to modify their marketing strategies under new privacy regulations to avoid potential liability, which could affect their ability to offer personalized and effective recommendations to consumers. This could reduce consumer satisfaction with recommendations and may change consumer behavior, such as increasing product search efforts (Zhao et al., 2021) and decreasing ad clicks (Aridor et al., 2020). However, little is known about how privacy regulation influences post-purchase behaviors, such as product returns, which measure purchase satisfaction. Therefore, our study seeks to fill this gap by examining *whether privacy regulations impact the volume of returns made by consumers*.

3 Research Context and Data

3.1 Background: the California Consumer Privacy Act (CCPA)

The California Consumer Privacy Act of 2018 (CCPA) is a state law intended to provide consumers with the right to protect their personal information gathered by firms. The bill, which took effect on January 1, 2020, grants Californian customers four rights: 1) to know what personal information is collected and how it is used and shared, 2) to delete the data, 3) to opt out of the sale of their personal data, and 4) to not face discrimination for exercising their CCPA rights. The act only covers Californians, but it applies to any firm that does business in California regardless of where they are located. Firms must comply if they have more than $25 million in annual revenue or collect or sell data from over 50,000 Californians, households or devices, or make more than half of their revenue from selling Californians' data. Firms can face penalties for violating the law. The CCPA covers a wide range of personal data and sales. Personal data includes not only physical identity information but also online activity and profile data. Sales include any communication of personal data to another entity for any benefit[2]. This could expose firms to more liability for handling sensitive data and affect consumer behavior more than expected.

[2] It defines sale as "selling, renting, releasing, disclosing, disseminating, making available, transferring, or otherwise communicating orally, in writing, or by electronic or other means, a consumer's personal information by the business to another business or a third party for monetary or other valuable consideration."

One advantage of using the CCPA as the context for an empirical study is that its localized influence allows for a more controlled analysis. While recent studies about the impact of privacy regulations mostly utilize the GDPR, our study using the CCPA provides a unique perspective. The wide-ranging application of the GDPR across numerous European countries and its spillover effects worldwide create challenges for event studies that aim to identify suitable control groups to examine the regulation's causal impact (Johnson, 2022). In contrast, the CCPA is limited to a single state, California, within the United States, ensuring that treatment and control groups are relatively homogeneous, except for the treatment. It enables a more accurate evaluation of the changes before and after the regulation within the same country.

3.2 Data

To study the impact of the CCPA on consumption behavior, we construct a proprietary panel dataset using individual-level transaction data obtained from a large financial data provider. The transaction data include an individual customer's daily purchase and return transactions covering credit card and bank account (debit card) transactions. Each observation in the data corresponds to a single card swipe, such as a debit or credit card. In addition to the transaction history, the data provider provides consumer location information for each month, predicted by their transaction history. It allows us to identify whether a consumer is under the CCPA's effectiveness or not and examine the influence of the privacy regulation on consumer behavior. We focus on consumer transactions that are influenced by firms' targeted promotions by limiting them to 12 focal categories such as cable, subscriptions, or entertainment, and removing business accounts (see appendix A).

We create a balanced panel data set of 101,389 users from January 2019 to December 2020, 12 months before and after the implementation of CCPA, by aggregating transaction data at a month-user level. As noted earlier, the panel dataset includes the treated group consisting of Californians and the control group of those residing in four neighboring states – Arizona, Oregon, Nevada, and Washington. The treated group has 47,133 Californians, and the control group has 54,256 users from the other four states.

Further, we restrict the treated group to the Californians who stayed in California for the study duration. It is unclear whether one is under the effect of CCPA or not if we include him who is moving between California and the other states over the study time window. Besides, it may cause bias in the staggered DiD model suggested by Goodman-Bacon (2021), because each individual may have a different treatment timing. The DiD estimator employing a two-way fixed effects model estimates a weighted average of various effects, and the weights can be negative if we do not exclude those who move across the treatment and control groups. We believe it is appropriate to exclude those samples, as it is unlikely that an individual would relocate to another state solely for the purpose of avoiding or being affected by the new privacy regulation.

Each transaction record includes a timestamp of the transaction, dollar amount, name and city-level address of the merchant, description of the transaction, and whether it is an online or offline transaction. Besides, the panel data also includes each user's monthly location and income classes. The data provider estimates the location and income class based on the consumer's transaction history. The income class is divided into seven brackets, with the higher income class denoting the higher income level. Based on the transaction data, we calculate the ratio of the number of online purchases to the number of total purchases for each individual every month.

We employ an additional dataset from ComScore to study changes in consumers' web browsing behavior post the CCPA. This dataset represents the browsing history of PCs in California and the other four states for the same study periods (from January 2019 to December 2020).

3.2.1 Variable Definitions

Table 1 describes the definition of the variables used in the main analysis.The summary statistics is presented in Table 2.

Table 1: Variable Explanation

PANEL A: Variable Definitions of Consumption Behavior	
Variable	**Description**
Dependent Variables	
$Purchase_{it}$	Dollar amount that consumer *i* purchased at month *t*
$Return_{it}$	Dollar amount that consumer *i* returned at month *t*
Explanatory/Control Variables	
$Treat_i$	Whether a consumer *i* is a Californian
$Post_t$	Whether a month *t* is post the CCPA
$IncomeClass_{it}$	Predicted income class of consumer *i* at month *t*
$OnlinePurchase_{it}$	Fraction of the number of the online purchases and the number of total purchases of consumer *i* at month *t*
PANEL B: Variable Definitions of Browsing Behavior	
Variable	**Description**
Dependent Variables	
$Duration_{jt}$	Number of minutes spent in web browsing for machine *j* at month *t*
$PagesViewed_{jt}$	Number of pages viewed for machine *j* at month *t*
Explanatory Variables	
$Treat_j$	Whether a machine *j* is set in Californian
$Post_t$	Whether a month *t* is post the CCPA

Dependent Variables: In the analysis of consumption behavior, the dependent variables are the individual consumer's monthly expenditure, denoted as $Purchase$, and the monthly refunds, represented by $Return$, both expressed in dollar amount. For the investigation of web browsing behavior, the dependent variables comprise the quantity of minutes dedicated to web browsing within a month , $Duration$ and the count of pages accessed , $PagesViewed$.

Control Variables: To account for potential confounding factors, this study incorporates a consumer's monthly income classification, $IncomeClass$ and the proportion of online purchases relative to the total number of purchases executed each month , $OnlinePurchase$. The details of income class brackets can be found in Table A.1 in Appendix A.

