ETLA Working Papers | 77 25.2.2020

Short-term Impacts of the GDPR on Firm Performance

Heli Koski
The Research Institute of the Finnish Economy and Aalto University, Finland
heli.koski@etla.fi

Nelli Valmari
The Research Institute of the Finnish Economy
nelli.valmari@etla.fi

Suggested citation:
Koski, Heli & Valmari, Nelli (25.2.2020). "Short-term Impacts of the GDPR on Firm Performance".

ETLA Working Papers No 77.
http://pub.etla.fi/ETLA-Working-Papers-77.pdf

Abstract

This paper uses extensive firm-level data on European and US companies from 2014–2018 to explore the short-term impacts of the General Data Protection Regulation (GDPR) on European companies' financial performance.

Our empirical analysis suggests that the costs of the GDPR during the first year of its implementation were substantial, at least for some European companies. The profit margins of the data-intensive firms increased, on average, by approximately 1.7 to 3.4 percentage points less than the profit margins of their US counterparts. The European data-intensive SMEs were the most disadvantaged group regarding their post-GDPR profit developments, while the large European data-intensive companies' short-term post-GDPR profit margins dropped relatively less.

We do not find any statistically significant difference in the profit margin developments of the very large European and US companies. This finding is consistent with the view that the very large, multinational US companies that often have European customers and deal with the personal data of EU citizens also faced substantial costs when they needed to comply with the GDPR.

Tiivistelmä

Tietosuoja-asetuksen lyhyen aikavälin kannattavuusvaikutukset

Tämä tutkimus tarkastelee EU:n yleisen tietosuoja-asetuksen eli GDPR:n lyhyen aikavälin vaikutuksia yritysten kannattavuuteen laajan vuodet 2014–2018 käsittävän yritysaineiston avulla.

Empiirinen analyysimme viittaa siihen, että tiukentuneen tietosuojasääntelyn kustannukset olivat sen voimassaolon ensimmäisenä kalenterivuotena mittavat ainakin osalle eurooppalaisia yrityksiä. Dataintensiivisten yritysten voittomarginaalit kasvoivat keskimäärin 1,7–3,4 prosenttiyksikköä vähemmän kuin vastaavissa yhdysvaltalaisissa yrityksissä. GDPR vaikutti eniten pieni- ja keskikokoisten dataintensiivisten eurooppalaisten yritysten kannattavuuteen, kun taas isojen dataintensiivisten eurooppalaisten yritysten voittoihin GDPR vaikutti suhteellisesti vähemmän.

Eurooppalaisten ja yhdysvaltalaisten hyvin suurten dataintensiivisten yritysten kannattavuuskehityksessä ei näy tilastollisesti merkitsevää eroa. Tämä löydös on odotettu, koska hyvin suuret yhdysvaltalaiset yritykset toimivat eurooppalaisilla markkinoilla ja käsittelevät EU-kansalaisten henkilötietoja, joten myös niille on koitunut mittavia kustannuksia GDPR:n noudattamisesta.

Dr.Sc. (Econ) **Heli Koski** is a Research Director at The Research Institute of the Finnish Economy and a Research Director at Aalto University.

D.Sc. (Econ) **Nelli Valmari** is a Researcher at The Research Institute of the Finnish Economy.

KTT **Heli Koski** on tutkimusjohtaja Elinkeinoelämän tutkimuslaitoksessa ja tutkimusjohtaja Aalto yliopistossa.

KTT **Nelli Valmari** on tutkija Elinkeinoelämän tutkimuslaitoksessa.

Acknowledgement: This report has been drafted as a part of the BRIE-Etla 2019–2022 – Shaping the Future in the Era of Intelligent Tools: AI and Beyond research project funded by Business Finland.

Kiitokset: Tämä raportti on osa BRIE-Etla 2019–2020 – Älykkäät teknologiat, globaalit ICT-toimijat ja markkinoiden toimivuus -tutkimusprojektia, jonka on rahoittanut Business Finland.

