Information Economics and Policy 60 (2022) 100985

Contents lists available at ScienceDirect

Information Economics and Policy

journal homepage: www.elsevier.com/locate/iep

Online privacy and market structure: Theory and evidence[☆]

Lorien Sabatino[a], Geza Sapi[b,1,*]

[a] Department of Management and Production Engineering, Politecnico di Torino Italy
[b] European Commission DG COMP - Chief Economist Team and Düsseldorf Institute for Competition Economics (DICE), Heinrich Heine University of Düsseldorf. Belgium

ARTICLE INFO

Article history:
Received 28 January 2021
Revised 4 April 2022
Accepted 10 June 2022
Available online 12 June 2022

JEL classification:
D43
L86
M37
M38

Keywords:
Privacy
Competition
Regulation
Eprivacy directive

ABSTRACT

This paper investigates how privacy regulation affects the structure of online markets. We empirically analyse the effects of the 2009 ePrivacy Directive in Europe on firm revenues. Our results indicate that, if any, only large firms were weakly negatively affected by the implementation of the Directive. We also provide a simple theoretical model predicting an avenue how privacy regulation may predominantly influence the revenues and profits of larger firms, even if - as some of our evidence indicates - these larger firms may actually offer more privacy than smaller rivals. Our results suggest that while privacy regulation is not without costs to businesses, it need not distort competition to the favour of larger firms.

1. Introduction

Firms in the digital economy collect customer data at an unprecedented rate. Electronic commerce in physical and digital goods is fuelled by recommendation engines: algorithms that rely on user data on demographics, previous purchases, and other preferences to predict products and services an online shopper may be interested in. Commentators often attribute a large share of the stellar success of internet giants like Amazon and Netflix to the ability of these firms to recommend products to their users based on data and analytics (Arora, 2016).

At the same time, consumers are increasingly mindful about online privacy. While in 2011 around 40% of surveyed Europeans were concerned about their behaviour being recorded through the internet when browsing, downloading files, and accessing content online (European Commission, 2011, page 67), in 2015 less than a quarter of Europeans reported trusting online businesses to protect their personal data (Eurobarometer, 2015, page 25).[2]

As a response to the increased privacy concerns, the European Union (EU) put into force a series of privacy regulations since the early 2000s. The 2018 General Data Protection Regulation (GDPR) empowered European data protection authorities to issue hefty fines comparable to those in antitrust on firms violating data protection rules. In the same year, the European Commission put forward a proposal for an EU-wide ePrivacy Regulation to replace the current ePrivacy Directive of 2009 (Eur, 2018). The proposal received a lot of criticism from industry representatives, expressing concerns about the effect of stricter online privacy rules on the competitiveness of European businesses, adding that the regulation may benefit large firms.[3]

[☆] We are grateful to the Editor and two anonymous referees for their valuable suggestions which greatly improved the paper. We would also like to thank Carlo Cambini, Grazia Cecere, Szabolcs Lorincz, Ulrich Laitenberger, and Mario Pagliero, as well as seminar participants at the 11th Paris Conference on Digital Economics, Bergen Competition Policy Conference 2019, X Workshop on Individual Behaviour and Economic Outcomes, 17th ZEW Conference on the Economics of Information and Communication Technologies, CRESSE Conference 2019, and 17th Media Economics Workshop in Braga for helpful comments and discussions. All remaining errors are ours.

[*] Corresponding author.
E-mail addresses: lorien.sabatino@polito.it (L. Sabatino), sapi@dice.uni-duesseldorf.de (G. Sapi).

[1] The views expressed in this article are solely those of the authors and may not, under any circumstances, be regarded as representing an official position of the European Commission.

[2] Growing online privacy concerns likely followed the general trend in internet use, and likely reach back to the early 2000s. While in 2002 around 39% of households in the EU15 had internet access, this increased to 59% in 2007, 79% in 2012, and 89% in 2017. By 2020 almost all Western-European households had internet access (Eurostat, 2021).

https://doi.org/10.1016/j.infoecopol.2022.100985
0167-6245/© 2022 Elsevier B.V. All rights reserved.

We do three things in this paper. First, we empirically investigate the effect of the 2009 revision of the European ePrivacy Directive (2009/136/EC) on the market structure in e-commerce. In particular, we employ a *difference-in-difference-in-differences* (DDD) model to identify the effect of increased privacy regulation on the revenues of European firms active in the online retail sector. We exploit time variation in the implementation of the ePrivacy Directive by EU Member States. Our data allow comparing European e-commerce businesses to a control group consisting of firms primarily active in North America (U.S. and Canada) as well as brick-and-mortar firms selling similar consumer discretionary products as their online counterparts. Second, we propose a simple theoretical model of competition in the online retail sector that captures the main trade-off between the informativeness of advertising and the degree of privacy intrusion (Tucker, 2012). Our model predicts that privacy regulation affects primarily the profits of larger firms, even if - as some of our evidence indicates - these larger firms may actually offer more privacy. Third, we review a large body of qualitative and quantitative evidence that relates to the assumptions in our theoretical model and empirical analysis.

Our empirical results indicate that the 2009 ePrivacy Directive had on average no significant effect on the revenues of European e-commerce firms. However, it had significant heterogeneous effects among large and small firms. Only the revenues of large firms reduced, while those of small firms were essentially unaffected. This stands in strong contrast to other empirical studies of (even the same) privacy regulation that tend to emphasise negative effects on the industry (Goldfarb and Tucker, 2011; Jia et al., 2021; Lambrecht, 2017), and especially on small firms (Campbell et al., 2015).

Our results carry strong implications for the intersection of competition and data protection policy. They allow informing the ongoing debate regarding the most recent round of revision of the ePrivacy Directive in Europe. This revision is at the time of writing this article finally pushed out of a long stalemate stretching over several years in the European Council.[4] Discussions came to a halt mostly due to concerns raised by industry groups regarding a loss of competitiveness vis-à-vis online firms outside Europe (Ghosh, 2018; Singer, 2018; Gwynn, 2017; McConnell, 2019). Our empirical results looking at the last (2009) round of revision of the very same Directive suggest a more nuanced and optimistic view for businesses. While revenue losses cannot be excluded, these were small and confined to large firms only. Second, we emphasise the potential role of technology - in particular, firms' ability to monetise user data - to drive asymmetries in market shares and even impact which firms are affected most by privacy regulation.

The remainder of the paper is organised as follows. Section 2 introduces the relevant literature. Section 3 describes the institutional background. Section 4 introduces the data and the empirical model, and reports our main findings. Section 5 provides a theoretical model that rationalises our empirical results. Section 6 concludes.

2. Literature review

Our research is related to the rich and growing body of literature on the economics and marketing aspects of privacy, surveyed extensively by Acquisti et al. (2016). A closely related theoretical research line in this strand investigates how the ability of firms to recognise customers and send targeted offers affects market outcomes in an oligopolistic setting (Thisse and Vives, 1988; Kox et al., 2017; Campbell et al., 2015; Shy and Stenbacka, 2016; Baye and Sapi 2017).

A growing body of research looks at the effects of the GDPR in Europe on various market outcomes and firm behaviour. Goldberg et al., 2021 analyse website traffic data and e-commerce revenue for around a thousand websites before and after the GDPR implementation. The study finds an approximately 12% reduction in website visits as well as revenues after the enforcement of the GDPR in Europe. Peukert et al. (2022) look into the effects of privacy regulation in the vertical digital value chain. The authors report that websites more affected by the GDPR tend to reduce their reliance on third parties. Technology firms dominant in several market segments, such as Google, actually gained prominence.

Particularly close papers to ours are Campbell et al. (2015); Lambrecht, 2017; Goldfarb et al. (2011); Jia et al. (2021). Similar to our paper, these articles revolve around the effects of the ePrivacy Directive.

Lambrecht, 2017 provides an empirical impact assessment of the 2002 enactment of the ePrivacy Directive and looks at whether and how the Directive affected venture capital investment into start-ups operating in online advertising, online news, and cloud computing. Using similar investments in the U.S. as a benchmark and controlling for drivers of venture capital investment, the author finds that the passage of the 2002 ePrivacy Directive significantly dampened EU venture capital investments in the analysed sectors. As in Lambrecht, 2017, our empirical assessment takes U.S. and Canadian firms as the control group. However, instead of investment, we focus on revenues in the online retail sector. The results of Lambrecht, 2017 are consistent with the view that privacy regulation affects predominantly small firms and reduced expected revenues may be a reason why venture capital investments into online start-ups have been found to decrease.

We find the opposite: small European E-commerce firms were not significantly affected while larger firms were hit somewhat harder. This difference in result suggests that there may be sectorial heterogeneities in the impact of privacy regulation, as well as potentially different effects on existing firms and start-ups. A further difference of our empirical analysis is in our identification strategy. Lambrecht (2017) employs a difference-in-differences approach comparing EU and U.S. firms before and after the regulation, while we adopt a DDD model further distinguishing between comparable online and offline firms.

Goldfarb et al. (2011) use data on 3.3 million survey-takers randomly exposed to 9596 online display (banner) advertising campaigns to investigate the effect of the 2002 ePrivacy Directive (2002/58/EC) on the effectiveness of advertising campaigns. The authors report that the 2002 ePrivacy Directive significantly reduced the effectiveness of online banner ads by curbing the ability of advertisers to track users and offer targeted advertisements.

Jia et al. (2021) provide an early assessment of the European General Data Protection Regulation (GDPR) that came into effect in May 2018. The authors look at the effect of the GDPR on venture capital investment activity, and argue that the regulation reduced EU ventures, relative to their U.S. counterparts.

Our results point against the empirical results of Goldfarb et al. (2011); Lambrecht, 2017; Jia et al. (2021). While these papers attribute a negative effect to privacy regulation on businesses, we find on average no significant negative effect on revenues. There are however significant heterogeneous effects on large and small firms. Small firms' revenues were unaffected or may have even increased slightly, while those of large firms experienced either no or negative effects. We argue based on a simple yet new theory that under realistic conditions large firms may

[3] As an executive of Adform, a leading independent advertising technology company, put it: "*for the other [small] players, advertising revenues will diminish as cross-platform reach via tracking & measurement, essential for providing advertising success metrics, will slowly die. Only if you are big enough with respect to reach (and potentially still data), you will be able to attract advertising budgets. If you are a medium or small publisher, you are likely out of that game. As a result, the walled gardens will grow even stronger, they will increase their dominance of the Internet; even fewer players will own even more data*" (Schlosser 2017).

[4] Source through the following https://europa.eu/!qh98db.

be carrying the heaviest burden. Our theoretical model explicitly takes into account that data may allow firms to increase revenues by targeting offers at users.

3. The 2009 eprivacy directive

Our research is motivated by the long ongoing debate in Europe about the ePrivacy Regulation (Apostle, 2018; Khan, 2018; Singer, 2018). The ePrivacy Regulation builds on the former ePrivacy Directive of 2002 and intends to regulate how online businesses handle data and use cookies.[5] At the time of writing this article, the adoption of the ePrivacy Regulation is staggering mainly due to concerns about its implications on the performance of European online businesses (Ghosh, 2018; Singer, 2018; Gwynn, 2017; McConnell, 2019).

Our empirical assessment focuses on the predecessor of the proposed ePrivacy Regulation, namely the 2009 revision of the ePrivacy Directive in the European Union (Directive 2009/136/EC), also known as the *Cookie Law*. Following its 2002 enactment, the ePrivacy Directive was amended in 2006 (Directive 2006/24/EC) and a major change followed in 2009 (Directive 2009/136/EC). The 2009 revision constitutes the subject of our empirical analysis. We summarise the main implications of the 2009 ePrivacy Directive for online businesses.[6]

Online businesses only: The 2009 ePrivacy Directive applies to *electronic* communication services only.[7] Due to its provisions related to spam and cookies, it has strong implications for e-commerce. It does not, however, apply to offline businesses such as brick-and-mortar stores.

Informed opt-in to storing files (e.g. cookies) on the user's electronic device: Article 5.3 of Directive 2009/136/EC included a revolutionary novelty. It introduced the requirement in Europe for e-commerce firms to obtain explicit and informed user opt-in to the use of cookies and other storing of files on a user's electronic device.[8] Furthermore, the default settings of the browser, if set to automatically accept cookies, did not count as lawful consent, because such presets do not fulfill the *"clear and comprehensive information"* as by the ePrivacy Directive requirement.

Limiting spam and unsolicited marketing messages: Article 13 of the ePrivacy Directive regulates spam and unsolicited marketing and messaging. It rendered any form of electronic marketing communication illegal unless users gave prior consent.

Obligation to notify data breaches: Article 4.3 of the ePrivacy Directive obliges electronic communication service providers to notify breaches of personal data to national regulators, typically to national Data Protection Authorities. These authorities are also authorised to audit the compliance of service providers.

Enforcement and fines: Since EU directives are transposed into national law, the institutions (and timing) of enforcement may differ from country to country. In general, however, *"the national [...] data protection authority may impose fines or undertake other actions"* Papakonstantinou and de Hert (2011). Several authorities across Europe effectively issued fines for breaches of the national laws that implement the 2009 ePrivacy Directive.[9]

Table 1
Results from Numerical Simulation.

t	1	1	2	2
Δ	2	3	2	3
n_1 (*change*)	−0.029	−0.053	−0.04	−0.07
Π_1(%change)	−14	−22	−17	−24
Π_2(%change)	−6	−11	−7	−16
q_1(before)	0.575	0.493	0.842	0.689
q_2(before)	0.653	0.625	1.03	0.992
q_1(after)	0.556	0.468	0.811	0.654
q_2(after)	0.575	0.493	0.842	0.689
f_1(change)	−0.013	−0.025	−0.019	−0.033
f_2(change)	−0.003	−0.006	−0.006	−0.011

Private enforcement: A highly important novelty of the ePrivacy Directive was to substantially enlarge the circle of parties with a right to *sue* spammers and senders of unsolicited messages. For the first time, it authorised practically all parties - email service providers, e-commerce firms, their rivals, consumers, and trade and consumer associations - directly or indirectly involved in a typical spamming and email marketing activity to sue spammers.