Table 2: Descriptive Statistics

PANEL A: Descriptive Statistics of Consumption Behavior					
	count	mean	s.d.	min.	max.
Purchase	2,433,336	2,652.87	1,921.76	0	14,999
Return	2,433,336	47.97	146.24	0	2,000
IncomeClass	2,433,336	4.56	1.87	1	7
OnlinePurchase	2,433,336	0.24	0.18	0	1
PANEL B: Descriptive Statistics of Browsing Behavior					
	count	mean	s.d.	min.	max.
Duration	20,328	1,709.71	1,711.38	9	10,204
PagesViewed	20,328	1,154.68	1,100.99	10	6,592

4 Empirical Methodology

4.1 Difference-in-Differences (DiD) Analysis

The California Consumer Privacy Act of 2018, approved in June 2018, went into effect on January 1, 2020. The bill creates a natural experimental setting that separates the treated group under the bill's effectiveness – Californians – from the control group who are not under its effectiveness – non-Californians, which allows us to evaluate the influence of the CCPA on consumers' purchase and return behaviors. As noted in Section 3, we remove consumers who relocate between California and the states in our analysis to ensure that one's treatment status is clearly defined.

We use a DiD analysis which is commonly used to identify the effect of "treatment" on treated (Meyer,

1995) by implementing a two-way fixed effects model to estimate the treatment effect:

$$y_{it} = \beta \times (\,Treat_i \times Post_t) + X_{it}^{\prime}\Gamma + \lambda_t + \mu_i + \epsilon_{it}, \tag{1}$$

i refers to a consumer, and t refers to a month: $t \in T = \{-12, \cdots, 11\}$, where $t = 0$ is the month when the CCPA was implemented (January 2020).

The dependent variable y_{it} is consumer i's monthly dollar amount of purchases (returns) at month t for the purchase (return) model. $Treat_i = 1$ is the dummy variable indicating the treated group, and $Post_t$ is the dummy indicating post-treatment periods. β is the coefficient of interest to estimate the effect of the privacy regulation on consumers' shopping behaviors in purchase and return. To account for potential time-invariant unobserved heterogeneity among consumers that may affect their consumption patterns, we incorporate individual consumer fixed effects, μ_i, into our model. Moreover, certain time-variant factors, such as changes in consumers' income or shifts in their usage of shopping channels (online vs. offline), may also impact their consumption patterns. Specifically, the outbreak of COVID-19 in early 2020 substantially increased consumers' online activities due to health concerns and government-imposed lockdown policies, potentially influencing their shopping behaviors as well. To control for these time-variant confounders, we include a vector of variables, X_{it}, which comprises the consumer's income bracket, $IncomeClass_{it}$, and the fraction of online purchases to total purchases, $OnlinePurchase_{it}$. λ_t is year-month fixed effects controlling for common time trends across time; and ϵ_{it} is an error term. The return equation additionally includes the amount of monthly purchases, which can influence the consumer's return amounts and be influenced by the CCPA simultaneously, as a control.

We use a similar model to estimate changes in browsing behavior after the CCPA. The treated group is the machines set up in California, and the control group is those in the other four states. The dependent variable y_{jt} is the browsing duration or pages viewed for machine j at month t. The model includes the interaction of $Treat_j$ and $Post_t$; its coefficient, β captures how Californians' browsing changes after the implementation of the CCPA in contrast to non-Californians. The model, similar to Model 1, includes year-month fixed effects λ_t, machine fixed effect, μ_j, and error term, e_{jt}, but does not include other covariates:

$$y_{jt} = \beta \times (\,Treat_j \times Post_t) + \lambda_t + \mu_j + \epsilon_{jt}. \tag{2}$$

4.2 Matching

Since our setting is quasi-experimental, there is a possibility that the treatment assignment is "nonrandom". To address all concerns related to the quasi-experimental setting, we follow recommendations in the literature. Before conducting the DiD analysis, we perform a Propensity Score Matching (PSM) to account for any potential systematic difference between the treated and control groups for the shopping behavior models (Dehejia and Wahba, 2002) (Equation 1). Using PSM in conjunction with DiD for causal analysis is a frequently used and well-established procedure in literature (Liu and Lynch, 2011; Smith and Todd, 2005). We use one-to-one nearest neighbor matching without replacement to match a sample in the treatment group similar to one in the control group. It matches each sample in the treated group with the closest propensity score in the control group within a given caliper. If a sample in the treated group does not have a matched control within the caliper, it is discarded from our data.

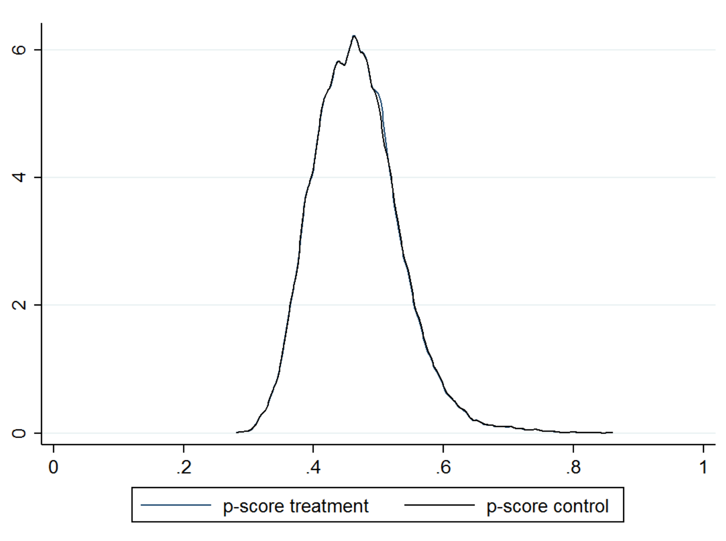
We estimate the propensity score with probit regression. The independent variables in the probit model are income class, the fraction of online purchases, and the number of categories returned and purchased before treatment periods. The model also includes pre-treatment dependent variables (purchase amounts and return amounts) and log transformation of them. The matching procedure yields 43,502 samples in the treated group and control group, respectively. Covariate balance between the treated and control groups is evaluated, and Table 3 summarizes the results, indicating a significant reduction in bias following the matching procedure. Figure 1 illustrates that the distribution of propensity scores is similar between the two groups after matching.