Keywords: Regulation, GDPR, Compliance costs, Firm profitability

Asiasanat: Sääntely, GDPR, Sääntelyn noudattamiskustannukset, Yritysten kannattavuus

JEL: K2, L2, L5

1. Introduction

In May 2018, the General Data Protection Regulation (GDPR; Regulation (EU) 2016/679) came into force in all countries of the European Union. This regulation, passed in April 2016, meant a drastic change in European data privacy regulation because it gave individuals substantial control over their personal data.[1] Moreover, it obliged any organization active in the area of the European Union to ensure that it collects, stores and processes personal data in compliance with the GDPR.[2] The idea was further to unify data regulation within the EU and consequently to simplify the regulatory domain. The European Commission's ex ante assessment was that the single pan-European law concerning data protection would mean benefits reaching 2.3 billion euros per year.[3]

On the other hand, a recent discussion suggests that the compliance costs of the GDPR may be substantial. Some sources indicate that during the first year of implementation of the GDPR, approximately 500,000 European organizations hired a data protection officer, i.e., a DPO (whose average annual salary is approximately 80 000 euros in Europe).[4] This means an additional cost of 40 billion euros solely from the employment of the DPOs. In addition, organizations have needed to use their resources to adapt their information systems and practices to comply with the requirements of the GDPR. Fines from failing to comply may be considerable, as much as 20 million euros or 4% of the company's turnover. During the first nine months of the GDPR, the total fines issued amounted to approximately 56 million euros. Fifty million euros of this total amount is due to a single fine that the French data protection agency CNIL issued to Google.[5] However, given that European data protection agencies have handled approximately 100 000 self-reported breaches and user complaints under the GDPR, it becomes clear that the amount of fines issued does not give a complete picture of the extent to which European firms have achieved and maintained compliance with the GDPR.

Achieving and maintaining compliance with the GDPR is likely to increase costs and, hence, deteriorate the financial performance of all EU companies collecting and handling personal data. However, it is not clear what

[1] This means that an organization needs to comply with the rights of the data subjects. These include the right to be informed on how one's personal data is processed, the right to request one's data to be updated or corrected, the right to request the erase of one's data, the right to restrict the processing of one's data, the right to transfer one's data for reuse across different services, the right to object to the processing of one's personal data in certain circumstances, and the right not to be subject to decision making based merely on automated processing.

[2] The GDPR has been criticized for not defining what a "reasonable" level of protection means. Firms' responsibilities regarding the GDPR are discussed in the second paragraph of Section 2.

[3] Source: https://europa.eu/rapid/press-release_MEMO-17-1441_en.htm, accessed 2.10.2019. Alternative reference: "LIBE Committee vote backs new EU data protection rules", 22 October 2013, European Commission memo

[4] Source: https://iapp.org/about/approaching-one-year-gdpr-anniversary-iapp-reports-estimated-500000-organizations-registered-dpos-in-europe/, accessed 2.10.2019.

[5] Source: https://gdpr.eu/gdpr-fines-so-far/, accessed 17.11.2019

the actual magnitude of these effects is or whether the effects vary across different types of firms (e.g., small vs. large) and across different industries (consumer data intensive vs. non-intensive). In this paper, we discuss the following questions:

1) How has the implementation of the GDPR affected the financial performance of European firms in 2018, as opposed to their US counterparts?
2) Do the economic impacts of the GDPR differ across industries and firms of different sizes?

Currently, as the GDPR came into force only recently, the empirical literature on the economic impacts of the GDPR is scarce. Yuan and Li (2019) look at the early effects of GDPR compliance investments on the group of organizations dealing with highly sensitive personal data: they estimate the pre-2018 financial impacts of the GDPR on EU hospitals that provide digital health services as their primary business. They use the difference-in-difference method and employ hospitals with at most a small share of digital services in their operating business as their control group. Yuan and Li find a negative effect on the financial performance of the hospitals in the treatment group. Another group of firms that is heavily affected by the GDPR is online firms collecting and managing vast quantities of individual-level user data. Goldberg, Johnson and Shriver's (2019) study among approximately 1500 online firms shows that the enforcement of the GDPR further adversely affected European web traffic and e-commerce sales. Their weekly data for the 2nd through 38th weeks of 2017 and 2018 indicate an approximately 10% post-GDRR drop in recorded pageviews and revenues in Europe.