Firms in electronic commerce and other online sectors voiced strong concerns about eroding the profitability of online advertising as a result of the Directive (Goldfarb and Tucker, 2011; Lambrecht, 2017). Our main question of interest is how the 2009 ePrivacy Directive affected firms and potentially market structure, by impacting large and small firms differently. We aim to empirically document any such differential effects by firm size (Section 4). We also emphasise differences in technology that may have acted as *one possible channel* to strengthen the differential impact of privacy regulation by firm size (Section 5).

4. Empirical analysis

This Section discusses our data and empirical strategy used to identify the causal effect of the introduction of the 2009 ePrivacy Directive on revenue in the online retail sector. We exploit variation in the timing of implementation of the ePrivacy Directive and construct a *difference-in-difference-in-differences* (DDD) model with variation in treatment timing. We provide evidence that the DDD model is valid in identifying the causal impact of the policy change through an event study design that shows parallel trends before the ePrivacy implementation. Finally, estimation results and robustness checks are presented.

4.1. Data

Our dataset consists of an unbalanced panel of firms active in the online and offline retail sector either in the U.S., Canada, or the EU for the period 2003–2017 from the S&P CapitalIQ database. This database contains detailed financial information on listed firms worldwide. For each firm, we observe revenues, total assets, current assets, operating status as well as the classification of the field of activity. Along with SIC codes by primary activity, the data provider offers its proprietary classification of business domains.

To implement our empirical strategy, we select firms active in the retail sector, either in North America (NA) or in the EU. We focus exclusively on firms selling consumer discretionary products such as clothes, electronics, and furniture. The sectors covered by the data are displayed in Table 2. This classification allows us to distinguish between firms selling similar products but differing in their respective reliance on online and brick-and-mortar distribution channels. We will refer to firms classified by our data provider

[5] Cookies are small files placed by visited websites on the users' computer that allows the website to track the user's activity and use this information for segmenting the audience and make targeted offers.

[6] Papakonstantinou and de Hert (2011), as well as Kosta (2013), provide excellent detailed reviews.

[7] Article 2(d) of Directive 2009/136/EC.

[8] In online markets where default choices are hard to resist (Lohr, 2011), the practical consequences of such a regime change can be large. Users browsing European websites will be familiar with a practical implication of the 2009 revision of the ePrivacy Directive: this regulation introduced the widespread use of pop-ups asking for consent to cookies that have been in place ever since.

[9] See Appendix B for the timing of transposition and Appendix C for enforcement actions.

Table 2
Retail Segments Covered by the Data.

CapitalIQ Classification	Freq.	Percent	Cum.
Apparel Retail	540	24.32	24.32
Computer and Electronics Retail	195	8.78	33.11
Home Furnishing Retail	285	12.84	45.95
Home Improvement Retail	180	8.11	54.05
Internet and Direct Marketing Retail	465	20.95	75
Specialty Stores	555	25	100
Total	2220	100	

Source: Selected data from S&P CapitalIQ database.

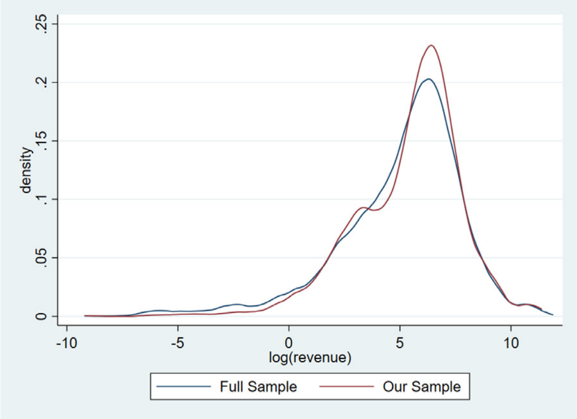


Fig. 1. Revenue Distribution across Samples This figure shows the distribution of the natural logarithm of revenues from the full sample, including 380 retail firms active in the EU and in NA, and from the restricted sample used in our analysis, which includes 148 firms with at least 75% of revenues realised in one single country.

as active in "*Internet and Direct Marketing Retail*" as "*Online*", while firms operating in other retail sectors will be defined as "*Offline*".[10] From this first selection by country and industry sector, we obtain financial information for 347 retail firms.

We then select firms that have been active before and after the implementation of the 2009 ePrivacy Directive. However, since the ePrivacy Directive applies only in the European Union, multinational firms active both in the EU and in NA pose a challenge to our data. As they typically report global financial figures, it is difficult to classify them as treatment (EU) or control (NA) regions. For these reasons, we restrict attention to the subset of firms that report financials separately by geographic segment and realise at least 75% of their revenues in one particular national geographical market. Our final dataset includes 148 firms active in the retail sector for which we observe annual total revenues. Summary statistics are displayed in Table 3.

Given that we focus on the narrow subset of listed businesses reporting financial results by region, the question of to what extent the restricted sample is representative of the population of retail firms requires attention. Fig. 1 displays the distribution of revenues for both the full and the restricted samples. We observe that the revenue distributions in the two samples are very close, lending some support to the view that our selected firms are representative of listed retail businesses active in NA and the EU in our financial database.

Our empirical strategy relies on firms operating either online or offline, either in North America or in the EU. Table 4 shows the distribution of firms in our sample across the four main categories necessary to implement the DDD analysis. We consider this distribution balanced across EU and NA as well as online and offline retail. Fig. 2 displays the trends of log revenue over these categories. We observe large heterogeneity in the dynamics of firm revenues. In particular, the figure shows a positive trend in NA compared to the EU (top-centre panel), which is mainly driven by the offline sector (bottom-right panel). When focusing on firms operating online (bottom-centre panel), we see a high fluctuation in the revenue of NA firms. This suggests that a simple comparison between NA and the EU would be misleading because of pre-existing differential trends unrelated to the ePrivacy Directive implementation. The DDD estimator is less prone to the same bias.

Finally, we investigate the implementation timing of the 2009 ePrivacy Directive across the different EU Member States. Table 5 shows when the ePrivacy Directive has been transposed into national law by the various EU Member States where the firms in our dataset operate.[11] Interestingly, EU Member States implemented the Directive between 2011 and 2013, years after its adoption by the EU. Finland, the United Kingdom, Sweden, and France were the first to convert the Directive into national law, while Poland and Norway were the last to do it.[12]

4.2. Empirical strategy

We first assess the impact of the ePrivacy Directive 2009/136/EC in a *difference-in-differences* (DiD) framework.[13] We aim to estimate the causal effect of the ePrivacy Directive on the revenues of firms operating in the online retail sector. The introduction of the ePrivacy Directive is regarded as an exogenous shock affecting European retail businesses, particularly those operating primarily online, as it influences the capability of online firms to acquire data on potential customers. This is expected to affect their capability to match consumer preferences by providing targeted offers.

Our empirical analysis relies on the ability to find a suitable control group, namely a set of comparable firms that have not been affected by the ePrivacy Directive. As the ePrivacy Directive applies only to EU businesses, one possible approach could be comparing retail firms operating primarily in the EU compared to NA retail firms. This would lead to a standard DiD approach commonly used in the privacy literature (Goldberg et al., 2021; Lambrecht, 2017, Peukert et al. 2022), but with variation in treatment timing (Goodman-Bacon, 2021), due to the differential implementation dates across EU Member States. Moreover, further interactions with the treatment indicator would also allow estimating differential effects for EU firms primarily active online. However, the main limitation is that the evolution of revenues in the retail sector might be systematically different in North America (NA) and the EU for reasons other than the policy change.

We test the validity of the DiD research design through an event study (Schmidheiny and Siegloch, 2019) that includes leads and lags from treatment timing, that is dummies identifying relative times from the introduction of the ePrivacy Directive in the EU

[10] Clearly, many brick and mortar businesses have also online shop surfaces. Inspecting the companies in our dataset, it appears that our classification corresponds well to the split of revenues these firms achieve online and offline. Online firms include names such as Amazon, Buch.de, Groupon, 1-800-flowers.com, NetonNet, Travel24.com, and Otto.de. Offline firms include among others Toys R Us and Foot Locker.

[11] The implementation dates were taken from publicly available national laws. A summary list of the transpositions of the ePrivacy Directive across Member States can be found in Appendix C.

[12] Although Norway is not part of the European Union, the country transposes EU Directives. It implemented both the 2002/58/EC Directive in July 2003 and its revision 2009/136/EC ePrivacy Directive, which is the object of our analysis.

[13] Early applications of this approach are found in Ashenfelter and Card, 1985; Card (1992) cand Card and Krueger (1993).

Table 3
Summary Statistics.

CapitalIQ Classification	Mean	s.d.	Min	Max	N
REVENUES					
Apparel Retail	1176.42	1824.43	2.69	12032.99	489
Computer and Electronics Retail	1622.75	2748.78	0.01	11640.73	176
Home Furnishing Retail	1061.61	1718.49	0.37	11547.56	262
Home Improvement Retail	9108.51	19594.97	0.39	88465.24	171
Internet and Direct Marketing Retail	997.02	2660.80	0.00	14766.78	402
Specialty Stores	625.41	1059.69	0.00	7408.64	512
Total	1698.60	6417.48	0.00	88465.24	2012
TOTAL ASSETS					
Apparel Retail	730.37	1116.49	1.18	7848.81	487
Computer and Electronics Retail	884.91	1444.40	0.33	6075.67	175
Home Furnishing Retail	651.22	1007.57	0.23	6449.46	260
Home Improvement Retail	5214.74	10857.30	1.64	40464.63	170
Internet and Direct Marketing Retail	998.90	2881.31	0.00	17264.65	397
Specialty Stores	379.06	544.71	0.01	3795.32	508
Total	1079.37	3736.31	0.00	40464.63	1997
CURRENT ASSETS					
Apparel Retail	348.11	506.65	1.06	3840.77	487
Computer and Electronics Retail	387.64	704.01	0.22	3407.79	175
Home Furnishing Retail	327.97	608.85	0.23	3641.96	260
Home Improvement Retail	1848.47	3791.48	0.22	16575.48	170
Internet and Direct Marketing Retail	564.41	1748.07	0.00	12771.86	397
Specialty Stores	200.18	289.33	0.01	1866.36	508
Total	482.04	1478.96	0.00	16575.48	1997

Source: Selected data from S&P CapitalIQ database. Data are in million U.S. dollar.

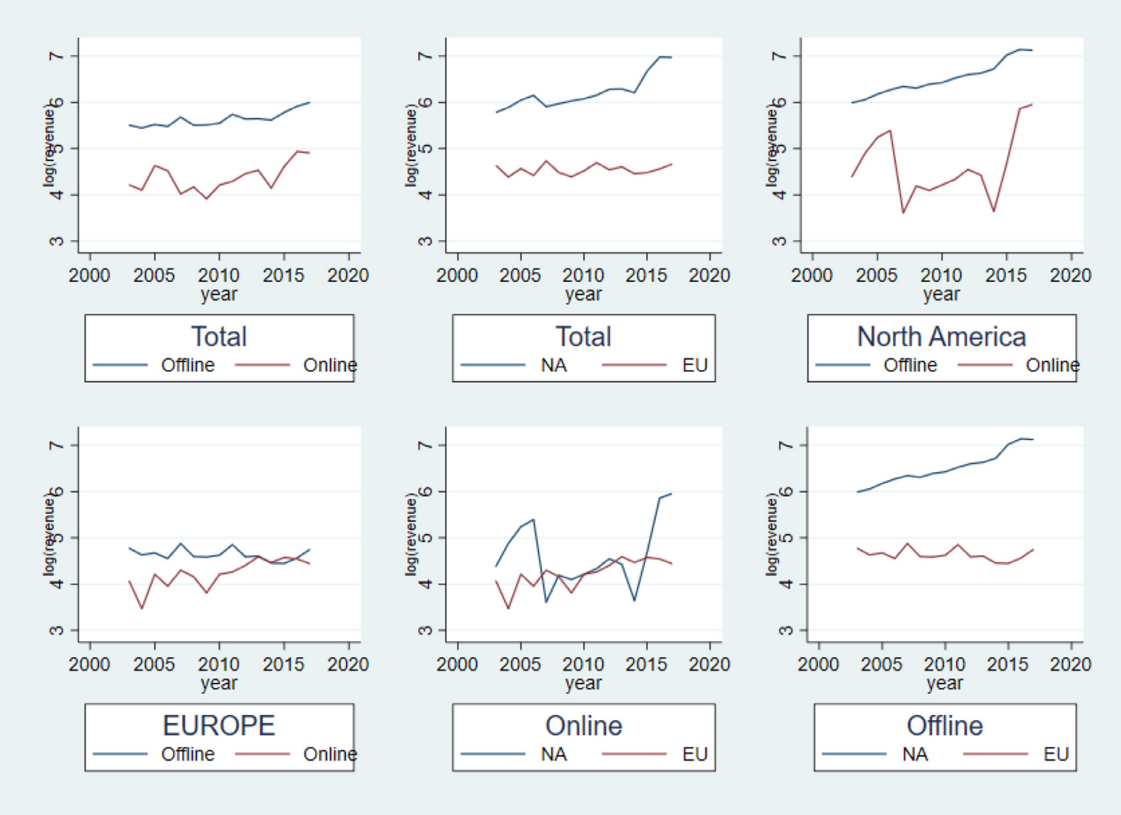


Fig. 2. Trends Over Categories This figure shows average yearly values for log-revenues across the main categories defining our DDD model. NA refers to North America, while EU indicates European Union. Online revenues come from firm operating in "Internet and Direct Marketing Retail." Source: Selected data from S&P CapitalIQ database.