Table 3: Covariate Balance Check: Pre-Treatment

PANEL A: Pre-Matching				
	Mean(Treated)	Mean(Control)	Difference	t-stats
OnlinePurchase	0.217	0.207	-0.010	-33.19***
IncomeClass	4.653	4.354	-0.299	-87.29***
debit_categories	7.373	7.456	0.083	25.21***
credit_categories	0.500	0.461	-0.039	-29.09***
PANEL B: Post-Matching				
	Mean(Treated)	Mean(Control)	Difference	t-stats
OnlinePurchase	0.213	0.213	-0.001	-1.67*
IncomeClass	4.543	4.546	0.003	0.75
debit_categories	7.447	7.455	0.008	2.26**
credit_categories	0.484	0.490	0.006	4.05***

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

Figure 1: Distribution of Propensity Score



4.3 Parallel Trend

The DiD analysis relies on the assumption of a parallel trend between the treated and control groups in the absence of treatment to interpret the estimated treatment effect as a causal impact (Autor, 2003). In our case, if the CCPA had not been adopted, the purchase and return amounts should have moved parallelly over time between Californians and non-Californians. Although this assumption is not generally testable due to the unobservability of counter-factual of post-treatment outcomes for the treated group, it is widely adopted to evaluate the validity of the parallel trend assumption by examining pre-treatment periods.

Our panels include multiple periods of pre- and post-treatment, enabling us to test the parallel trend during the pre-treatment periods. To assess the validity of this assumption, we replace $Post_t$ with λ_t and β with β_t in Equation 1. We compare the difference in the purchase (return) amounts between the treatment and control groups during all other periods with the difference in $t = -1$. As a baseline, we normalize β_{-1} to zero:

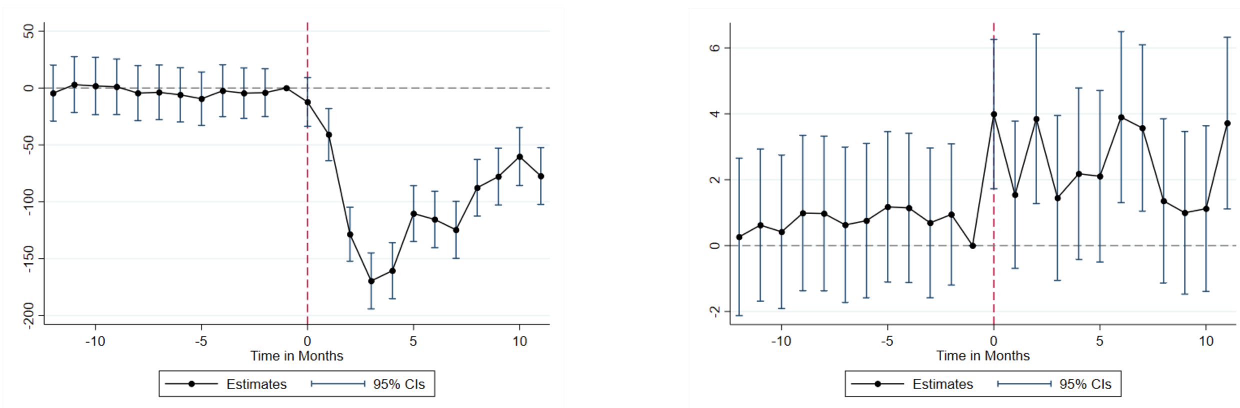
$$y_{it} = \sum_{t \neq -1} \beta_t \times (Treat_i \times \lambda_t) + X_{it}^{'}\Gamma + \lambda_t + \mu_i + \epsilon_{it}. \tag{3}$$

Similarly, we transform the Equation 2 as

$$y_{it} = \sum_{t \neq -1} \beta_t \times (Treat_i \times \lambda_t) + \lambda_t + \mu_i + \epsilon_{it}. \tag{4}$$

The coefficients before the treatment ($t = -2, \cdots, -12$) should be insignificant in Equation 3 and 4 if our models meet the parallel trend assumption.

In Figure 2a and 2b, the dots represent point estimates, and the gray area around each estimate is the 95% confidence interval. As shown in the figures, no coefficients before the treatment were significant except for a few periods in each model. It implies that the treated groups would have changed evenly from the control group if they had not received the treatment. Figure 3a and 3b summarize the estimated results of the browsing behavior models, and they show that the dataset also meets the parallel trend assumption. The F-test fails to reject the null hypothesis that the all pre-treatment coefficients are jointly zero for each respective model, thereby reinforcing the validity of the parallel trend assumption.



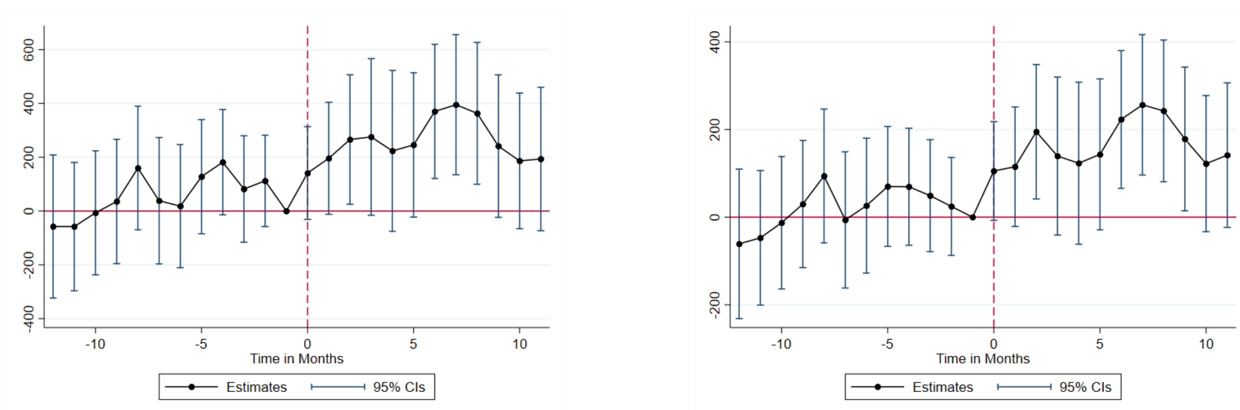
F test that all pre-treatment coefficients are jointly zero: $F(11, 87003) = 0.18, p > F = 0.9984$

(a) Purchases

F test that all pre-treatment coefficients are jointly zero: $F(11, 87003) = 0.19, p > F = 0.9983$

(b) Returns

Figure 2: Parallel Trends: Consumption Behavior



F test that all pre-treatment coefficients are jointly zero: $F(11, 846) = 1.24, p > F = 0.2590$

(a) Duration

F test that all pre-treatment coefficients are jointly zero: $F(11, 846) = 0.82, p > F = 0.6199$

(b) Pages Viewed

Figure 3: Parallel Trends: Browsing Behavior

5 Results

5.1 Main Results

Tables 4 and 5 summarize the effect of the CCPA on consumers' monthly purchase amounts and return amounts, respectively. The estimated results in Table 4 reveal the consistent negative effect of the privacy regulation on purchase regardless of model specifications. The regulation suppresses about \$94 per month; it is about a 4.3% decrease in the monthly spending. The results, presented in Table 5, show that the monthly return amounts increases by about \$2 per month; it is about a 3.0% increase in the monthly return.