The GDPR may affect the performance of not only incumbent companies but also, via investors' future profit expectations, the funding available for newly established companies. Jia et al. (2019) compare venture capital investments in new and emerging technology firms in Europe and in the US before and after the enforcement of GDPR. Their estimations using monthly investment data from January 2014 to April 2019 suggest that the short-term negative impacts of the GDPR on EU technology venture investments were substantial.

The empirical research of these previous papers focuses on a single industry (Yuan et al.: hospitals; Goldberg et al.: e-commerce) or on certain types of firms (Jia et al.: new and emerging technology firms). Instead, we use comprehensive data on firms active in a variety of industries to estimate the short-term impacts of the GDPR on firms' financial performance.

We employ extensive firm-level data concerning European and US firms from 2014 to 2018 to explore the short-term profitability impacts of compliance with the GDPR. We do not find any statistically significant differences in the pre- and post-GDPR profit margin developments between the samples of European firms and their US

counterparts. The short-term evidence from the first year of the GDPR suggests that the implementation costs of the new privacy regulation affected the profitability of primarily data-intensive European companies. The European data-intensive SMEs suffered more in terms of their profit margins than the data-intensive SMEs in the US. Our data further suggest that the compliance costs of the GDPR for very large data-intensive companies did not differ substantially between European and US-based companies.

The rest of the paper is organized as follows. Section 2 discusses the actions that firms may need to take and the costs that they may incur to achieve and maintain compliance with the GDPR, which affect their financial performance. Section 3 introduces and describes the data used in the empirical analysis. Section 4 first presents the estimation model and then discusses the empirical findings. Section 5 concludes.

2. GDPR compliance and firm performance

The GDPR obliges firms to protect the personal data of consumers they may have and, in line with consumers' rights as set by the GDPR, respond to the inquiries and requests of consumers regarding their personal data. Achieving and maintaining compliance with the GDPR can be expected to incur significant costs and hence deteriorate firms' financial performance. Compliance may incur costs in various ways, as discussed below.

Human resources. Achieving GDPR compliance requires, first, that decision makers in firms understand what the GDPR implies and what kind of actions it necessitates. Second, it may require training of employees for the new tasks required for achieving and maintaining compliance. The actual responsibilities depend on whether the firm is what the GDPR defines as a data controller or a data processor. A firm that owns data is called a data controller. It is responsible for determining the purposes for which the data are used and the means of data processing. The data controller is also responsible for responding to consumers' requests concerning their personal data. A firm that processes data for a third party is called a data processor. It processes the data only on instructions from the data controller and keeps records of how the data are processed. Regardless of whether a firm is a data controller or a data processor, it is responsible for securing the data and keeping records of the actions taken for data security. A firm may be and have the responsibilities of both a data controller and a data processor. Under certain conditions, the firm needs to designate a data protection officer who is responsible for designing and implementing the firm's data protection plan as well as monitoring the implementation. In short, compliance calls for a considerable allocation of human resources. A firm may allocate the tasks that GDPR compliance brings about either to the current employees or to new hires. Reallocating current employees to take care of GDPR

compliance reduces output, while hiring new employees raises salary costs. Either way, GDPR compliance decreases the firm's financial performance.

IT investments. In addition to human resources, achieving and complying with the GDPR imposes needs on IT infrastructure and software. If the current IT infrastructure is insufficient or the current software does not cater to the requirements imposed by the GDPR, the firm needs to invest in IT capital. Similar to human resources, IT investments made to comply with the GDPR incur costs without increasing output. In other words, the obligation to invest in this type of non-productive capital has a negative effect on firm profitability.

Total factor productivity. Implementing new systems and practices may temporarily disrupt or slow down the normal operations of a firm. If a firm uses consumer data intensively, GDPR compliance may require the firm to make substantial changes to its procedures. In other words, GDPR compliance may cause a negative productivity shock or even a lasting productivity drop. A decrease in productivity reduces the firm's output and affects the firm's performance negatively.

Costs of legal or technical services. Many firms lack legal and technical expertise to implement the changes that GDPR compliance necessitates. These firms are likely to purchase legal and technical services from a third party. Even if a firm has expertise, it may be more cost-effective to outsource the design and implementation of the changes that GDPR compliance requires. An increase in service expenses naturally has a negative effect on firm profitability.