Member States where the firms in our sample operate. The event study equation takes the following form:

$$y_{i,t} = \delta_0 + \sum_{l=-6}^{+4} \gamma_l \mathbb{I}\{t - ePriv_i = l\} + X'_{i,t}\delta_1 + \alpha_i + \tau_t + u_{i,t} \qquad (1)$$

where $y_{i,t}$ is the natural logarithm of revenue for firm i at time t. $ePriv_i$ is the time when ePrivacy Directive starts to be effective for firm i, and $l = t - ePriv_i$ are the relative times from the enactment of the ePrivacy Directive. $X'_{i,t}$ is a vector of potential controls, whereas α_i and τ_t are firm and year fixed effects, respectively. The parameters γ_l measure the percentage variations of our dependent

	Online	Offline	Total
EU	19	57	76
US & Canada	12	60	71
Total	31	117	148

Number of firms distributed among DDD groups.

Table 5
Post Dummy Identification.

Country	Date of Implementation	*Post* = 1
Cyprus	18th May 2012	2012
Denmark	14th December 2011	2012
Finland	25th May 2011	2011
France	24th August 2011	2011
Germany	10th May 2012	2012
Greece	10th April 2012	2012
Ireland	1st July 2011	2011
Italy	30th May 2012	2012
Norway	1st July 2013	2013
Poland	22nd March 2013	2013
Romania	26th July 2012	2012
Slovenia	15th January 2013	2013
Spain	2nd April 2012	2012
Sweden	1st July 2011	2011
UK	26th May 2011	2011

The table shows when Member States in our sample data implemented the ePrivacy Directive 2009/136/EC. The variable *Post* takes value one from the related year onward. The detailed list of National laws can be found in Appendix C.

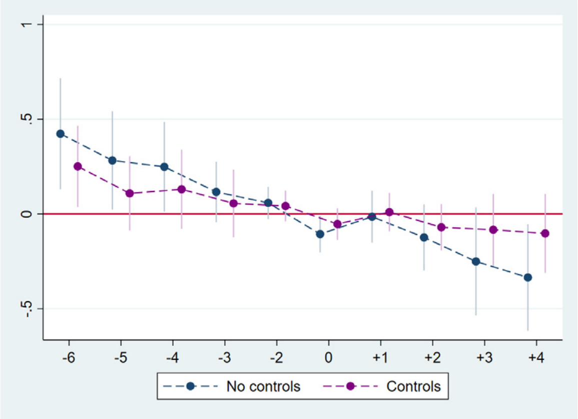


Fig. 3. Event Study Estimates of the DiD Model Presented are estimated event study coefficients from Eq. (1) and the associated 95% confidence interval. The dependent variable is the natural logarithm of firm revenues. Standard errors clustered by country-industry.

variable in the relative period l. We bin coefficients at $l = -6$ and $l = +4$, so that γ_{-6} and γ_{+4} capture the mean effect before and after the relative time window $\{-6, -5, \ldots, 0, +1, \ldots, +4\}$. Because of collinearity we need to drop $l = -1$.

The DiD research design identifies the causal impact of the ePrivacy Directive on firm revenues if firms active in NA provide a valid counterfactual for EU firms. This implies that, before the transposition of the ePrivacy Directive, we should observe no systematic variation between EU and NA revenues. Hence, we would expect flat trends before firm i is subject to the ePrivacy Directive. That is, γ_l should not be statistically different from zero for $l < 0$.

Fig. 3 reports estimated coefficients and the associated 95% confidence interval of Eq. (1). Blue dots refer to a simple specification without controls – apart from firm and time fixed effects, while purple dots refer to a specification that adds the natural logarithm of firm current and total assets. Both with and without controls, we observe pre-trends in our variable of interest, as the coefficient γ_{-6} is never statistically different from zero. Adding controls attenuates the issue, although without solving it. Post-treatment coefficients are negative and statistically significant in the *naive* specification, but they approach zero once we add further controls, suggesting a null mean effect. However, since the parallel trend assumption is not satisfied, the event study estimation suggests that a simple DiD comparison between EU and NA firm revenues is not feasible, consistently with the heterogeneous behaviour of firm revenues observed in Fig. 2.

Our data allow us to estimate a *difference-in-difference-in-differences* (DDD) model. In this setting, we use retail firms - online and offline - operating in NA, along with EU offline retail firms, as controls. Online firms operating in the EU remain in the treatment group. This approach allows controlling for potential confounding factors, including changes in revenues due to idiosyncratic, time-varying differences between the NA and the EU, and other factors affecting online firms other than the policy change.

Let $y_{i,t}$ be the natural logarithm of revenue for firm i at time t. Our baseline empirical specification is of the form:

$$y_{i,t} = \beta_0 + \beta_1 Post_{i,t} \times Online_i + \\ + \sum_t \beta_{2,t} Time_t \times EU_i + \sum_t \beta_{3,t} Time_t \times Online_i + X'_{i,t}\beta_4 + \alpha_i + \tau_t \\ + \epsilon_{i,t} \qquad (2)$$

where EU_i is a dummy taking value one if firm i operates in the EU, while $Online_i$ is an indicator taking value 1 if firm i operates online; α_i and τ_t are firm-specific and time fixed effects respectively. The variables $Time_t \times EU_i$ and $Time_t \times Online_i$ are derived by the interaction of time-specific dummies with EU_i and $Online_i$, respectively. The coefficients $\beta_{2,t}$ capture the percentage difference of the dependent variable between firms operating in the EU and NA at time t. Coefficient $\beta_{3,t}$ captures the percentage difference of the dependent variable for online firms versus offline firms at time t. The variable $Post_{i,t}$ is a dummy that identifies the period covered by the policy change for firm i, depending on whether the Member State where i operates has implemented the ePrivacy Directive.[14] Finally the vector $X_{i,t}$ collects time-varying controls for each firm in our sample. Specifically, it includes the natural logarithm of a firm's total and current assets, which are strongly positively correlated with firm revenue.

The coefficient of interest in expression (2) is β_1, capturing the average causal effect of the policy change on the dependent variable. In particular, it represents the average causal effect of the introduction of the ePrivacy Directive on online firms in the EU.[15]

Compared to a simple DiD approach, the DDD model in Eq. (2) adds a new set of firms acting as controls, i.e. EU offline firms, providing a better counterfactual for the estimation. We test whether this is the case by looking for parallel trends in a new event study design that generalises Eq. (2). Therefore, we estimate the following equation:

$$y_{i,t} = \beta_0 + \sum_{l=-6}^{+4} \mu_l \mathbb{I}\{t - ePriv_i = l\} \times Online_i + \\ + \sum_t \beta_{1,t} Time_t \times EU_i + \sum_t \beta_{2,t} Time_t \times Online_i + X'_{i,t}\beta_3 + \alpha_i + \tau_t \\ + \epsilon_{i,t}, \qquad (3)$$

as before, we drop $l = -1$, and we want $\mu_l = 0$ for $l < 0$.

[14] As shown in Table 5 in next Section, the ePrivacy Directive 2009/136/EC has been implemented between 2011 and 2013.

[15] As shown by Goodman-Bacon (2021), the resulting estimate is a weighted average of all the simple two-period treatment effects in each DDD, where weights depend on treatment variances and group sizes.

Table 6
Results on Revenues.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post×Online	0.421**	0.049	0.085	-0.162	0.170	0.478**	0.098	0.121
	(0.173)	(0.126)	(0.111)	(0.101)	(0.145)	(0.210)	(0.127)	(0.117)
(Post×Online)$\mathbb{I}_{>p50}$						-0.230	-0.193***	-0.147**
						(0.200)	(0.055)	(0.073)
Current Assets		0.697***	0.431***	0.361*	0.487*		0.696***	0.432***
		(0.061)	(0.153)	(0.185)	(0.254)		(0.061)	(0.154)
Total Assets			0.318*	0.302	0.294			0.316*
			(0.164)	(0.194)	(0.260)			(0.164)
Data	All	All	All	Above Median	Below Median	All	All	All
R-squared	0.944	0.967	0.968	0.954	0.934	0.944	0.967	0.968
Observations	2012	1997	1997	1023	974	2012	1997	1997

Presented are OLS estimated coefficients from Eq. (2). All specifications include time and firm-specific fixed effects. Clustered standard errors at industry-country level are reported in parenthesis below coefficient. *** p<0.01, ** p<0.05, * p<0.1

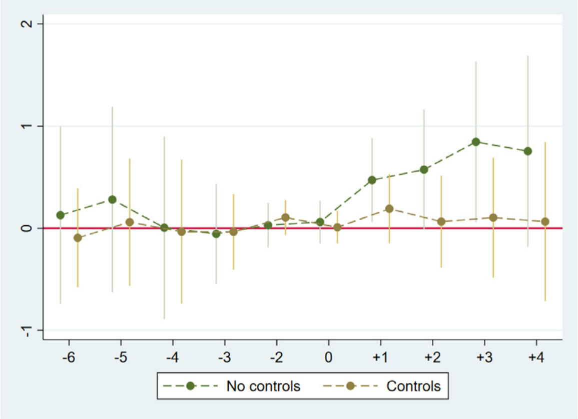


Fig. 4. Event Study Estimates of the DDD Model Presented are estimated event study coefficients from Eq. (3) and the associated 95% confidence interval. The dependent variable is the natural logarithm of firm revenues. Standard errors clustered by country-industry.

Fig. 4 presents the event study coefficients of Eq. (3) and the associated 95% confidence interval with and without firm-level controls. Both specifications show flat trends before the introduction of the ePrivacy Directive, supporting the validity of the DDD design. Thus, Eq. (2) correctly estimates the impact of the ePrivacy Directive on firm revenue. Furthermore, the inclusion of firm-level controls nullifies the positive impact suggested by the "*naive*" specification, implying a mean null effect of the ePrivacy Directive on the revenues of EU online firms.

Although event study estimates suggest that the ePrivacy Directive had a negligible effect on firm revenues overall, potential heterogeneous effects might arise across firms. We investigate this issue in two ways. First, we split the dataset, and we estimate Eq. (2) only on large firms and only on small firms[16] Second, we include additional covariates derived from the interaction of our main variable of interest $Post_{i,t} \times Online_i$ with an indicator $\mathbb{I}_{>p50}$ that identifies large firms. The resulting covariate measures the average deviation in the post-treatment period for firms identified by $\mathbb{I}_{>p50}$ from the average treatment effect on firms for which $\mathbb{I}_{>p50}$ equals zero (i.e., small firms), measured by β_1.

4.3. Results

Table 6 reports results from the ordinary least squares (OLS) estimation of the econometric model (2), showing the impact of the ePrivacy Directive on revenues. In particular, columns 1–3 display results from the estimation over the whole dataset, identifying the mean effect of the ePrivacy Directive on European online retail businesses. In columns 4–8 we attempt to identify the heterogeneous effect on the treated in different ways: columns 4 and 5 show the results when we run the econometric model only on large and small firms, respectively. In columns 6–8 we include interactions of our variable of interest with an indicator function identifying large firms.[17]

The coefficient of interest, *Post* × *Online*, captures the average causal effect of the introduction of the ePrivacy Directive on the dependent variable in percentage terms. When we estimate the model over the full sample of firms in our dataset without control variables (column 1), we find a sizeable positive effect induced by the ePrivacy Directive on firm revenues. However, once we include total and current assets as further controls (columns 2 and 3), the estimated coefficient is much lower in magnitude and no more statistically significant.

We next run the model on large firms (column 4) and small firms (column 5) only. The estimated coefficients change in sign when moving from large to small firms, highlighting potential heterogeneous effects based on firm size. Standard errors are relatively high, which is not surprising given that we have reduced significantly the number of observations by splitting the sample.

In columns 6–8 we add an additional interaction of the variable of interest with an indicator function $\mathbb{I}_{>p50}$ that identifies large firms as those with total assets above the sample median. In this setting, the coefficient associated to *Post* × *Online* captures the causal effect of the ePrivacy Directive on firms for which $\mathbb{I}_{>p50}$ is equal to zero, i.e. *small* firms. On the other hand, the coefficient associated to $(Post \times Online)\mathbb{I}_{>p50}$ measures the additional variation for *large* firms in the post-treatment period from *Post* × *Online*.

Once we control for firm assets, we find a negative and statistically significant coefficient on the interacted term, while *Post* × *Online* is positive but never statistically different than zero. The estimated coefficient in column (8) for $(Post \times Online)\mathbb{I}_{>p50}$ implies an average deviation of large firms from the rest of the sample of −14.7%, implying a negative causal impact of the ePrivacy Directive on large firms of about 2.6%. Hence, our results suggest that small retail firms are not significantly affected by privacy policy restrictions and that only larger firms may be weakly negatively affected.

In summary, we find that the implementation of the 2009 revised ePrivacy Directive had on average little impact on revenues in the retail sector. However, when we analyse the heterogeneous

[16] We exploit the distribution of total assets. We identify small (large) firms, as those with total assets below (above) the sample median.

[17] As before, large (small) firms are defined by those with total assets above (below) the sample median. It is useful to recall that all firms in our sample are listed. Small firms, therefore, are small compared to large listed firms in the sample and are likely relatively large compared to non-listed firms.