Table 4: Effect on Purchase (\$)

	(1)	(2)	(3)	(4)
Treat $\times$ Post	-101.932***	-101.932***	-101.559***	-94.349***
	(5.669)	(5.669)	(5.581)	(5.584)
Post	86.993***			
	(4.003)			
IncomeClass=2			101.541***	103.619***
			(6.288)	(6.334)
IncomeClass=3			170.146***	173.435***
			(8.113)	(8.162)
IncomeClass=4			215.593***	219.974***
			(9.244)	(9.295)
IncomeClass=5			270.371***	275.412***
			(9.822)	(9.871)
IncomeClass=6			342.842***	348.244***
			(10.766)	(10.811)
IncomeClass=7			422.225***	427.425***
			(12.163)	(12.201)
OnlinePurchase				-554.180***
				(12.119)
N	2,088,096	2,088,096	2,088,096	2,088,096
Individual fixed effects	yes	yes	yes	yes
Year-month fixed effects	no	yes	yes	yes
Robust standard errors	yes	yes	yes	yes

Clustered standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

These findings imply that the privacy regulation meant to defend customers' rights to their personal data may have an unanticipated detrimental influence on their consumption and satisfaction. Firms aiming to comply with the new regulation may reduce the collection of customer data and the delivery of tailored advertisements and their quality. As a result, the fewer targeted advertisements after the CCPA's implementation cannot attract customers as effectively as those before the regulation, consequently making Californian consumers purchase less than before. Furthermore, customers are less satisfied with their purchases due to poorly customized recommendations, and they return more than before.

Table 5: Effect on Return ($)

	(1)	(2)	(3)	(4)	(5)
Treat × Post	0.600	2.198***	2.176***	2.186***	1.764***
	(0.478)	(0.464)	(0.464)	(0.464)	(0.463)
Post	5.933***	4.569***			
	(0.346)	(0.338)			
Purchase		0.016***	0.015***	0.015***	0.016***
		(0.000)	(0.000)	(0.000)	(0.000)
IncomeClass=2				-1.346*	-1.495**
				(0.722)	(0.721)
IncomeClass=3				-0.896	-1.134
				(0.862)	(0.861)
IncomeClass=4				-0.836	-1.151
				(0.937)	(0.936)
IncomeClass=5				0.143	-0.224
				(0.964)	(0.962)
IncomeClass=6				1.915*	1.510
				(1.023)	(1.021)
IncomeClass=7				5.608***	5.199***
				(1.117)	(1.115)
OnlinePurchase					34.083***
					(0.884)
N	2,088,096	2,088,096	2,088,096	2,088,096	2,088,096
Individual fixed effects	yes	yes	yes	yes	yes
Year-month fixed effects	no	no	yes	yes	yes
Robust standard errors	yes	yes	yes	yes	yes

Clustered standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The findings in Table 6 – the increase in web surfing duration and pages viewed – provide indirect evidence to explain the changes in consumption behaviors after the regulation. Since Californian consumers receive fewer targeted promotions than before, they should look for commodities that fit their preferences by themselves. Therefore, they may need to spend more time on web surfing and visit more pages to seek products that suit their taste.

Table 6: Effect on Browsing Behavior

	(1)	(2)	(3)	(4)
	Duration	Duration	Pages Viewed	Pages Viewed
Treat × Post	205.346***	205.346***	145.737***	145.737***
	(73.468)	(73.508)	(46.573)	(46.598)
Post	0.577		-73.959*	
	(59.602)		(38.133)	
N	20,328	20,328	20,328	20,328
Individual fixed effects	yes	yes	yes	yes
Year-month fixed effect	no	yes	no	yes
Robust standard errors	yes	yes	yes	yes

Clustered standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5.2 Robustness Check

5.2.1 Controlling the Impact of the COVID-19 Pandemic

The COVID-19 pandemic has greatly impacted the U.S. economy and society since March 2020, resulting in significant shifts in consumer behavior due to social distancing and lockdown policies. Thus, it is essential to consider its impact when analyzing the effects of privacy regulation on consumer behavior as the pandemic-induced changes in consumption patterns and preferences. For example, consumers may have shifted their purchases online due to stay-at-home orders or concerns about contracting the virus in public spaces. This shift may confound the analysis if not properly addressed.

We employ two strategies in our analysis to account for these potential confounding effects. First, we include a control variable in Equation 1 that indicates whether each state government implemented a stay-at-home order each month. This variable is set to 1 on and after the month when the state government implemented the order. By controlling for stay-at-home orders, we can isolate the effect of privacy regulation from changes in consumer behavior caused by these orders. Second, we control the COVID-19 cases per state population for each state. By doing so, we can adjust for any differences in consumer behavior across states that could have resulted from variations in the COVID-19 prevalence. For instance, states with higher COVID-19 cases may have experienced more significant changes in consumption patterns than states with lower case numbers. By controlling for the stay-at-home orders or the COVID-19 cases per state population, we can more accurately estimate the effect of privacy regulation on consumer behavior while accounting for potential confounding factors introduced by the pandemic.

The results in Table B.1 and B.2 in the Appendix B show that controlling for the effect of the COVID-19 does not significantly alter the significance and magnitude of the effects in the main analysis. In Table B.1, we can see that the coefficient of the treatment effect remains negative and significant across all three specifications. The results indicate that privacy regulation has a negative effect on purchases even after controlling for the stay-at-home orders or the COVID-19 cases per state population. Similarly, Table B.2 indicates that privacy regulation has a positive effect on returns even after controlling for the stay-at-home orders or the COVID-19 cases per state population. Overall, these results suggest that our findings of the impact of the CCPA on consumers' consumption behavior are robust to potential confounding effects introduced by the COVID-19 pandemic.