Allocative efficiency. In the longer term, the GDPR may affect firms' performance by changing the efficiency of resource allocation among firms and sectors. The capability of the firm to cost-effectively meet the requirements of the GDPR affects its financial performance and, in the end, whether the firm can operate in the given market or industry or not. The GDPR may thus affect market structure and competition. Campbell, Goldfarb and Tucker (2015) propose a theoretical model to examine how regulation regarding the privacy of consumers' data affect the competitive structure of data-intensive industries. They assume that consumers incur a cost when prompted to give consent to use their data. They make this assumption following the EU's Data Protection Directive (95/46/EC) and Privacy and Electronic Communications Directive (2002/58/EC) and its amendment (2009/135/EC), which were active before the GDPR came into force. Their model implies that privacy regulation imposes transaction costs that fall disproportionately on small and new firms, which may deter the entry of new firms. This conclusion is in line with the empirical finding of Jia et al. (2019) that the adverse effects of the GDPR on venture capital investments in emerging technology companies in Europe may block the entry of innovative companies.

The GDPR concerns any firm that owns or processes personal data of EU consumers. In other words, an organization, irrespective of its location, needs to comply with the GDPR if it tracks, collects, stores, uses or analyses the data of citizens and residents of the EU. Consequently, the implementation of the GDPR has generated compliance costs not only to European firms but also to companies outside of the EU dealing with the data of individuals located in the EU area. Hence, the largest US companies, active also in the EU, have been hit hard by the GDRP compliance requirements. The International Association of Privacy Professionals and Ernst & Young suggest that the combined GDPR compliance costs of Fortune 500 firms amount to 7.8 billion US dollars.[6] It seems that it has primarily been the largest US multinational companies and data-centric firms dealing with EU citizen data that have swiftly acted on the GDPR in the United States.[7] Consequently, it seems possible that large US-based firms, particularly those that are data-intensive high-technology companies, have not had a change in their financial performance different from that experienced by their European counterparts. Instead, we expect to observe the greatest performance difference after the GDPR among the groups of European SMEs and their US counterparts.

The compliance costs of the GDPR are likely to differ across industries. The required actions and costs of GDPR compliance are presumably higher the more intensively the firm collects, processes and/or manages consumer data. The financial implications of GDPR compliance are therefore likely to differ across industries. Information and communication as well as finance are examples of sectors where consumer data are used intensively and where we expect the financial impacts of the GDPR to be significant. We also note that as the GDPR brings about an increase in the demand for legal and technical counselling – perhaps even gives birth to new kind of services – these counselling industries can, in fact, gain from the GDPR.

In addition, if there are returns to scale in adopting GDPR compliant systems and processes, the financial effects of GDPR compliance are likely to differ across firms of different sizes. Irrespective of whether becoming GDPR compliant requires the firm to reallocate human resources or hire new employees, invest in IT capital, undergo a drop in total factor productivity, or purchase technical or legal services, the financial impacts show in the firm's profit margin.

[6] See, e.g., https://www.ft.com/content/0d47ffe4-ccb6-11e7-b781-794ce08b24dc, accessed 13 Nov, 2019.

[7] "This GDPR-compliance ambiguity is fostering a lack of urgency in U.S. companies. From small business to large firms, companies are finding it hard enough to adhere to their current security posture and U.S. regulatory requirements. The exceptions are firms who have historically had robust risk and compliance practices (Fortune 500), or companies whose operations are data-centric and handle EU citizen data. In general, companies have weighed the cost of compliance with the potential for realizing a fine, and have so far taken a wait-and-see approach." Source: https://threatpost.com/gdprs-first-150-days-impact-on-the-u-s/138739/)

3. Data

We use the differences-in-differences method to estimate the effect of the GDPR on firms' profit margins. To this end, we need data on firms that are exposed to the GDPR (the treatment group) and firms that are not (the control group). Our treatment group consists of firms in the old EU-15 countries, except Luxembourg (i.e., Austria, Belgium, Denmark, Finland, France, Germany, Great Britain, Greece, Ireland, Italy, the Netherlands, Portugal, Spain, and Sweden), and Norway[8]. As a control group, we use US firms.

We use data on private and listed firms extracted from Bureau van Dijk's Orbis database. The data[9] used in our estimations comprise the financial statement information of close to 267,000 firms (i.e., 794,886 observations) active in all non-agricultural sectors during 2014 - 2018 in the treatment and control groups. The dependent variable of the estimated equations is the firm's profitability measured by its profit margin, or EBIT divided by turnover. EBIT is calculated by subtracting the firms' costs, including raw material and total operating expenses (incl. wages), from its turnover.