Table 7
Robustness Checks.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Post×Online	0.489*	0.065	0.112	0.207	0.108	0.074
	(0.261)	(0.143)	(0.137)	(0.200)	(0.168)	(0.144)
(Post×Online)$\mathbb{I}_{>p50}$	-0.087	-0.220**	-0.186**	-0.224	-0.270**	-0.218**
	(0.254)	(0.085)	(0.092)	(0.193)	(0.103)	(0.083)
Current Assets		0.700***	0.426**		0.474***	0.220
		(0.066)	(0.166)		(0.087)	(0.212)
Total Assets			0.330*			0.359
			(0.179)			(0.280)
Data	Germany Excluded			Drop U.S. fluctuation		
R-squared	0.945	0.968	0.969	0.981	0.985	0.986
Observations	1895	1880	1880	861	858	858

Presented are OLS estimated coefficients from Eq. (2). All specifications include time and firm-specific fixed effects. Clustered standard errors at industry-country level are reported in parenthesis below coefficient. *** p<0.01, ** p<0.05, * p<0.1

Table C1
Share of email, display ads and direct traffic on total visits by website size category, 2017.

Size category	Email	Display ads	Direct	N(websites)
small	6.77%	2.15%	36.70%	183
medium	6.33%	4.58%	40.39%	74
large	4.54%	1.82%	47.30%	70
giant	4.25%	0.96%	44.48%	19

Source: SimilarWeb, data for 346 domains from France, Germany, Italy, Netherlands, United Kingdom and United States, in website categories *"E-commerce related"* and *"E-commerce and shopping"*.

effects of such a policy change, we find that the null mean effect is due to qualitative differences in the treatment effect between large and small firms. Small firms are not significantly affected by the introduction of the 2009 ePrivacy Directive. On the contrary, large firms may suffer from more stringent privacy regulations.

4.4. Robustness checks

We run a battery of robustness checks, all of which confirm our main results. The first check relates to a possible ambiguous application of the ePrivacy Directive in Germany. We run this test because Germany is often claimed to have partly avoided the implementation of the 2009 ePrivacy Directive.[18] Our results are not affected by German firms. Table 7 columns 1–3 collect estimated coefficients of Eq. (2) when we drop Germany. We do not observe any significant variation compared to our main results.

Another potential concern may arise from the fluctuation in online NA firm revenues observed in Fig. 2. Although the event study estimates in Fig. 4 suggest that we are carefully controlling for differential trends between EU and NA firms, we estimate (2) focusing only on a the *stable* period 2008–2013. Our results remain intact, and somewhat increasing in magnitude.

To further validate our empirical model, we also run a placebo test in which we randomise the Online identifier. We randomly shuffle the Online assignment 1000 times and we estimate our DDD model including also the covariate $(Post \times Online)\mathbb{I}_{>p50}$. If the implementation of the ePrivacy Directive is really the driver of our main results, and brick-and-mortar retailers are not significantly influenced by the Directive, then we should find a zero effect arising from such a "fake" assignment. Fig. 5 shows the distribution of $(Post \times Online)$ and $(Post \times Online)\mathbb{I}_{>p50}$ coefficients when the dependent variable is log-revenue. Estimated coefficients in Table 6 column (8) lie in the tails of the distribution, implying that we can reject the null to get the same results by using such "fake" assignments. What is more, the mean of $(Post \times Online)\mathbb{I}_{>p50}$ lies slightly above zero, in contrast with the negative effect identified by our DDD model. Thus, a random assignment of the Online identifier yields statistically different results compared to the DDD estimates, implying a causal interpretation of the ePrivacy Directive on the revenues of European online firms.

In conclusion, the robustness checks validate the goodness of our DDD model. The placebo tests confirm that we cannot obtain the same results of Table 6 when we randomise the online assignment, suggesting that the ePrivacy Directive has affected only European online firms. Table 7 shows that our main results are not sensitive to different cuts of the data, confirming the negative impact of the ePrivacy Directive on large firms primarily operating online.

5. A theoretical model of privacy in E-commerce

Why may larger firms be affected differently by privacy regulation than smaller firms? In this section, we provide a simple yet novel theoretical model that predicts, as we find empirically, that larger firms may lose more revenues due to the privacy regulation (both in absolute and percentage terms) than smaller firms.[19]

Our theoretical framework is motivated by e-commerce environments in which firms of the treatment group of our empirical analysis operate. These firms are typically retail platforms operating e-commerce websites selling a vast array of third-party products offered by various brands. In particular, we focus on a market consisting of two competing multi-product online retailers $i = \{1, 2\}$. The retailers sell the products of several brands on their websites and finance themselves from slotting fees these brands pay in exchange for listing their products. Retailers provide services at zero marginal cost and realise profits

$$\Pi_i = a_i f_i, \tag{4}$$

where a_i is the number of brands choosing to be listed at the retailer and f_i is the uniform slotting fee of Retailer i.

Consumers regard retailers as differentiated. Retailers can be thought of as being located at the endpoints of a line of unit length along which consumers are uniformly distributed with unit mass. For convenience, we assume that Retailer 1 is located at

[18] See Appendix C for more on this issue.

[19] We do not suggest that the differential effects by firms size we observed in Section 4 are necessarily driven by the theory we propose in this section. We aim to raise awareness about technology factors being one possible explanation for the documented differential effect across firms of different sizes.

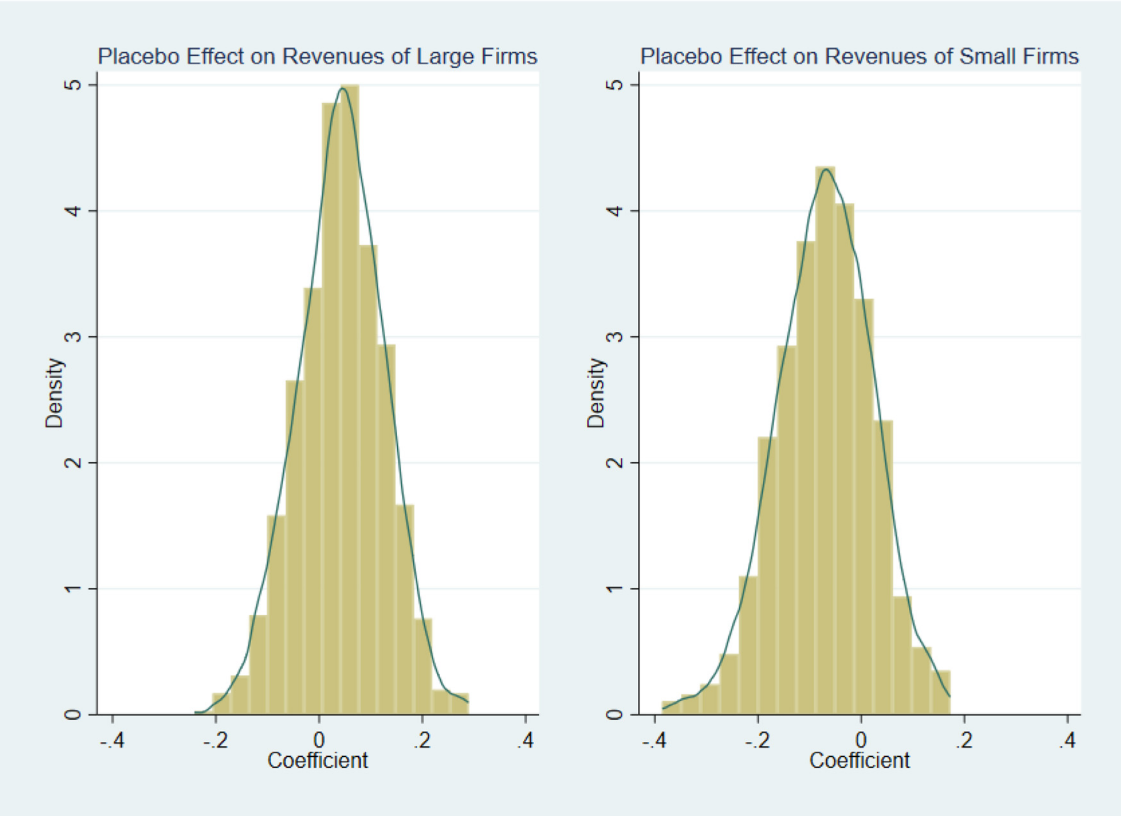


Fig. 5. Placebo Test Distribution of *Post* × *Online* (right) and $(Post \times Online)\mathbb{I}_{>p50}$ (left) coefficients deriving from randomization of the Online dummy. The dependent variable is the natural logarithm of revenues. Number of permutations: 1000.

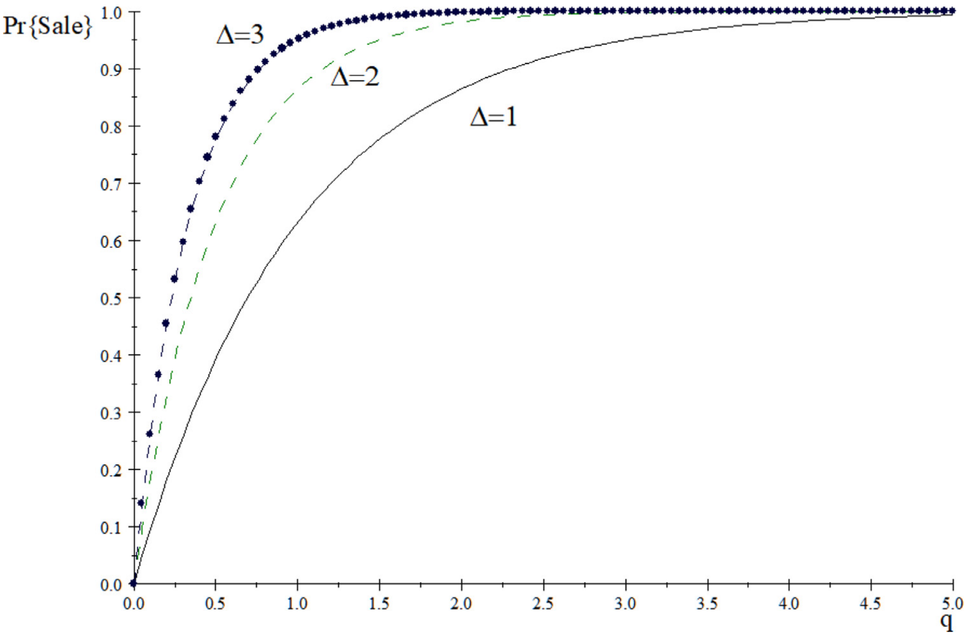


Fig. A1. Illustration of the function $\Pr\{Sale_1\} = 1 - E^{-\Delta q_1}$, with $\Delta = 1, 2, 3$.

endpoint 0 and Retailer 2 at endpoint 1 of the unit line. Consumers are characterised by an address on the line so that their distance to the endpoints represents their preference for each retailer.

When purchasing from a retailer, the consumer incurs a disutility that increases linearly in proportion to the distance to the retailer. Retailers 1 and 2 are free to use for consumers but they collect data on their users. These data in turn enable brands carried by the retailer to better target products to users. In particular, retailers choose their privacy policy $q_i \geq 0$, where a larger value represents more intense use of data and consequently less user privacy: $q_i \geq 0$ can be seen as the retailer's (or its website's) *privacy-intrusiveness*. Consumers value privacy and are informed about the retailers' use of data and choice of q_i. Privacy in this setup is a retailer-specific feature (Farrell, 2012). It affects demand, as more privacy implies a higher willingness to pay for the retailer's product.

When choosing between the retailers, consumers single home. A consumer at location x faces the choice between realizing the following utilities at Retailers 1 or 2:

$$\begin{aligned} U_1 &= V - tx - q_1 \\ U_2 &= V - t(1-x) - q_2, \end{aligned} \tag{5}$$

where V is a basic utility from visiting a retailer and t is a transportation cost parameter per unit distance in the preference space. We assume that V is high enough so that in equilibrium every consumer visits one of the retailers.

The retailers operate websites that provide information about products of different brands. For the brands the retailer's website is a marketing channel to consumers, allowing them to target products to individual users based on the data the retailer's website collects.[20] Brands decide on whether to advertise at Retailers 1 and 2 and face no capacity constraint. In particular, they may decide to list their products on either retailer's website, or both websites or refrain from listing at the retailers. If brand j lists its product on Retailer i's website, the brand incurs two types of costs. First, the slotting fee f_i that is uniform to all brands at Retailer i. Second, a retailer-specific cost c_{ij}. The latter captures the brand's fixed cost associated with listing its product on Retailer i's website other than the slotting fee, such as costs to comply with the technical requirements of the retailer and designing a digital advertisement for the product sold with the retailer. We assume that c_{1j} and c_{2j} are uniformly distributed on the interval $c_{ij} \in [0, \infty)$. This means that brand j expects the following profit from offering its product priced at $\overline{p}$ on the website of Retailer i:

$$\pi_{ji} = \Pr\{Sale_i\} n_i \overline{f_j} - f_i - c_{ij}, \tag{6}$$

where n_i and $\overline{f_j}$ respectively denote Retailer i's share among users and the average price of brand j's product. $\Pr\{Sale_i\}$ is the probability of successfully selling the product to a consumer through Retailer i. In particular, we assume that this probability depends on q_i, the privacy policy of that retailer, so that $\Pr\{Sale_i\} = s_i(q_i)$, with $\partial s_i(q_i)/\partial q_i > 0$ and $\partial^2 s_i(q_i)/\partial q_i^2 < 0$. The more data the retailer collects through its website, the higher the probability that the brand realises a sale via the retailer, but data shows decreasing returns.