5.2.2 Addressing Serial Correlation and Estimating Standard Error

To address the potential presence of serial correlation in the dependent variable, we apply a collapsed DiD analysis recommended by Bertrand et al. (2004) to our dataset. Specifically, we collapse our data into pre-and post-treatment periods by calculating the average of the respective variables. For the factor variable, $IncomeClass$, the mode is utilized as a substitute for the average. The findings from the collapsed analysis, presented in Table B.3 in the appendix, consistently support the main analysis, indicating a significant negative impact on purchase and a significant positive impact on return.

We also estimate standard errors using bootstrap followed by Austin and Small (2014) and Smith and Todd (2005), which recommended to estimate standard errors using bootstrap for the PSM without replacement. Table B.4 and B.4 exhibit the findings derived from 200 replications, utilizing bootstrap standard errors, and the results are consistent.

5.3 Potential Mechanism

Our primary analysis indicates that privacy regulation intended to protect consumers' personal data has an unintended negative impact on their welfare. Consumers purchase less but return more and spend more time searching for products that meet their preferences. However, the mechanism behind the shift in consumer behavior after the new privacy regulation's implementation still needs to be discovered. One plausible explanation is that firms reduce data collection and trade to avoid future liability and comply with the new regulation, which reduces the quantity and quality of tailored advertising. As discussed by Aridor et al. (2020) and Goldfarb and Tucker (2011), new privacy regulations have decreased the effectiveness of online advertisements, suggesting that consumers can be less attracted to post-regulation recommendations from advertisers. This section explores this mechanism by reviewing two critical aspects of business websites: 1) advertising web technology usage and 2) product descriptions.

5.3.1 Changes in Advertising Web Technology Usage

We explore the mechanisms behind the shift in consumer behavior following the implementation of the CCPA by examining changes in firms' web advertising technologies used on their websites. As online advertisements are typically personalized to individual consumers, a decrease in advertising web technology may suggest that firms reduce personalized advertising. To explore this issue, we use the threshold requirement of the CCPA, which applies to firms with an annual gross revenues of over $25 million. We distinguish firms subject to the CCPA and those not and exploit a DiD approach to compare changes in the number of advertising web technologies employed by firms on their websites.

We follow several steps to classify firms as subject to the CCPA. We collect the U.S./North American portion of annual gross revenue for 31 public companies in 2019 from their 10-k filings. We then compare the revenues with the transaction amount in our primary dataset to obtain the ratio between revenue and transaction amount. We use this ratio and the transaction data to estimate the revenue for other companies and classify firms with an annual revenue of over $25 million as subject to the CCPA or not. Finally, we scrape the web technology usage of firms' websites and match it to our estimated revenue.

The results in Table 7 indicate that firms subject to the CCPA reduce the number of advertising web technologies employed on their websites after the CCPA compared to those not subject to the regulation. This finding suggests that the CCPA may prompt CCPA-affected firms to reduce the use of personalized advertising. Firms may reduce the use of advertising web technologies to comply with the new privacy regulation. Specifically, the analysis reveals that CCPA-affected firms used 1.04 fewer advertising web technologies on their websites post-CCPA enforcement.[3]

Table 7: Changes in Advertising Web Technology Usage

	(1)	(2)
Treat $\times$ Post	-1.040***	-1.040***
	(0.379)	(0.379)
Post	0.461	
	(0.333)	
N	117720	117720
Individual fixed effects	yes	yes
Year-month fixed effect	no	yes
Robust standard errors	yes	yes

Clustered standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

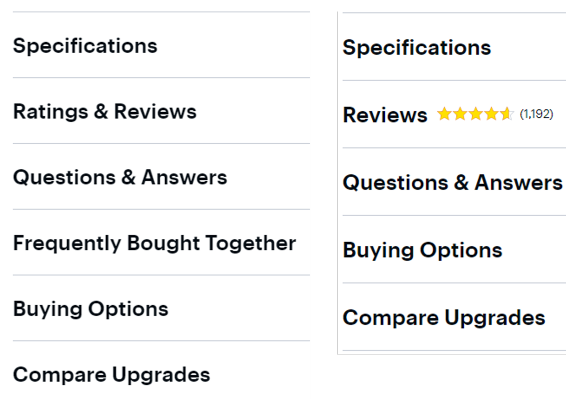
5.3.2 Anecdotal Evidence: Changes in Product Description

To explore the possible mechanisms, we review product description pages on two prominent e-commerce platforms, BestBuy.com and Amazon.com. Our analysis aims to identify any modifications in web or recommendation technologies employed by these platforms which can influence consumer behavior.

[3]Firms use an average of 7.79 advertising web technologies on their websites before the treatment.

Our examination reveals that companies might have restricted product recommendations and product information derived from users' personal data, potentially contributing to the observed shift in purchasing patterns. A careful analysis indicates that modifications in product pages occurred following the implementation of new privacy regulations. For instance, BestBuy.com eliminated the "Frequently Bought Together" section from its product descriptions, while Amazon.com transitioned from displaying product ratings based on customer groups and interests to ratings centered on product features. Figures 4a and 4b depict product description pages on Amazon.com and BestBuy.com, respectively. The figure on the left represents the year 2019, while the one on the right portrays the year 2020. These adjustments suggest that firms proactively minimize the collection of personal information to avert potential future liabilities.

(a) Product Rating Tab on Amazon.com



(b) Product Description Tab on Bestbuy.com

Figure 4: Changes in Product Description

6 Discussion and Conclusion

6.1 Discussion

In this study, we have investigated the impact of data protection regulations on consumers' consumption behavior, with a specific focus on the California Consumer Privacy Act. The CCPA grants Californians new rights to control their personal data, which had previously been freely collected and sold by firms. While data serves as a valuable resource in the digital economy, allowing firms to improve their businesses and explore new opportunities, it also has the potential to be misused or abused in ways that may harm consumers' welfare, such as through price discrimination. Undoubtedly, the new regulation enhances consumers' data privacy and offers protection against potential misuse. However, it also gives rise to unintended consequences. We examined how privacy regulations might impede firms from discovering latent consumer preferences embedded in personal data, which can subsequently affect consumer welfare.