We use firms' size, age, and industry, as defined by the NACE 2 Rev. at the 2-digit level, as control variables. We divide firms into size categories according to the number of employees as follows: firms with 249 employees or less (size category 1, i.e., small and medium-sized firms), firms with 250 to 999 employees (size category 2, i.e., large firms), and firms with at least 1000 employees (size category 3, i.e., very large firms). We also divide firms into age categories as follows: 5 years old or less (age category 1), 6 to 9 years old (age category 2), 10 to 14 years old (age category 3), 15 to 24 years old (age category 4), 25 to 49 years old (age category 5), and firms at least 50 years old (age category 6).

Table 1a shows that profit margins are, compared by the mean and the median, considerably lower among European firms than among their US counterparts. The mean (median) profit margin among European firms is 7% (4%), while the mean (median) profit margin among US firms is as high as 19% (18%). There are likely to be several supply- and demand-side factors that build up this difference between European and US firms, including various regulatory differences between the two continents.

[8] The GDPR is applicable throughout the Internal Market, including the EEA EFTA States of Iceland, Liechtenstein and Norway.

[9] The observations with the most extreme profit margins have been excluded from this sample. More precisely, firms in the profit margin distribution below the 1st percentile (profit margin lower than -37.6%) and above the 99th percentile (profit margin higher than 89%) have been excluded from this sample.

TABLE 1a HERE

The European and US firms also differ considerably in their operating revenues, as Table 1b demonstrates. The mean operating revenue of US firms is approximately 15 times greater than the mean profit margin of European firms. However, the median operating revenues of the two groups are not so different: the median operating revenue of a European firm is approximately 8 701 million US dollars, while the median operating revenue of a US firm is approximately 11 120 million US dollars. This implies that the control group includes firms with very high operating revenue. Table 1c shows that European firms are also smaller when measured by the number of employees. A high share reaching 87% of European firms employ fewer than 250 employees, and only 3% have at least 1000 employees. Among the US firms, by contrast, 72% employ less than 250 employees, and 17% of the firms employ at least 1000 employees.

TABLE 1b AND TABLE 1c HERE

The firms in the treatment and control groups also differ by age. Table 1d shows that the European firms are, on average, considerably younger than the firms in the control group. There are relatively more European firms in each age category except the category of the most mature firms at least 50 years old. Only 8% of the European firms are in this category, while 60% of the US firms are at least 50 years old. Table 1e summarizes the industries in which the firms operate, as defined at the two-digit level by the NACE Rev. 2 classification. It shows that not all industries are equally represented in the treatment and the control groups.

TABLE 1d AND 1e HERE

Using the differences-in-differences method requires that we are confident in assuming that the European and the US firms have parallel trends in their profit margins before the GDPR is introduced and that the trends would have remained parallel had not the regulatory change taken place. However, the firm characteristics, as discussed above, suggest that the European firms are rather different from the US firms. Hence, a natural concern is that the two different groups of firms may have different trends in profit margins. In the next section, we explain how the imbalance between the treatment and control groups is controlled for.

4. Empirical analysis

4.1 Econometric modelling

We use the differences-in-differences (DID) method to evaluate the effect of the GDPR on European firms' profitability or, more precisely, profit margin. As a control group, we use a sample of US companies obtained by CEM, as discussed below. There is some rather strong evidence that, in general, companies did not begin to act on the requirements of the GDPR until 2018, the year of the GDPR's enforceability (see, e.g., Jia et al., 2019, p. 11). We consequently used only 2018 as our post-treatment time. However, it is possible that a large share of firms were not yet GDPR compliant by the end of 2018, as various surveys conducted concerning the implementation of the GDPR suggest. [10] Nevertheless, as the GDPR was enacted in 2016, it is likely that some firms began preparing for the new privacy regulation in 2017. Therefore, we exclude 2017 from our analysis and use the years from 2014 to 2016 as the pre-GDPR years. That is, the estimated effect of the GDPR on European firms' profit margin in 2018 does not fully capture the (sunk) compliance costs of the GDPR.