Given that brands face no capacity constraint, they will decide to be listed at each retailer as long as doing so entails positive expected profits, which is the case when $\Pr\{Sale_i\} n_i \overline{f_j} - f_i \geq c_{ij}$. There will be a marginal brand at each retailer with fixed cost $\overline{c_i}$ for which this relationship holds with equality so that the brand is indifferent between being present at the retailer or not. The number of brands listed on the platform equals the fixed costs of the marginal brand, with $a_i = \overline{c_i}$. The demand function of brands for retail space is therefore given by

$$a_i = \Pr\{Sale_i\} n_i \overline{f_j} - f_i. \tag{7}$$

To economise on parameters we without loss of generality normalise $\overline{f_j}$ to 1. To obtain closed form solutions we assume an explicit functional form for the probability of successful sale. In particular, let

$$\Pr\{Sale_i\} = s_i(q_i) = 1 - e^{-\Delta_i q_i}, \tag{8}$$

with $\Delta_2 = 1$ and $\Delta_1 = \Delta > 1$. Parameter Δ represents Retailer 1's *technology advantage* in enabling brands to convert user data into sales.[21] In the following we will sometimes refer to Retailer 1 as the firm with *superior data technology*.

Superior data technology can stem from a range of factors, including a better ability to obtain, collect, store and analyse customer data, or access to a better algorithm that recommends products the customer may be interested in. The relevance of these factors is well documented in e-commerce (Akter et al., 2016).

This theoretical setup is simple and tractable, and captures the main trade-off in electronic commerce regarding privacy: retailers value data because it increases revenues. Users however value privacy and prefer to reveal less data. To focus on the essentials, we consciously abstract away from possible feedback loop effects arising from consumers anticipating how revealing their data may affect prices and offers at the retailers, and form no expectations about the prices they expect to see at the websites. Consumers are therefore *uninformed but rational* (Gomes and Tirole, 2018): they are conscious of websites' privacy offering but do not know what product content and prices to expect before they visit the websites. This corresponds to mild consumer myopia, which we consider both practical and realistic in e-commerce.[22]

The sequence of decisions is as follows. Retailers simultaneously and independently decide on the privacy-intrusiveness of their websites, q_i. They subsequently simultaneously and independently decide on the uniform slotting fee f_i. Brands choose whether to list their product on a retailer's website and consumers choose which retailer's website to visit.

5.1. Equilibrium analysis

Consumers single home when they decide which retailer's website to visit and surf to the website of the retailer offering higher utility. We can find the address $\bar{x}$ of the marginal consumer that is indifferent between the retailers. Under our assumptions this address directly determines the market share of the retailers among users, so that $\bar{x} = n_1$ and $1 - \bar{x} = n_2$. The demand for each retailer is

$$\begin{aligned} n_1 &= (q_2 - q_1 + t)/2t, \\ n_2 &= (q_1 - q_2 + t)/2t. \end{aligned} \tag{9}$$

Plugging these values together with Expressions (7) and (8) into Expression (4) allows us to obtain the profit retailers seek to maximise by setting slotting fees:

$$\begin{aligned} \Pi_1 &= f_1\left[(1 - e^{-\Delta q_1})\frac{q_2 - q_1 + t}{2t} - f_1\right], \\ \Pi_2 &= f_2\left[(1 - e^{-q_2})\left(1 - \frac{q_1 - q_2 + t}{2t}\right) - f_2\right]. \end{aligned} \tag{10}$$

Maximizing with respect to the slotting fees yields

$$\begin{aligned} f_1^*(q_1, q_2) &= \frac{\left(1 - e^{-\Delta q_1}\right)(q_2 - q_1 + t)}{4t}, \\ f_2^*(q_2, q_1) &= \frac{(1 - e^{-q_2})(q_1 - q_2 + t)}{4t}. \end{aligned} \tag{11}$$

Plugging these back into Expression (10) results in the following reaction functions:

$$q_1(q_2) = q_2 + t + \frac{1 - W\left(e^{\Delta(q_2+t)+1}\right)}{\Delta}, \tag{12}$$

[20] Nothing in this model would change if retailers tailored the targeted offers using the customer data.

[21] We follow Reinganum (1983) and a large body of subsequent R&D literature with this functional form assumption. Appendix A contains a graphical illustration of the function $s_i(q_i)$.

[22] A survey by Episerver, 2018 found that only 17 percent of people say that making a purchase is their primary purpose for visiting a brand's website for the first time. The primary purpose of visiting an e-commerce website is in the vast majority of cases is not directly related to purchase intent, but involves looking for information on store openings, shipping, or payment. It, therefore, seems unlikely that consumers would strategically refrain from website visits anticipating that doing so may affect prices.

$$q_2(q_1) = q_1 + t + 1 - \frac{W\left(e^{q_1+t+1}\right)}{\Delta},$$

where $W(.)$ is the *Lambert W* function that satisfies $W(ze^z) = f^{-1}(ze^z) = z$.[23] Notably, this function is positive and concave over the domain of real numbers. Using this property, we can take the partial derivatives of the reaction functions with respect to the rival retailer's privacy intrusiveness to establish that reaction functions are upward sloping and hence privacy decisions are strategic complements:

$$\frac{\partial q_1(q_2)}{\partial q_2} = \left[W\left(e^{\Delta(q_2+t)+1}\right)\right]^{-1} > 0,$$
$$\frac{\partial q_2(q_1)}{\partial q_1} = \left[W\left(e^{q_1+t+1}\right)\right]^{-1} > 0.$$

Having set up the basic model, the following proposition describes the equilibrium absent privacy regulation.

Proposition 1. *In equilibrium the retailer with superior data technology (Retailer 1) has higher market share among consumers ($n_1^* > n_2^*$), adopts a less intrusive privacy policy ($q_1^* < q_2^*$), offers brands higher probability of sale ($s_1^* > s_2^*$), has higher slotting fees ($p_1^* > p_2^*$), offers more products ($a_1^* > a_2^*$) and realises higher profits ($\Pi_1^* > \Pi_2^*$) than the rival.*

Proof. See Appendix A. □

The main result is that the firm with superior data technology is larger than the rival, yet it offers a higher level of privacy.[24] This is an important insight that goes against the prevailing intuition in competition policy, where market power is traditionally regarded as a precondition for the ability and incentive of firms to exploit users for their data. In our case, the contrary holds: the higher level of privacy is the precise reason why Retailer 1 is larger than the rival. Since it needs to obtain less data on consumers due to its superior technology to turn those data into increased sales, Retailer 1 can outcompete Retailer 2 in the privacy policy and provide services in a less intrusive manner.[25]

5.2. Regulating eprivacy

A regulation of ePrivacy has the aim of increasing the privacy level of online services. We can think of it analytically as a cap on q_i, the *data intrusiveness* of web retailers. Can such regulation affect competition between large and small firms, so that they are affected differently? We extend our theoretical model to address this question. The following proposition sums up our main insight.

Proposition 2. *Privacy regulation that caps the privacy intrusiveness of e-commerce retailers may reduce the profits and revenues of larger firms more than those of smaller firms, both in absolute and percentage terms.*

Proof. By numerical example provided in the Appendix and the text below. □

This proposition is modest, as we do not allege privacy regulation always affects larger firms more. Our aim is merely to demonstrate that such an outcome is possible in theory, and would be aligned with our empirical results. Yet, the claim that small firms may not be the main victims of such regulation is novel both in the literature as well as in the public discourse. A numerical example is sufficient to prove the proposition. To do so, we use the model above to numerically calculate the equilibrium without regulation. For simplicity, we then assume that the regulation imposes a cap on privacy, which equals the privacy level offered by the less intrusive firm (Retailer 1) in the absence of privacy regulation. We then assess how key model variables change when the regulation is introduced. The main variable of interest for the following empirical analysis is the percentage change in profits at each firm. The results of the numerical simulation are in Table 1.

The main result emerging from this numerical simulation is as follows. If privacy regulation caps privacy intrusiveness at the pre-regulation level ($\overline{q}$ of the less intrusive firm), then this regulation may not be binding for the retailer with superior data technology (Retailer 1) and only binds for the rival $q_2^* = \bar{q}$. In this case, the retailer with superior data technology (Retailer 1) loses market share to the rival (n_1^* decreases), its profits are decreased in percentage terms more than those of the rival ($|\Pi_1\%| > |\Pi_2\%|$), the probability of sale, slotting fees and number of products sold decrease at both retailers.

We assumed for the numerical simulations that the regulation caps the privacy intrusiveness of Retailer 2 at the pre-regulation level of the retailer with superior data technology (Retailer 1). Since privacy choices are strategic complements, the reduction of the privacy intrusiveness of Retailer 2 by the regulation also induces Retailer 1 to offer more privacy to consumers. This also means that such a regulation is not binding for Retailer 1 and only binds for Retailer 2.

We summarise the main insights emerging from this simple setup. The main result of the proposed model is that, surprisingly, the profits of the larger Retailer 1 may be hit harder by the regulation than those of the smaller rival (Retailer 2), even if the former offered more privacy. The retailer with superior data technology may experience a higher percentage reduction in profits than the rival.

The reason is that the binding privacy regulation at Retailer 2 makes the latter attractive for consumers who value privacy. Retailer 2, therefore, gains market share from Retailer 1. Privacy decisions being strategic complements results in both firms offering more privacy following the regulation, even if the regulation is binding for one firm only. However, since Retailer 1 is more productive in converting consumers into sales due to its superior data technology, losing these consumers implies a relatively high-profit reduction for Retailer 1 that exceed the profit gains of Retailer 2 even in percentage terms.

Our model is based on the idea that larger firms are better able to turn data into increased sales. This may be so due to technological reasons, such as economies of scale in data, or better analytical capabilities. If larger firms are more productive in data use, stripping them from their ability to gather customer data affects their revenues stronger (negatively) than those of smaller rivals. Our model also predicts that smaller firms offer on average less privacy than larger ones, a conjecture for which we can provide stylised evidence in the Appendix. Even if - as *sceptics* argue - the privacy regulation is not directly binding for large firms, competition forces these businesses to reduce their privacy intrusiveness in response to their smaller rivals doing so.

In terms of theory, our finding that privacy regulation may negatively affect predominantly large firms - even even though these larger firms may pre-regulation actually provide more privacy - is novel. A useful comparison is with the model of Dimakopoulos and Sudaric (2018), looking at the privacy decisions of competing two-sided platforms. While that model is symmetric and hence allows no immediate comparison of large and small firms, it predicts that firms competing more intensely for either side of the market (in terms of lower transport costs) offer better pri-

[23] See Corless et al. (1996) for a detailed analysis of the *Lambert W* function.

[24] We provide stylised evidence for larger websites being more privacy-friendly than smaller ones in the Appendix. This includes privacy ratings for thousands of websites.

[25] The Appendix includes further discussion on some detailed assumptions of the model, and how those compare to related work.

vacy. If smaller firms in that model were those facing less competition, privacy regulation may reduce the profits of these smaller firms stronger than those of larger ones. The reason is that in their model smaller firms are more privacy-intrusive, hence a privacy regulation is more likely to affect their privacy choice.[26] We find that this intuition turns around if we allow asymmetry between firms, in our case stemming from a single technology parameter, Δ.

Our model presented here draws attention to technological factors that may drive the differential empirical effect of the ePrivacy Regulation documented empirically in Section 4. It is important to note that these technological factors are not necessarily the only driving forces of that differential effect on large and small firms, and other explanations may exist.

One such alternative avenue could be that enforcement differed by firm size. While we may expect public enforcement to focus predominantly on large firms in order to prioritise intervention, large firms may be better able to engage in private enforcement against smaller rivals. Reputation damage and consumer damage claims in principle apply to firms of all sizes. We also observe public fines for small firms for breaching the 2009 ePrivacy Directive.[27] We cannot exclude that the 2009 ePrivacy Directive may have had differential effect on large and small firms due to *institutional* reasons. Our aim in this article is to empirically document differences in the effect of the regulation by firms of different sizes, and to provide a possible novel theoretical explanation that would explain these differences through technology-related factors.

6. Concluding remarks

In this paper, we investigate the relationship between privacy regulation and market structure. We aim to provide some policy guidance in the discussion surrounding the European Commission's proposed ePrivacy Regulation. Our analysis contributes to the broader discussion of the effect of privacy regulation on market outcomes.

Our empirical assessment focuses on the revision of the 2009 ePrivacy Directive (*Cookie Law*) that introduced an opt-in system for cookies in the European Union and finds that the privacy regulation had little effect on the revenues of e-commerce firms in Europe. Our empirical analysis relies on a *difference-in-difference-in-differences* model with variation in treatment timing. Our results show that small firms have not been significantly affected, while large firms suffered relatively minor revenue losses due to stricter privacy rules. This goes against the arguments of industry representatives who have harshly criticised privacy regulation as being harmful to businesses.

We provide a simple yet novel theory to argue that if the newly proposed EU ePrivacy Regulation should have any effect on businesses, it may well affect primarily larger firms, even if these may offer more privacy than smaller ones. Our theoretical model is based on the idea that larger firms are better able to turn data into increased sales, for example, due to technological reasons, such as economies of scale in data, or better analytical capabilities. If larger firms are more productive in data use, the regulation affects them disproportionately by reducing the amount of data available to these firms, as users are less likely to consent the cookie use.

Our model also predicts that smaller firms offer relative to the large firm less privacy. Even if - as *sceptics* argue - the privacy regulation may not be directly binding for large firms, competition forces these businesses to reduce their privacy intrusiveness in response to their smaller rivals doing so, for whom the regulation may be binding.

Our results carry relevance for the ongoing debate about the proposed European ePrivacy Regulation, whose adoption is staggering predominantly due to concerns about the competitiveness of European online businesses. Our empirical results looking at the last round of revision of the very same regulation suggest a more nuanced and optimistic view: the 2009 revision of the ePrivacy Directive had only minor negative effects on the revenues of large European online businesses, and had no measurable effect on smaller firms.

Declaration of Competing Interest

I declare not having received financial support from anybody for writing this publication.