Our findings reveal that consumers in California reduce their purchases following the enactment of the CCPA compared to their neighboring non-Californian counterparts. We demonstrate that this decrease in consumption is accompanied by a decline in consumer satisfaction with their purchases, as evidenced by an increase in returns. Furthermore, the regulation leads to extended web surfing hours and visits, possibly suggesting that consumers need to devote more time to shopping to find products that meet their needs.

We propose a potential mechanism that leads to changes in consumer behavior. Firms affected by the CCPA reduce their use of advertising-related web technologies on their websites after the regulation comes into effect. Additionally, we provide anecdotal evidence that companies may restrict product recommendations and the product information based on user's personal data, as observed through changes in web pages on two major e-commerce platforms. Firms' efforts to minimize potential liability and comply with the new regulation impede their ability to uncover latent preferences within consumer data. As a result, they may offer fewer personalized advertisements, and the effectiveness of these ads may be diminished, ultimately impacting consumer welfare negatively.

The research findings from our study align with extant literature, indicating that privacy regulation has an adverse impact on the effectiveness of online advertisements in attracting consumers (Goldfarb and Tucker, 2011; Aridor et al., 2020) and raises search costs for consumers during shopping (Zhao et al., 2021). Our study contributes novel insights to the literature by demonstrating that firms' diminished targeting capabilities not only decrease consumers' purchasing intent but also lead to an actual reduction in consumption behavior and satisfaction. Taken together, our results underscore the trade-offs that arise from implementing privacy regulations. While they serve to protect individuals' data privacy and limit potential misuse, they may also lead to unintended consequences for consumer behavior and overall satisfaction.

6.2 Conclusion

In conclusion, this study contributes to both theoretical and practical understanding of the impact of privacy regulations on consumer behavior and firm performance. Our research adds to extant studies on the unintended consequences of privacy regulation. Our study suggests that privacy regulations, intended to protect consumers' rights to their privacy, may have unintended negative impacts on both firms and consumers.

Privacy regulations can have detrimental effects on firms by reducing revenues. Moreover, the regulations negatively impact consumers by altering purchasing patterns, resulting in reduced purchase satisfaction and increased search costs.

Our findings also offer practical implications. These results provide crucial managerial insights for firms. Businesses, particularly those heavily relying on consumer data for advertising purposes, may experience a drop in sales and revenue due to the regulation. Consequently, these companies may need to increase their marketing efforts to compensate for the decline in sales and revenue and diversify their marketing strategies beyond targeted advertisements. Furthermore, the impact can be more severe for smaller firms, which may not have in-house data and rely on third-party institutions for consumer data, as the regulation explicitly protects consumers' right to opt-out of selling their personal information.

Our findings also suggest implications for policymakers in recognizing that privacy regulation has unintended consequences, as previously discussed. For example, firms may face increased costs to protect data privacy to comply with the regulation and experience a drop in sales and revenue. The regulation also imposes additional search costs on consumers for shopping and reduces purchase satisfaction due to increased information friction between firms and consumers following the regulation. Policymakers should assess these unintended costs and weigh them against the regulation's intended benefits, such as reduced future risks of data breaches and their associated social harms.

By examining the impact of privacy regulations on consumer behavior and firm performance, this research offers valuable insights for both scholars and practitioners. It highlights the need for a holistic approach to privacy regulation, taking into account not only the intended benefits but also the potential unintended consequences for consumers and businesses.

References

Acquisti, A., Taylor, C., and Wagman, L. (2016). The economics of privacy. *Journal of economic Literature*, 54(2):442–492.

Ananthakrishnan, U. M., Proserpio, D., and Sharma, S. (2023). I hear you: Does quality improve with customer voice? *SSRN ID 3467236*.

Aridor, G., Che, Y.-K., and Salz, T. (2020). The economic consequences of data privacy regulation: Empirical evidence from gdpr. *NBER working paper*, (w26900).

Austin, P. C. and Small, D. S. (2014). The use of bootstrapping when using propensity-score matching without replacement: a simulation study. *Statistics in medicine*, 33(24):4306–4319.

Autor, D. H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of labor economics*, 21(1):1–42.

Bar-Gill, O. (2021). Price discrimination with consumer misperception. *Applied Economics Letters*, 28(10):829–834.

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

Bertschek, I. and Kesler, R. (2022). Let the user speak: Is feedback on facebook a source of firms' innovation? *Information Economics and Policy*, 60:100991.

Bessen, J. E., Impink, S. M., Reichensperger, L., and Seamans, R. (2020). Gdpr and the importance of data to ai startups. *NYU Stern School of Business*.

Bonatti, A. and Cisternas, G. (2020). Consumer scores and price discrimination. *The Review of Economic Studies*, 87(2):750–791.

Cadwalladr, C. and Graham-Harrison, E. (2018). Revealed: 50 million facebook profiles harvested for cambridge analytica in major data breach. *The Guardian*, 17(1):22.

Campbell, J., Goldfarb, A., and Tucker, C. (2015). Privacy regulation and market structure. *Journal of Economics & Management Strategy*, 24(1):47–73.

Canayaz, M., Kantorovitch, I., and Mihet, R. (2022). Consumer privacy and value of consumer data. *Swiss Finance Institute Research Paper*, (22-68).

Cimpanu, C. (2018). Google restricts which android apps can request call log and sms permissions.

Cusumano, M. A., Gawer, A., and Yoffie, D. B. (2021). Can self-regulation save digital platforms? *Industrial and Corporate Change*, 30(5):1259–1285.

Dehejia, R. H. and Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics*, 84(1):151–161.

Doblas-Madrid, A. and Minetti, R. (2013). Sharing information in the credit market: Contract-level evidence from us firms. *Journal of financial Economics*, 109(1):198–223.

Feng, J., Li, X., and Zhang, X. (2019). Online product reviews-triggered dynamic pricing: Theory and evidence. *Information Systems Research*, 30(4):1107–1123.

Fleder, D. and Hosanagar, K. (2009). Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management science*, 55(5):697–712.

Goldberg, S., Johnson, G., and Shriver, S. (2022). Regulating privacy online: An economic evaluation of the gdpr. *SSRN ID 3421731*.

Goldfarb, A. and Tucker, C. E. (2011). Privacy regulation and online advertising. *Management science*, 57(1):57–71.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

Janssen, R., Kesler, R., Kummer, M. E., and Waldfogel, J. (2022). Gdpr and the lost generation of innovative apps. Technical report, National Bureau of Economic Research.

Jappelli, T. and Pagano, M. (2002). Information sharing, lending and defaults: Cross-country evidence. *Journal of Banking & Finance*, 26(10):2017–2045.