The difference-in-difference method relies on the assumption that the European firms' profit margins developed similarly to that of the US firms before the "treatment" and that they would have developed similarly had not the GDPR been implemented. We acknowledge that this is a strong assumption because the economic environments of European and US firms are somewhat different. Moreover, as the data description in the previous section reveals, the European and US firms in the sample vary in their characteristics. Hence, a natural concern is whether the parallel trends assumption is realistic.

To reduce the imbalance between the two groups, i.e., to obtain treatment and control groups that are roughly similar in their characteristics, we prepare the data using coarsened exact matching (henceforth CEM, Iacus, King and Porro (2011, 2012)). CEM temporarily coarsens each variable into groups, matches the coarsened data, and finally retains the original values of the matched data. The variables we use in CEM are firms' operating revenue, size as measured by the number of employees, age, and industry. The variable on firm employees is coarsened into four categories: 249 employees or fewer (size category 1, i.e., small and medium-sized firms), 250 to 999 employees (size category 2, i.e., large firms), and at least 1000 employees (size category 3, i.e., very large firms). Firm age is coarsened into categories as follows: 5 years old or less (age category 1), 6 to 9 years old (age category 2), 10 to 14 years old (age category 3), 15 to 24 years old (age category 4), 25 to 49 years old (age category 5),

[10] See, e.g., https://datacenterfrontier.com/50-percent-of-firms-still-not-gdpr-compliant-how-about-your-data-center/. Accessed Jan 29, 2020.

and at least 50 years old (age category 6). The industry identifiers are coarsened into fourteen categories[11]. The data are matched annually and without replacement for the years 2014-2016 (i.e., for the sample years prior to the implementation of the GDPR). The CEM weights[12] are used in the estimation of the following difference-in-differences model:

$$Profit\_margin_{it} = \alpha_i + \delta_0 \times treated_{it} + \delta_1 \times post_{it} + \delta_2 \times treated\_post_{it}$$
$$+ \sum_j \beta_{0_j} \times 1[size_{it} = j] + \sum_k \beta_{1_k} \times 1[age_{it} = k] + \sum_l \beta_{2_l} \times 1[industry_{it} = l]$$
$$+ \sum_m \beta_{3_m} \times 1[country_{it} = m] + \sum_n \beta_{4_n} \times 1[year_{it} = n] + \varepsilon_{it}$$

where i denotes firm and t denotes time, α_i accounts for firm fixed effects, $treated_{it}$ is a dummy for the European countries, $post_{it}$ is a dummy for the year 2018, $treated\_post_{it}$ is a dummy variable that takes the value 1 for the European countries in 2018, 0 otherwise, and ε_{it} is a clustered standard error. We further control for firm size (by size group), age (by age group), industry (at 2-digit level), country and calendar year by adding respective dummies to the model. We estimate the model with firm fixed effects and clustered standard errors to allow for arbitrary correlation within observation units. The form of the differences-in-differences function is linear. The motivation for this choice is the very short time frame we have in this study.

The model is estimated first for all (non-agricultural) industries. We also estimate a variant of the model where the $treated\_post_{it}$ dummy is replaced by the product of the $treated\_post_{it}$ dummy and the share of data professionals in the sector in 2016.[13] We use the share of data professionals as a proxy for the relevance of the GDPR to the sector to allow for heterogeneous treatment intensity across sectors. We then estimate the model separately for firms in different size categories (firms with 249 employees or less, i.e., small and medium size firms; firms with 250 to 999 employees, i.e., large firms; and firms with at least 999 employees, i.e., large firms). The motivation for this is that the costs of GDPR compliance may vary by firm size.

[11] The industry groups were defined as follows (2-digit NACE 2 Rev. codes in parentheses): Manufacturing (10-33), Utilities: (35-39), Construction (41-43), Retail and Wholesale (45-47), Transport (49-53), Accommodation and food services (55-56), Information & communication (58-63), Finance (64-66), Real estate (68), Professional services (69-75), Administrative services (76-82), Public administration (84), Education (85), and Health (86-88).

[12] The weight for each non-treated firm is calculated as the product of the total number of matched control group firms in relation to the total number of matched treated firms and the number of the treated firms in the stratum in relation to the number of control group firms in the firm's stratum (i.e., a group of similar firms with respect to selected coarsened observable characteristics). Unmatched units are removed from the sample.