Appendix A

Proof of Proposition 1

We first show that in equilibrium $q_1^* < q_2^*$. Since it is not possible to algebraically calculate the equilibrium values, we resort to an alternative proof. In particular, we prove that $q_1^* < q_2^*$ by demonstrating that the reaction function of Retailer 1 $(q1(q2))$ crosses the 45 ° line at a lower value for q_2 than where the Reaction function of Retailer 2 crosses the 45 ° line. Since - as we established in the main text - both reaction functions are upward sloping, this implies that the intersection of the reaction curves is below the 45-degree line, which means that $q_2^* > q_1^*$. This is illustrated in Fig. A2.

From the main text, the reaction functions are as follows:

$$q_1(q_2) = q_2 + t + \frac{1 - W\left(e^{\Delta(q_2+t)+1}\right)}{\Delta}, \tag{A.1}$$

$$q_2(q_1) = q_1 + t + 1 - \frac{W\left(e^{q_1+t+1}\right)}{\Delta},$$

We first calculate the value for q_2 at which these reaction functions intersect with the 45 ° line (with q_1 plotted on the vertical axis and q_2 on the horizontal). Let $\widehat{q}_2^1$ and $\widehat{q}_2^2$ denote the values of q_2 at which the respective reaction functions of Retailer 1 and Retailer 2 cross the 45-degree line. To obtain $\widehat{q}_2^1$ we solve $q_1(q_2) = q_2$. To obtain $\widehat{q}_2^2$ we solve $q_1^{-1}(q_2) = q_2$. We then have

$$\widehat{q}_2^1 = \frac{\ln(1+t\Delta)}{\Delta},$$

$$\widehat{q}_2^2 = \ln(1+t).$$

Note that $\widehat{q}_2^1 < \widehat{q}_2^2 \iff \ln(1+t\Delta) < \Delta\ln(1+t) \iff \ln(1+t\Delta) < \ln(1+t)^{\Delta} \iff 1+t\Delta < (1+t)^{\Delta} \iff \frac{1+t\Delta}{(1+t)^{\Delta}} < 1$.

The lhs of the last inequality decreases is t, since $\frac{\partial lhs}{\partial t} = -t(1+t)^{-\Delta-1}(\Delta-1)\Delta < 0$. It is therefore sufficient to show that $lhs = 1$ holds for $t = 0$. Since lhs decreases in t, any positive value of t will render $lhs < 1$. We plug $t = 0$ into lhs and get $lhs = \left.\frac{1+t\Delta}{(1+t)^{\Delta}}\right|_{t=0} = 1$. It follows that $lhs < 1$ if $t > 0$ and so $\widehat{q}_2^1 < \widehat{q}_2^2$. This in turn implies that $q_2^* > q_1^*$, Q.E.D.

Having established that $q_1^* < q_2^*$, it follows from Expression (9) that $n_1^* > n_2^*$. Q.E.D.

It is immediate from Expression (8) that $s_1^* > s_2^*$ iff $q_1^*\Delta > q_2^*$. To prove that $q_1^*\Delta > q_2^*$ it is sufficient to show that $\Delta\widehat{q}_2^1 > \widehat{q}_2^2$. This is so because $q_1(\widehat{q}_2^1) < q_1^*$ and $q_2^* < \widehat{q}_2^2$. The relationship $\Delta\widehat{q}_2^1 > \widehat{q}_2^2$ corresponds to $\ln(1+t\Delta) > \ln(1+t)$ which holds for any $\Delta > 1$. Q.E.D.

Having proven that $s_1^* > s_2^*$, we now turn to proving that $f_1^* > f_2^*$. Notice in Expression (11) that $f_i^* = s_i^* n_i^* / 2t$. With $s_1^* > s_2^*$ and $n_1^* > n_2^*$ it is therefore immediate that $f_1^* > f_2^*$. Q.E.D.

[26] A polar example may be if regulation only affected the firm offering less privacy, which in the setup of Dimakopoulos and Sudaric (2018) is the firm facing less intense competition.

[27] See Appendix C.

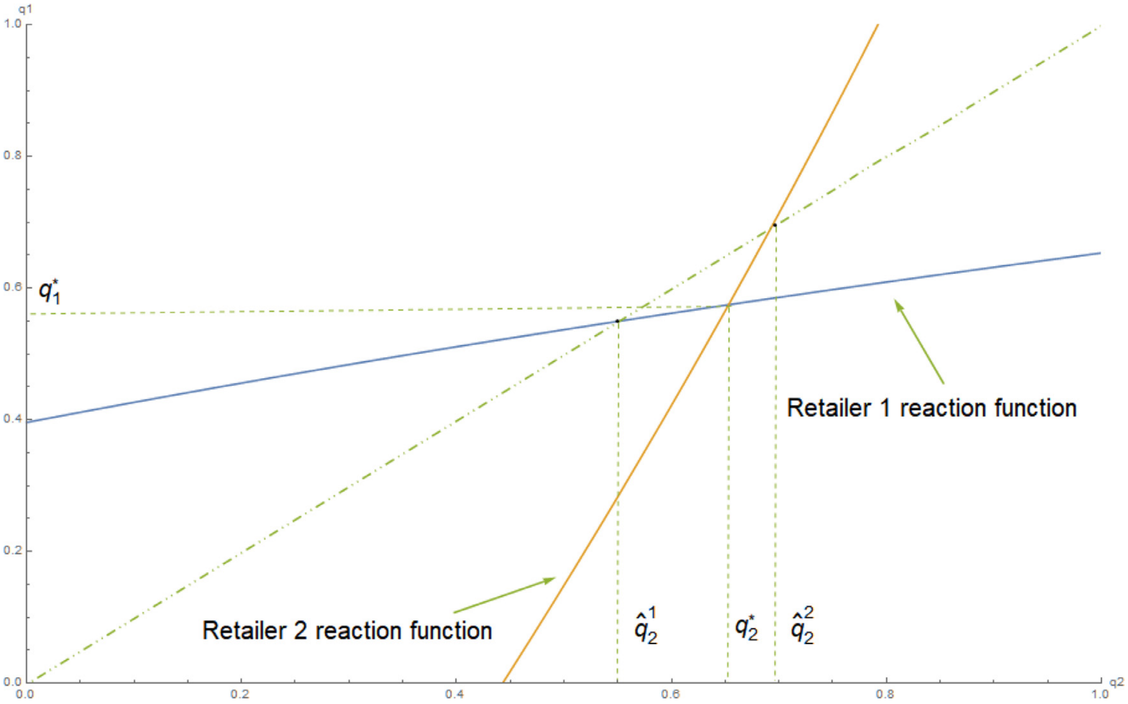


Fig. A2. The intersection of the reaction functions is below the 45 ◦ line if $\hat{q}_2^1 < \hat{q}_2^2$.

We now turn to the proof of the claim that $a_1^*> a_2^*$. We can conveniently re-write Expression (7) as $a_i^* = s_i^* n_i^* f_i^*$. With $s_1^* > s_2^*$, $n_1^* > n_2^*$ and $f_1^* > f_2^*$ we therefore have $a_1^* > a_2^*$. Q.E.D.

Finally, we prove that $\Pi_1^* > \Pi_2^*$. Since $\Pi_i^* = a_i^* f_i^*$ and $a_1^* > a_2^*$ as well as $f_1^* > f_2^*$ we have $\Pi_1^* > \Pi_2^*$. Q.E.D.

Appendix B. National Transpositions of the 2009 ePrivacy Directive

Cyprus: Cypriot ePrivacy laws are contained in Part 14 of the Law on the Regulation of Electronic Communications and Postal Services, Law No. 112(I)/2004 as amended, which came into force on 18 May 2012 (the æePrivacy Lawg). Source: https://www.linklaters.com/en/insights/data-protected/data-protected---cyprus.

Denmark: The Danish Executive Order on Cookies No. 1148 of 9 December 2011 entered into force on 14 December 2011. Source: https://www.linklaters.com/en/insights/data-protected/data-protected---denmark.

Finland: Amendment about cookies to the The Act on the Protection of Privacy in Electronic Communications (516/2004, in Finnish: Sähköisen viestinnän tietosuojalaki) entered into force on 25 May 2011. Source: https://www.finlex.fi/en/laki/kaannokset/2004/en20040516_20110365.pdf.

France: France has implemented the EU Cookies Directive by Order N× 2011-1012, dated August, 24, 2011. Source: https://www.linklaters.com/en/insights/data-protected/data-protected---france.

Germany: Special case. See Appendix C.

Greece: Greece has implemented the EU Cookies Directive through the Law 4070/2012, which entered into force on the 10th of April 2011. Source: https://gdprhub.eu/Data_Protection_in_Greece.

Ireland: ePrivacy Directive was implemented into Irish law with effect from the 1st of July 2011 through the European Communities (Electronic Communications Networks and Services) (Privacy and Electronic Communications) Regulations 2011 (SI 336 of 2011). Source: http://www.irishstatutebook.ie/eli/2011/si/336/.

Italy: On May 28, 2012 the Italian Government issued Legislative Decree no. 69/2012 (æDecreeg) implementing in Italy the Directive no. 2009/136/EC (which amended Directive 2002/58/EC æePrivacy Directiveg). Source: https://portolano.it/news/legislative-decree-692012-implements-in-italy-the-e-privacy-directive-2009136ec.

Norway: The Marketing Control Act, dated 9 January 2009, implemented Article 13 of the Privacy and Electronic Communications Directive. The Marketing Control Act came in to force on 1 June 2009. The Marketing Control Act, the Ecommerce Act and the Ecommerce Regulation were amended on 1 July 2013 to implement the amendments to the Privacy and Electronic Communications Directive. Source: https://www.linklaters.com/en/insights/data-protected/data-protected---norway.

Poland: The 2009 ePrivacy Directive, in particular article 5(3), was implemented into Polish law by Act of the 16th of November 2012 on the Change of the Telecommunication Law and Other Acts. This Act amended, among others, Article 173 of the Telecommunications Law (TL), which governs cookies. This change took effect on 22 March 2013. Source: https://www.uke.gov.pl/gfx/uke/userfiles/m-pietrzykowski/telecommunications_act_en.pdf.

Romania: The Law no. 506/2004 of 17 November 2004 regarding the processing of personal data and the protection of privacy in the electronic communications sector (the æPECRg). The PECR came into force on 28 November 2004 and has been amended in 2012 in order to implement the amendments to the Privacy and Electronic Communications Directive. Source: https://www.linklaters.com/en/insights/data-protected/data-protected---romania.

Slovenia: The 2009 ePrivacy Directive was implemented in Slovenia by an amendment to the Act on Electronic Communications (In Slovenian: Zakon o elektronskih komunikacijah; ZEKom-1). It came into force on the 15th of January 2013. Source: http://pisrs.si/Pis.web/pregledPredpisa?id=ZAKO6405#.

Spain: The Spanish Information Society Services and Electronic Commerce Law was amended on April 2012 to implement the changes required by the 2009 ePrivacy Directive. Source: https://www.hldataprotection.com/2012/04/articles/international-eu-privacy/at-last-the-eu-cookies-regulation-is-implemented-in-spain-2/.

Sweden: Sweden implemented the ePrivacy Directive through amendments to the Electronic Communications Act (2003:389), which came into effect on 1 July 2011. Source: https://www.linklaters.com/en/insights/data-protected/data-protected---sweden.

UK: The 2009 ePrivacy was implemented in the UK in May 2011 through The Privacy and Electronic Communications (EC Directive) Regulations 2003, as amended by the Privacy and Electronic Communications (EC Directive) (Amendment) Regulations 2011. Source: https://www.legislation.gov.uk/uksi/2011/1208/contents/made.

Appendix C. Privacy, Data Monetization, and Firm Size

We provide some qualitative and quantitative evidence related to assumptions and testable predictions of our model, including that larger firms seem to provide more privacy.

Larger Firms May Offer More Privacy. Some empirical facts seem to validate the prediction of our model that larger firms offer more privacy on their websites than smaller ones. We gathered traffic data for the 6000 most popular websites in the U.S. in terms of monthly visitors from QuantCast.org. We furthermore obtained a privacy score for these websites from PrivacyScore.org.[28] This organization rates websites according to four privacy dimensions: whether tracking services are used, whether selected attacks are prevented, the quality of encryption during data transmission to the website, and the quality when sending e-mails to an existing e-mail server. The precise methodology is outlined in more detail in Maass (2017). We could match the privacy score data with monthly traffic for 3952 websites. Privacy ratings are ordinally measured on a five-level scale. We associated numerical values with these categories as follows: critical (-3), bad (-2), warning (-1), neutral (0), and good (1). Fig. C1 depicts the relationship between the (log number of) monthly users and the lowest privacy rating of the website across the four evaluated categories.[29] Based on this simple analysis it appears that larger websites offer more privacy than less frequented ones.[30]

Large Firms Have an Advantage in Monetizing Data. Our theoretical model aims at isolating the effect of a firm having a technology advantage that allows it to turn data on website visitors into revenues. An empirically testable result emerging is that such firms are also larger as due to better privacy they are also able to attract more users.

There is some evidence suggesting that larger firms included in our financial dataset are indeed better able to monetise data. To show this, we combined publicly available data on websites owned by the firms in our financial dataset used in the empirical analysis in Section 4 with a proprietary dataset covering purchase conversions on and visits to these websites.[31] The use of cookie data is often motivated by the aim of increasing conversions on e-commerce websites (Shah, 2021).

We were able to identify 81 firms in our financial database owning in total 382 domains, such as eprice.it, otto.de, or blacks.co.uk. Of these, we found visit and conversion data for 95 domains.[32] We regressed the website conversion rate on the logarithm of website visits using OLS. Fig. C2 provides a graphical representation.[33] This is consistent with the view that larger websites in our dataset (typically owned by larger firms) are better able to convert visitors into e-commerce transactions.