Jia, J., Jin, G. Z., and Wagman, L. (2021). The short-run effects of the general data protection regulation on technology venture investment. *Marketing Science*, 40(4):661–684.

Johnson, G. (2022). Economic research on privacy regulation: Lessons from the gdpr and beyond.

Johnson, G., Shriver, S., and Goldberg, S. (2021). Privacy & market concentration: Intended & unintended consequences of the gdpr. *SSRN ID 3477686*.

Johnson, P. M., Kou, H., Paulding, M., Zhang, Q., Kagawa, A., and Yamashita, T. (2005). Improving software development management through software project telemetry. *IEEE software*, 22(4):76–85.

Ke, T. T. and Sudhir, K. (2022). Privacy rights and data security: Gdpr and personal data markets. *Management Science*.

Kuperman, G. J. and Gibson, R. F. (2003). Computer physician order entry: benefits, costs, and issues. *Annals of internal medicine*, 139(1):31–39.

Lee, M. (2017). Equifax data breach impacts 143 million americans. *Forbes*.

Lefrere, V., Warberg, L., Cheyre, C., Marotta, V., and Acquisti, A. (2022). Does privacy regulation harm content providers? a longitudinal analysis of the impact of the gdpr. *SSRN ID 4239013*.

Liu, X. and Lynch, L. (2011). Do agricultural land preservation programs reduce farmland loss? evidence from a propensity score matching estimator. *Land Economics*, 87(2):183–201.

McMillan, R. and Knutson, R. (2017). Yahoo triples estimate of breached accounts to 3 billion. *The Wall Street Journal*.

Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of business & economic statistics*, 13(2):151–161.

Miller, A. R. and Tucker, C. (2009). Privacy protection and technology diffusion: The case of electronic medical records. *Management science*, 55(7):1077–1093.

Miller, A. R. and Tucker, C. E. (2011). Can health care information technology save babies? *Journal of Political Economy*, 119(2):289–324.

Montgomery, A. L., Li, S., Srinivasan, K., and Liechty, J. C. (2004). Modeling online browsing and path analysis using clickstream data. *Marketing science*, 23(4):579–595.

Niebel, T., Rasel, F., and Viete, S. (2019). Big data–big gains? understanding the link between big data analytics and innovation. *Economics of Innovation and New Technology*, 28(3):296–316.

Perlroth, N., Tsang, A., and Satariano, A. (2018). Marriott hacking exposes data of up to 500 million guests. *The New York Times*, 30.

Peukert, C., Bechtold, S., Batikas, M., and Kretschmer, T. (2022). Regulatory spillovers and data governance: Evidence from the gdpr. *Marketing Science*.

Shiller, B. R. (2020). Approximating purchase propensities and reservation prices from broad consumer tracking. *International Economic Review*, 61(2):847–870.

Smith, J. A. and Todd, P. E. (2005). Does matching overcome lalonde's critique of nonexperimental estimators? *Journal of econometrics*, 125(1-2):305–353.

Tambe, P., Hitt, L. M., and Brynjolfsson, E. (2012). The extroverted firm: How external information practices affect innovation and productivity. *Management Science*, 58(5):843–859.

Van Ooijen, I. and Vrabec, H. U. (2019). Does the gdpr enhance consumers' control over personal data? an analysis from a behavioural perspective. *Journal of consumer policy*, 42:91–107.

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

A Appendix: Data

A.1 List of the 12 categories

- Automotive/Fuel
- Cable/Satellite/Telecom
- Electronics/General Merchandise
- Entertainment/Recreation
- Gifts
- Groceries
- Home Improvement
- Office Expenses
- Personal/Family
- Pets/Pet Care
- Subscriptions/Renewals
- Travel

A.2 Removing Business Accounts

To remove business accounts, we identify users who spent more than $15,000 or returned more than $2,000 once in a month and consider them as business accounts. We then remove these accounts as part of our standard procedure. In addition, we remove users who spent less than an average $100 per month to eliminate any potential unusual transactions.

A.3 Income Class Brackets

Table A.1: Income Class Brackets

Income Class	Range
1	$0 - 25k
2	$25k - 45k
3	$45k - 60k
4	$60k - 75k
5	$75k - 100k
6	$100k - 150k
7	$150k +

B Appendix: Robustness Check

B.1 Controlling the Impact of the COVID-19 Pandemic

Table B.1: Effect on Purchase: Controlling COVID-19

	(1)		(2)		(3)	
Treat × Post	-92.68***	(5.715)	-94.34***	(5.584)	-96.46***	(5.594)
IncomeClass=2	103.7***	(6.334)	103.7***	(6.334)	103.8***	(6.334)
IncomeClass=3	173.5***	(8.162)	173.5***	(8.162)	173.6***	(8.162)
IncomeClass=4	220.0***	(9.295)	220.1***	(9.295)	220.2***	(9.294)
IncomeClass=5	275.5***	(9.871)	275.5***	(9.871)	275.6***	(9.871)
IncomeClass=6	348.3***	(10.81)	348.4***	(10.81)	348.5***	(10.81)
IncomeClass=7	427.5***	(12.20)	427.6***	(12.20)	427.7***	(12.20)
OnlinePurchase	-554.2***	(12.12)	-554.0***	(12.12)	-554.1***	(12.12)
Covid19	-13.12	(10.01)	-85.14***	(25.57)	-5342.0***	(1252.6)
N	2,088,096		2,088,096		2,088,096	
Individual-fixed effects	yes		yes		yes	
Month-fixed effects	yes		yes		yes	
Robust standard errors	yes		yes		yes	
Covid19	stay-at-home		cases per population		deaths per population	

Clustered standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.2: Effect on Return: Controlling COVID-19

	(1)		(2)		(3)	
Treat × Post	1.696***	(0.472)	1.764***	(0.463)	1.784***	(0.465)
Purchase	0.0157***	(0.000138)	0.0157***	(0.000138)	0.0157***	(0.000138)
IncomeClass=2	-1.497*	(0.721)	-1.497*	(0.721)	-1.496*	(0.721)
IncomeClass=3	-1.135	(0.861)	-1.136	(0.861)	-1.136	(0.861)
IncomeClass=4	-1.152	(0.936)	-1.154	(0.936)	-1.153	(0.936)
IncomeClass=5	-0.226	(0.962)	-0.227	(0.962)	-0.226	(0.962)
IncomeClass=6	1.508	(1.021)	1.508	(1.021)	1.508	(1.021)
IncomeClass=7	5.194***	(1.115)	5.195***	(1.115)	5.196***	(1.115)
OnlinePurchase	34.08***	(0.884)	34.08***	(0.884)	34.08***	(0.884)
Covid19	0.537	(0.945)	2.102	(2.412)	49.68	(112.8)
N	2,088,096		2,088,096		2,088,096	
Individual-fixed effects	yes		yes		yes	
Month-fixed effects	yes		yes		yes	
Robust standard errors	yes		yes		yes	
Covid19	stay-at-home		cases per population		deaths per population	