The 2009 ePrivacy Directive Led to a Change in the Privacy and Data Protection Practice of European Firms. We found evidence directly or indirectly indicating that some firms in our dataset changed their behaviour in response to the 2009 ePrivacy Directive. In particular, we were able to retrieve the historic privacy policies for some websites owned by firms in our dataset. We furthermore gathered publicly available evidence and testimonies that may directly or indirectly indicate changes in firms' behaviour.

We relied on the *EntityMap* in the *DuckDuckGo TrackerRadar* dataset to identify websites operated by some of the firms in our dataset. Among these, we have eight online EEA firms owning 58 domains. We then manually inspected some of these websites using the *Wayback Machine*, if possible before and after the date we indicate in Table 5 to identify changes in the privacy policy texts.[34] A limitation is that the *Wayback Machine* may not always save scripts associated with a webpage, such as a cookie popup banner, one of the main targets of the 2009 ePrivacy Directive. It furthermore may not always record all webpages on a domain, such as the privacy policy page. Our inspection is best regarded as a one-way test: We cannot conclude from this exercise that no change occurred on the website due to the ePrivacy Directive, but we can confirm changes that took place on the websites. We find the following:

eprice.it: The first recorded version of the website on the *Wayback Machine* dates back to the 10th of February 2014.[35] From Table 5, the date we took as the event in Italy was the 30th of May 2012. Since archive.org did not crawl this page before and we know it tends to crawl more popular pages, it is possible that the website had no privacy policy before the ePrivacy Directive or that the link was not prominently placed to be frequently visited.

otto.de: The event date in our empirical analysis for Germany was the 10th of May 2012 (Table 5). We inspected the website on the *Wayback Machine* before and after this date, namely on the 7th of May 2011 and the 27th of November 2011.[36] We see no privacy policy link on the early version, but do see one on a later date.[37]

schlafwelt.de: We inspected the privacy policy page of the domain for three dates around the event date assumed for Germany

[28] Independently from us and in parallel to our research, Ramadorai et al. (2019) used the same source of website privacy ratings to argue, as we do, that larger firms provide more privacy.

[29] Since plotting such a large number of observations results in crowded graphs, we ordered websites into 50 bins and plot the bin average privacy rating.

[30] This relationship is largely confirmed in a series of OLS and ordered response regressions along with various privacy categories (e.g. web encryption, mail encryption, attack vulnerability, web tracking, both for individual categories and the average as well as the minimum score across these categories), explaining website privacy scores by the number of monthly users. The regression outputs are available upon request from the authors.

[31] The source of data on domains owned by firms is the public *DuckDuckGo TrackerRadar* dataset, available at https://github.com/duckduckgo/tracker-radar. Website conversions and visits for 2019–2020 were kindly provided by *SimilarWeb*, an industry-leading provider of website analytics and data. More information can be found under www.similarweb.com. Further detail on the method is available upon request from the authors.

[32] The conversion rate is calculated as *the share of "visits that included the completion of the primary business goal of the website"* as provided by SimilarWeb, over the period 2019–2020.

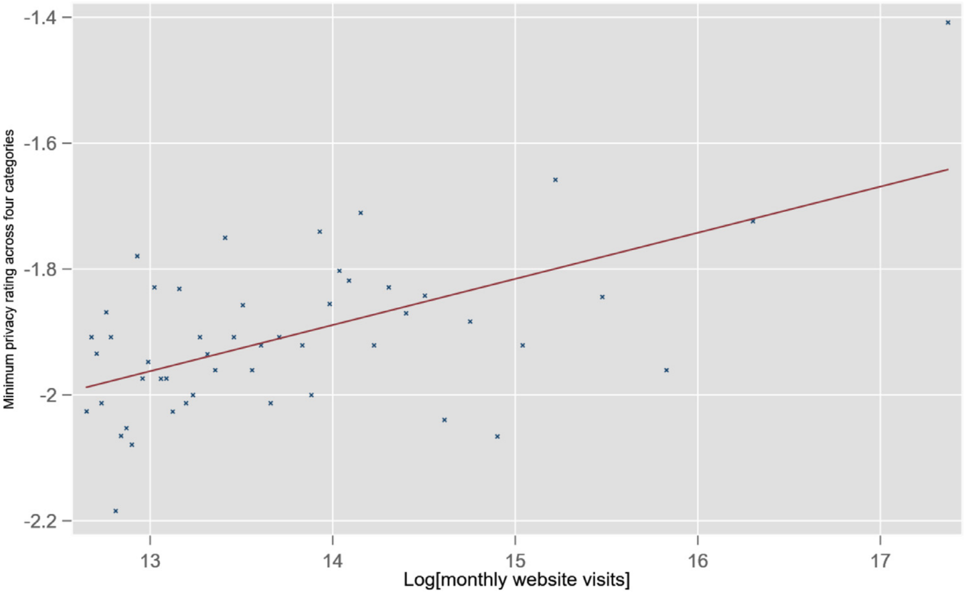
[33] The regression coefficient for $log(visits)$ is statistically significant at the 1% level.

[34] The *Wayback Machine* available under http://web.archive.org/ maintains an archive of websites logged at periodic intervals.

[35] Source: https://web.archive.org/web/20140210200813/https://www.eprice.it/default.aspx?zona=1&dove=24.

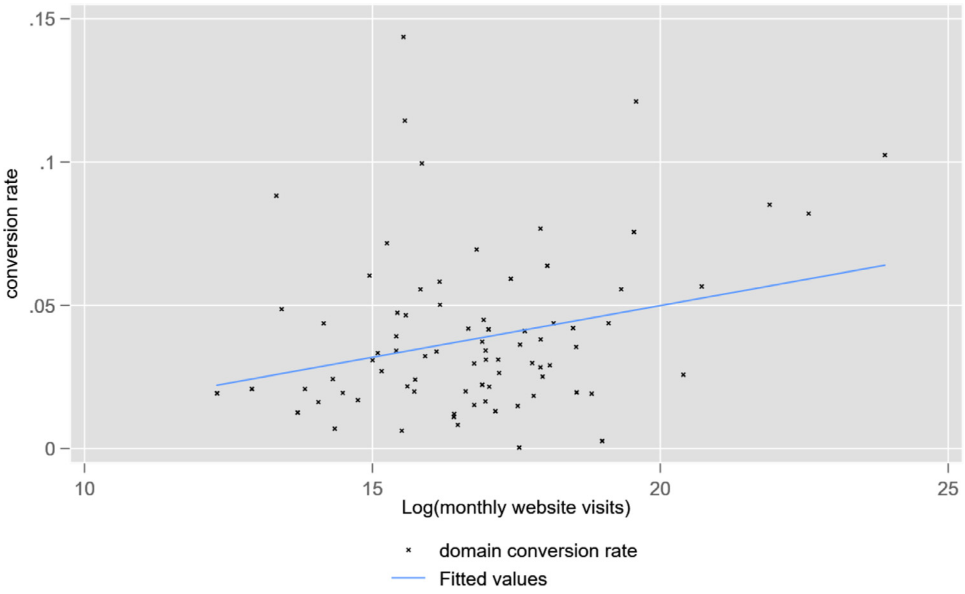
[36] Sources: https://web.archive.org/web/20110507221933/http://www.otto.de/ and https://web.archive.org/web/20121127225357mp_/http://www.otto.de/.

[37] Under *"Rechtliches, Datenschutz."*



3,952 US domains in 50 bins. The privacy rating for each domain is an integer -2: bad, -1: warning, 0: neutral, 1: good across categories 'tracking intensity', 'encryption quality', 'attacks prevented' and 'email encryption'.
Dots mark the bin-averages of the lowest rating for the included websites across the four categories. Outlier bins removed.

Fig. C1. Relationship between website popularity and privacy level.



95 domains owned by 49 firms in our dataset

Fig. C2. Relationship between e-commerce website visits and conversions.

(10th of May 2012). While in January 2012 the privacy policy page had no explicit language on the use of retargeting technologies (nor on the 11th of May), this changed at the latest by the 20th of September 2012.[38] Furthermore, the 2012 September version contains language on opt-in to the email newsletter that we did not find in the earlier versions of the privacy policy text. These additions reflect obligations under the 2009 update of the ePrivacy Directive to obtain consent before cookie use and opt-in to email communication.

Qualitative Evidence on Changed Firm Behaviour due to the 2009 ePrivacy Directive. While various EU member states may have had different approaches to implementing the 2009 ePrivacy Directive provisions, we found documentary evidence indicating that the Directive had consequences on firms. Spain was the first to issue fines for practices falling under the 2009 update of the ePrivacy Directive, in 2014.[39] The firm in question was fined for its conduct in 2012, failing to provide cookie policies on the websites. It is furthermore indicative of the effectiveness of the ePrivacy Directive that consumers were entitled to damage claims.[40]

[38] We looked at the archived pages https://web.archive.org/web/20120113225951/http://www.schlafwelt.de/static/privacy, https://web.archive.org/web/20120511014038/http://www.schlafwelt.de/static/privacy and https://web.archive.org/web/20120920091527/http://www.schlafwelt.de/static/privacy.

[39] Source: https://marketinglaw.osborneclarke.com/data-and-privacy/spain-imposes-first-fines-in-europe-for-breach-of-cookie-laws-2/.

[40] Source: https://www.loc.gov/law/help/online-privacy-law/2012/spain.php, retrieved on the 3rd of June, 2021.

We furthermore found documents of a meeting with e-commerce industry associations (IAB and EASA) and regulators in September 2011, where the industry association still tried to argue that users' inaction may be interpreted as consent to cookies. At that time the regulator made clear that this was *not* the case.[41] The same view was confirmed by the EU's privacy law guidance body (Working Party 29) in December 2011.[42]

Germany was a special case among EU member states. In particular, Germany took the view regarding the 2009 update of the ePrivacy Directive that its laws at the time already reflected the principles in the Directive. It, therefore, undertook no additional legislative changes. Even in 2019, the cookie regime in Germany was largely based on an opt-out principle, despite surrounding suspicions that this may not be in line with the ePrivacy Directive, a unique situation in the EU.[43] This changed in 2019, when the EU Court of Justice ruled that Germany failed to implement the ePrivacy Directive appropriately and the opt-out principle on cookies was unlawful.[44]

This, however, does not mean that in Germany the 2009 revision of the ePrivacy Directive would not have had any effect on firms. Beyond changing the regime on cookie consent, the 2009 revision round had further important implications. As noted by Papakonstantinou and de Hert (2011), *"perhaps the most important amendment concerning spam refers to the fact that the ePrivacy Directive substantially enlarged the circle of parties with a right to sue spammers[...] Practically, therefore, all parties directly or indirectly involved in a typical spamming activity will be authorised to sue independently of each other."* This in turn likely had serious implications on the use of email marketing. We know that such litigation took place in Germany, explicitly referring to the 2009/136 ePrivacy Directive.[45]

There is contemporary documentary evidence acknowledging that firms faced significant costs in association with Directive EC/2009/136.[46] These costs include cookie audits, due diligence of ad partners before contracting, legal review from clickwrap agreements, ensuring effective contracts that respect the privacy regulations, post-contract monitoring, testing, and evaluating agreements. The fact that these costs are estimated to be substantial is consistent with firms changing their behaviour in compliance with the privacy directive.

Finally, archives of U.S. SEC filings also testify that Directive EC/2009/136 had revenue implications for firms. For example, *Criteo*, a major online marketing firm, noted in relation to the 2009 ePrivacy Directive that *These existing and proposed laws, regulations and industry standards can be costly to comply with and can delay or impede the development of new products, result in negative publicity and reputational harm, increase our operating costs, require significant management time and attention, increase our risk of non-compliance and subject us to claims or other remedies, including fines or demands that we modify or cease existing business practices.*[47]

Similarly, *Rubicon Project*, another leading online marketing platform filed in relation to the 2009 ePrivacy Directive that *Limitations on the use or effectiveness of cookies, whether imposed by regulation or otherwise, may impact the performance of our solution. We may be required to, or otherwise may determine that it is advisable to, develop or obtain additional applications and technologies to compensate for the lack of cookie data, which may require a substantial investment on our part. However, we may not be able to develop or implement additional applications that compensate for the lack of cookie data. Moreover, even if we are able to do so, such additional applications may be subject to further regulation, time-consuming to develop or costly to obtain, and less effective than our current use of cookies."*[48]

The 2009 ePrivacy Directive Was More Likely to Affect Small Firms Directly by Limiting E-Mail Marketing and Display Ads. Our theoretical model predicts that privacy regulation may affect small firms directly, but large firms possibly only indirectly.

The 2009 ePrivacy Directive has been widely discussed to have limited display advertising.[49] Small websites however are more likely to rely on display ads than large websites as a source of traffic, given their limited brand reach: users are less likely to type in their URL directly into the browser.

In a similar vein, while the 2009 ePrivacy Directive became famous as the *Cookie Law*, one of its particularly important novelties was not related to cookies, but to the prohibition of unsolicited communication without prior consent in its Articles 6(3) and 13. This likely had significant implications for email marketing, which now required explicit opt-in. We hypothesise that the drying up of email marketing as a traffic acquisition channel likely affected predominantly small firms. This is, again, because large firms tend to have established brands that attract direct traffic.[50]

To demonstrate that small firms are indeed more likely to rely on email marketing and display ads as a source of traffic, we obtained data from SimilarWeb with the share of email and display ad traffic on total visits in the 1st week of November 2017 for 349 websites in the categories *"E-commerce and Shopping"* as well as *"E-commerce Related"* in France, Germany, Italy, Netherlands, United Kingdom and United States.[51] While we were not able to retrieve similar data from the time of the implementation of the 2009 ePrivacy Directive, our 2017 data stems from before the GDPR. This is important, as the GDPR likely further restricted e-mail marketing and display ad targeting.

It is very likely that before the 2009 ePrivacy Directive smaller websites relied heavier on both email marketing and display ads than larger websites. This is indeed what we observe in our data in 2017: while the share of email traffic on total visits was 6.77% for small websites, the largest websites only obtained 4.25% of visits via links embedded into email messages.[52]

Similarly, in 2017 display ads still served as a source of a relatively large share of website traffic for small and medium websites (respectively 2.15% and 4.58%), while large and especially gi-

[41] Source: https://web.archive.org/web/20111005162442/http://ec.europa.eu/justice/data-protection/article-29/press-material/press-release/art29_press_material/20110914_press_release_oba_industry_final_en.pdf.

[42] Source: https://web.archive.org/web/20120111082509/http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2011/wp188_en.pdf, under *"User choice over Online Behavioural Advertising"*.

[43] Source: https://bvdw.org/fileadmin/user_upload/BVDW_Datenschutzkonformes_Affiliate_Marketing_2020.pdf.

[44] EU Court of Justice Ruling in Planet49, C-673/17.

[45] Source: https://www.online-und-recht.de/urteile/Werbung-in-Autoreply-E-Mails-ist-Spam--Bundesgerichtshof-20151215/.

[46] Source: https://www.dlapiper.com/en/uk/insights/publications/2014/09/eu-law-on-cookies/.

[47] Source: https://www.sec.gov/Archives/edgar/data/1576427/000119312513369592/d541385df1.htm.

[48] Source: https://www.sec.gov/Archives/edgar/data/1595974/000119312514128315/d652651d424b4.htm.

[49] For example, one of the few studies discussing the effect of the ePrivacy Directive on firms explains that *"Display advertising depends not only to a large extent on behavioral targeting but also on access to the storage medium to ensure the correct display of ads. The ePrivacy Directive is expected to render these practically impossible"* (Arnold and Hildebrandt (2017), last paragraph of page 24, translated from the German original by the authors.)

[50] It is easy to interpret parameter q_i as the intrusiveness of spam emails sent or intrusive display ads placed by website i. Other things equal, intrusive spam emails or display ads cause a disutility to consumers (hence the negative sign in Expression 5). They however increase the probability of sale, as in Expression 8.

[51] We do not restrict this analysis to domains owned by the firms included in the empirical analysis in Section 4 as the overlap between the two datasets is too small.

[52] We use the size categories *small, medium, large* and *giant* based on total visits.

ant websites barely used this traffic channel (respectively 1.81% and 0.96% of total traffic).

This is likely largely due to the fact that smaller websites are less well known and therefore attract less traffic by users directly typing in their URL into the browser, hence relying more on intrusive marketing such as (spam) email and display ads. Indeed, direct traffic accounts for 36.7% of visits to small websites in contrast to 44.5% for giant websites.

The stricter rules on marketing emails, spam, and reduced targeting of display ads introduced by the 2009 ePrivacy Directive, therefore, were more likely to affect small firms directly than large firms, as small firms very likely relied heavier on these channels for traffic acquisition than larger firms.

Appendix D. Discussion of the Model

In our model asymmetry arises due to a single parameter governing firms' ability to monetise website visitors' data (Δ), that is, converting the data of each visitor into revenues.[53]

In reality, asymmetry may arise due to many factors such as intrinsic website popularity and reputation Kummer and Schulte (2019) or user-side network effects (Sapi and Schäfer, 2020). Furthermore, these factors may also interrelate: a platform benefiting from user-side network effects may also become better at monetizing visitors by placing more relevant ads. This in turn can increase its reputation, and generate further user-side network effects. To isolate effects in a tractable setup, we intentionally break this loop in our modelling approach. The aim of our modelling approach is to highlight the role of a specific aspect of technology - the ability to monetise visitors' data - in determining market shares, and eventually the impact of privacy regulation.

An important insight of our model is that larger firms may offer a higher privacy level than smaller firms.[54] This prediction is consistent with some empirical analyses conducted elsewhere. Goldberg et al., 2021 report empirical support for smaller e-commerce firms being less able to get consent to privacy practices than larger firms.[55] This is precisely what one would expect if users perceived the privacy practices of smaller websites as more intrusive than those of larger websites. It however appears at first glance inconsistent with the findings of Preibusch and Bonneau (2013), who - looking into the privacy practices of around 140 websites - found no significant difference in the privacy practices of websites charging higher and lower prices for the same products or services.[56]

In terms of theory, the finding that larger firms may offer more privacy is also confirmed by Casadesus-Masanell and Hervas-Drane (2015), who show that whether the larger or smaller market share firms offer more privacy may depend on how consumers intrinsically value the services of these firms.[57] In a related model, Dimakopoulos and Sudaric (2018) report that advertising-supported platforms facing less competition on either consumer or advertiser side tend to offer less privacy.[58] In the model of Dimakopoulos and Sudaric the platforms are symmetric. Our model adds to that analysis by showing that asymmetry may qualify and even turn around the relationship between market power and privacy provision.

References

Acquisti, Alessandro, Taylor, C., Wagman, L., 2016. The economics of privacy. J. Econ. Lit. 54.2, 442–492.

Akter, Shahriar, Wamba, S.F., 2016. Big data analytics in e-commerce: a systematic review and agenda for future research. Electron. Mark. 26.2, 173–194.

Apostle, J., 2018. We survived GDPR, now another EU privacy law looms. Financ. Time..

Arnold, R., Hildebrandt, C., 2017. Wirtschaftliche auswirkungen der regelungen der eprivacy-verordnung auf die online-werbung und werbefinanzierte digitale geschäftsmodelle.

Arora, A., 2016. Recommendation engines: how amazon and netflix are winning the personalization battle. MarTech Advis..

Ashenfelter, O., Card, D., 1985. Using the longitudial structure of earnings to estimate the effect of training programs. Rev. Econ. Statistics 67 (4), 648–660.

Baye, I., Sapi, G., 2017. Should mobile marketers collect other data than geo-location? Scand. J. Econ..

Campbell, J., Goldfarb, A., Tucker, C., 2015. Privacy regulation and market structure. J. Econ. Manage. Strat. 24.1, 47–73.

Card, D., 1992. Using regional variation in wages to measure the effects of the federal minimum wage. ILR Rev., 46.1, 22–37

Card, D., Krueger, A.B., 1993. Minimum wages and employment: a case study of the fast food industry in new jersey and pennsylvania. no. w4509. Natl. Bur. Econ. Res..

Casadesus-Masanell, R., Hervas-Drane, A., 2015. Competing with privacy. Manage. Sci. 61.1, 229–246.

Corless, R.M., et al., 1996. On the lambert w function. Adv. Comput. Math. 5.1, 329–359.

Dimakopoulos, P.D., Sudaric, S., 2018. Privacy and platform competition. Int. J. Ind Organiz 61, 686–713.

European Commission, 2011. Special Eurobarometer 359: Attitudes on Data Protection and Electronic Identity in the. European Union.

European Commission, 2018. Proposal for an eprivacy regulation. Digital Single Market.. https://ec.europa.eu/digital-single-market/en/proposal-eprivacy-regulation.

McConnell, E., 2019. The pending european eprivacy regulation (ePR). https://www.sgrlaw.com/ttl-articles/pending-european-eprivacy-regulation-epr/, retrieved on the 28th of March.

Episerver, 2018. Reimagining Commerce: Global Findings. Experience-Driven Commerce Outperforms Conversion-Focused Strategies. https://www.episerver.com/globalassets/assets-website-structure/resources/guides-and-reports/episerver-reimagining-commerce-report-2018.pdf (retrieved on the 20th of August 2019). An archived version is available under: https://web.archive.org/web/20190823223927/https://www.episerver.com/globalassets/assets-website-structure/resources/guides-and-reports/episerver-reimagining-commerce-report-2018.pdf

Eurobarometer, 2015. Special Eurobarometer 431: Data protection. Available at: http://data.europa.eu/88u/dataset/S2075_83_1_431_ENG [Accessed June 13, 2022].

Farrell, J., 2012. Can privacy be just another good. J. on Telecomm. & High Tech. L. 10, 251.

Ghosh, S., 2018. Europe will become a digital backwater: there's a new war over online privacy and metadata in europe right now. Bus. Insid.. May 30

Goldberg, S., Johnson, G., Shriver, S., 2021. Privacy & market concentration: Intended & unintended consequences of the GDPR. Available at SSRN, 3477686 doi:10.2139/ssrn.3477686.

Goldfarb, A., Tucker, C.E., 2011. Privacy regulation and online advertising. Manage. Sci. 57.1, 57–71.

Gomes, R., Tirole, J., 2018. Missed sales and the pricing of ancillary goods. Q. J. Econ. 133.4, 2097–2169.

Goodman-Bacon, A., 2021. Difference-in-differences with variation in treatment timing. J. Econom..

Eurostat, 2021. Households - level of internet access. Table ISOC_CI_IN_H, https://ec.europa.eu/eurostat/databrowser/view/isoc_ci_in_h$DV_512/default/table?lang=en [Accessed June 13. 2022]

Jia, J., Jin, G.Z., Wagman, L., 2021. The short-run effects of the general data protection regulation on technology venture investment. Market. Sci. 40.4, 661–684.

Khan, M., 2018. EU States urged to agree online privacy 'cookie law'. Financ. Time.. April 24

Kosta, E., 2013. Peeking into the cookie jar: the european approach towards the regulation of cookies. Int. J. Law Inform. Technol. 21.4, 380–406.

Kox, H., Straathof, B., Zwart, G., 2017. Targeted advertising, platform competition, and privacy. J. Econ. Manage. Strat. 26.3, 557–570.

[53] Dimakopoulos and Sudaric (2018) provide a similar setup, where firms are asymmetric due to one of them being more appealing to consumers. In that setup, a platform's market share decreases in the privacy costs it induces on users. This can easily be seen for example in Expression 4 of Dimakopoulos and Sudaric (2018), where $\kappa(d_i)$ corresponds to this privacy cost.

[54] Appendix C provides further supporting evidence that this is indeed the case.

[55] See Goldberg et al., 2021, section 6.2.3.

[56] Hypothesis 3 in Preibusch and Bonneau (2013). Other things equal, since the comparison involves websites offering very similar services and products, we would expect the lower-priced website to be larger in terms of visitors. If this is the case, no difference in privacy practices between high- and low-price websites offering the same products also implies no privacy difference between larger and smaller websites.

[57] Proposition 4 in Casadesus-Masanell and Hervas-Drane (2015). Larger firms may provide more privacy (less *disclosure*) if the average valuation of firms is high, but not so high to induce Bertrand competition.

[58] Less intense competition is captured by higher transport costs in a Hotelling setup.

Kummer, M., Schulte, P., 2019. When private information settles the bill: money and privacy in googles market for smartphone applications. Manage. Sci. 65.8, 3470–3494.

Gwynn, S., 2017. EU's eprivacy proposals 'would kill off half the digital ad market', campaign. September 11. 2017, https://www.campaignlive.co.uk/article/eus-eprivacy-proposals-would-kill-off-half-digital-ad-market/1444180 [Accessed June 13. 2022]

Lohr, S., 2011. The default choice, so hard to resist. N.Y. Times.

Maass, M., et al., 2017. Privacyscore: Improving privacy and security via crowd–Sourced benchmarks of websites. Annual Privacy Forum. Springer, Cham.

Lambrecht, A., 2017. E-privacy provisions and venture capital investments in the EU. https://www.ceps.eu/wp-content/uploads/2017/10/E-Privacy%20Provisions%20and%20Venture%20Capital%20Investments%20in%20the%20EU.PDF [Accessed June 13. 2022]

Papakonstantinou, V., de Hert, P., 2011. The amended EU law on eprivacy and electronic communications after its 2011 implementation; new rules on data protection, spam, data breaches and protection of intellectual property rights. J. Marshall J. Computer Info. L. 29, 29.

Peukert, C., 2022. Regulatory spillovers and data governance: evidence from the GDPR. Market. Sci..

Preibusch, S., Bonneau, J., 2013. The Privacy Landscape: Product Differentiation on Data Collection. In: Economics of Information Security and Privacy III.. Springer, New York, NY, pp. 263–283.

Ramadorai, T., Walther, A., Uettwiller, A., 2019. The Market for Data Privacy. CEPR Discussion Paper No. DP13588

Reinganum, J.F., 1983. Uncertain innovation and the persistence of monopoly. Am. Econ. Rev. 73.4, 741–748.

Shah, R., 2021. How personalization can increase website conversions for ecommerce websites. https://www.maropost.com/how-personalization-can-increase-website-conversions-for-ecommerce-websites/, January 14.

Schmidheiny, K., Siegloch, S. 2019, On Event Studies and Distributed-Lags in Two-Way Fixed Effects Models: Identification, Equivalence, and Generalization. CEPR Discussion Paper No. DP13477, Available at SSRN: https://ssrn.com/abstract=3324212

Schlosser, J.. GDPR is great - no question about it! but, the newer eprivacy directive might be the end of the internet as we know it.

Shy, O., Stenbacka, R., 2016. Customer privacy and competition. J. Econ. Manage. Strat. 25.3, 539–562.

Singer, N., 2018. The next privacy battle in europe is over this new law. N. Y. Time..

Thisse, J.-F., Vives, X., 1988. On the strategic choice of spatial price policy. Am. Econ. Rev. 122–137.

Tucker, C.E., 2012. The economics of advertising and privacy. Int. J. Ind. Organ. 30.3, 326–329.

Further Reading

TrackerRadar, DuckDuckGo TrackerRadar, https://github.com/duckduckgo/tracker-radar [Accessed June 13. 2022]