Clustered standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.2 Collapsed DiD Analysis

Table B.3: Collapsed DiD Analysis of Purchases and Returns

	(1) Purchase		(2) Return	
Post × Treat	-106.0***	(5.957)	1.761***	(0.464)
Purchase			0.0169***	(0.000210)
Post	11.73**	(4.472)	2.292***	(0.342)
Treat	-6.269	(8.127)	-1.363**	(0.499)
IncomeClass=2	625.0***	(8.077)	-1.410	(0.889)
IncomeClass=3	936.2***	(8.548)	-4.866***	(0.947)
IncomeClass=4	1307.4***	(9.819)	-8.063***	(1.028)
IncomeClass=5	1746.7***	(9.087)	-6.966***	(0.943)
IncomeClass=6	2443.6***	(10.09)	-5.780***	(0.977)
IncomeClass=7	3064.2***	(14.72)	2.843**	(1.056)
OnlinePurchase	316.5***	(34.12)	46.49***	(1.712)
Constant	851.1***	(9.427)	-5.187***	(0.948)
N	174,008		174,008	
Individual-fixed effects	yes		yes	
Robust standard errors	yes		yes	

Clustered standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.3 Bootstrap Standard Errors

Table B.4: Effect on Purchase ($): Bootstrap S.E.

	(1)	(2)	(3)	(4)
Treat × Post	-101.932***	-101.932***	-101.559***	-94.349***
	(5.613)	(5.613)	(5.538)	(5.524)
Post	86.993***			
	(3.918)			
IncomeClass=2			101.541***	103.619***
			(6.228)	(6.235)
IncomeClass=3			170.146***	173.435***
			(8.001)	(8.085)
IncomeClass=4			215.593***	219.974***
			(9.365)	(9.412)
IncomeClass=5			270.371***	275.412***
			(9.768)	(9.768)
IncomeClass=6			342.842***	348.244***
			(10.888)	(10.884)
IncomeClass=7			422.225***	427.425***
			(12.366)	(12.341)
OnlinePurchase				-554.180***
				(10.944)
N	2,088,096	2,088,096	2,088,096	2,088,096
Individual-fixed effects	yes	yes	yes	yes
Month-fixed effects	no	yes	yes	yes
Standard errors	bootstrap	bootstrap	bootstrap	bootstrap

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

Table B.5: Effect on Return ($): Bootstrap S.E.

	(1)	(2)	(3)	(4)	(5)
Treat × Post	0.600	2.198***	2.176***	2.186***	1.764***
	(0.508)	(0.492)	(0.492)	(0.491)	(0.488)
Post	5.933***	4.569***			
	(0.380)	(0.369)			
Purchase		0.016***	0.015***	0.015***	0.016***
		(0.000)	(0.000)	(0.000)	(0.000)
IncomeClass=2				-1.346*	-1.495*
				(0.768)	(0.768)
IncomeClass=3				-0.896	-1.134
				(0.889)	(0.887)
IncomeClass=4				-0.836	-1.151
				(0.986)	(0.985)
IncomeClass=5				0.143	-0.224
				(0.999)	(0.998)
IncomeClass=6				1.915*	1.510
				(1.080)	(1.077)
IncomeClass=7				5.608***	5.199***
				(1.189)	(1.190)
OnlinePurchase					34.083***
					(0.841)
N	2,088,096	2,088,096	2,088,096	2,088,096	2,088,096
Individual-fixed effects	yes	yes	yes	yes	yes
Month-fixed effects	no	no	yes	yes	yes
Standard errors	bootstrap	bootstrap	bootstrap	bootstrap	bootstrap

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

B.4 Log Transformation of Purchase and Return

Table B.6: Effect on Log (Purchase)

	(1)	(2)	(3)	(4)
Treat × Post	-0.045***	-0.045***	-0.045***	-0.043***
	(0.003)	(0.003)	(0.003)	(0.003)
Post	0.029***			
	(0.002)			
IncomeClass=2			0.101***	0.102***
			(0.005)	(0.005)
IncomeClass=3			0.155***	0.156***
			(0.006)	(0.006)
IncomeClass=4			0.181***	0.182***
			(0.006)	(0.006)
IncomeClass=5			0.206***	0.208***
			(0.007)	(0.007)
IncomeClass=6			0.233***	0.234***
			(0.007)	(0.007)
IncomeClass=7			0.254***	0.255***
			(0.007)	(0.007)
OnlinePurchase				-0.174***
				(0.008)
N	2,088,096	2,088,096	2,088,096	2,088,096
Individual fixed effects	yes	yes	yes	yes
Year-month fixed effects	no	yes	yes	yes
Robust standard errors	yes	yes	yes	yes

Clustered standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.7: Effect on Log (Return)

	(1)	(2)	(3)	(4)	(5)
Treat × Post	0.013**	0.037***	0.036***	0.036***	0.030***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Post	0.087***	0.071***			
	(0.004)	(0.004)			
Purchase_ln		0.543***	0.524***	0.523***	0.528***
		(0.004)	(0.004)	(0.004)	(0.004)
IncomeClass=2				-0.011	-0.013
				(0.009)	(0.009)
IncomeClass=3				0.008	0.004
				(0.011)	(0.011)
IncomeClass=4				0.023*	0.018
				(0.012)	(0.012)
IncomeClass=5				0.037***	0.032**
				(0.013)	(0.013)
IncomeClass=6				0.072***	0.066***
				(0.013)	(0.013)
IncomeClass=7				0.130***	0.124***
				(0.015)	(0.014)
OnlinePurchase					0.481***
					(0.011)
N	2,088,096	2,088,096	2,088,096	2,088,096	2,088,096
Individual fixed effects	yes	yes	yes	yes	yes
Year-month fixed effects	no	no	yes	yes	yes
Robust standard errors	yes	yes	yes	yes	yes

Clustered standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.