[13] The data on the share of data professionals are obtained from "The European Data Market", Final report, February 2017.

The financial effects of the GDPR are likely to differ across sectors. We are particularly interested in how the implementation of the GDPR affected data-intensive firms. We therefore estimate the differences-in-differences model for the most data-intensive industries comprising information and communication (or ISIC Rev 2. 58-63) and financial and insurance services (or ISIC Rev 2. 64-66). We also estimate the impact of the GDPR separately for i) firms in the information and communications sector, ii) banks and iii) other financial institutions. There are various reasons for this. First, financial companies hold and manage sensitive personal data. On the one hand, this means that the data privacy regulation concerns their core business and may have a significant impact on them. On the other hand, due to the sensitivity of the data held by banks and insurance companies, financial institutions have a history of dealing with data risk management and data protection, which may reduce their GDPR compliance costs compared to other data-intensive companies. Furthermore, during our "post-GDPR" period, there were other regulatory reforms in the EU concerning financial companies that were likely to generate compliance costs for European banks and insurance companies. The Markets in Financial Instruments Directive (MiFID II, (EU) 2014/65) and the Markets in Financial Instruments Regulation (MiFIR, (EU) No 600/2014), whose purpose is to improve investor protection and increase the transparency of trading activity, came into force in January 2018. The Second Payment Services Directive (PSD2, (EU) 2015/2366) also came into force in January 2018. The objective of the Directive is to extend the scope of regulation to the various types of payment services and to update payment services regulation in line with market developments. The Insurance Distribution Directive (IDD, (EU) 2016/97) became applicable in October 2018 to replace the Insurance Mediation Directive. The aim of the IDD is to harmonize national provisions on insurance and reinsurance. Consequently, the "Post-GDPR" effect for financial and insurance companies comprises the total effect of the GDPR and the other regulatory reforms that took place in the European finance and insurance industry in 2018. Finally, we estimate the impact of the GDPR separately for firms in the i) information and communications sector and ii) banking industry and other financial institutions by firm size category: 1) small and medium size, 2) large and 3) very large firms.

4.2 Empirical findings

We first estimate the differences-in-differences model for the profit margin in the total sample, including companies in all (non-agricultural) sectors. The estimated coefficient on the variable "Treated post-GDPR" is not statistically significant (see Table 2a). Similarly, for the variant of the estimation model where the "Treated post-GDPR" dummy is replaced by the product of the "Treated post-GDPR" dummy and the share of data professionals

in the sector in 2016, the estimated coefficient on the variable "Treated post-GDPR * data intensity weight" is not statistically significant (see Table 2a). The estimation results of Table 2b show the short-term impacts of the GDPR for different firm size categories. We do not find any statistically significant differences in the pre- vs. post-GDPR profit margin developments between the European and US companies in the different size groups.

TABLE 2a-b HERE

We next estimate the financial impact of the GDPR for firms active in the most data-intensive industries, i.e., the information and communications sector, banks and other financial services. The estimated coefficient on the variable "Treated post-GDPR" for data-intensive firms is approximately -1.9 and statistically significant (see Table 3a). We also estimate the model allowing for different treatment effects for firms in the information and communications sector, banks and other financial services. The estimated coefficients vary between -1.7 and -2.1 and are statistically significant (see Table 3a). We further undertake a t-test for the equality of the estimated coefficients on the variable "Treated post-GDPR" for the firms in the three different industries. The t-test does not indicate any statistically significant difference in the estimated coefficients. Note that the sample firms' profit margins have increased during the post-GDPR time period, as the positive and statistically significant coefficient on the "Post-GDPR" variable indicates. Consequently, the estimated negative coefficients on the post-GDPR variables for the treated firms suggest that the European companies' profit margins increased less than those of their US counterparts during the first year of GDPR implementation. In other words, the post-GDPR profit margins of European data-intensive companies increased by approximately 1.7 to 2.1 percentage points less than those of their US counterparts.

TABLE 3a-b HERE

Finally, we estimate the effects of the GDPR on firms in the three data-intensive industries in different size categories. The estimation results further suggest that the post-GDPR profit margins of European data-intensive SMEs increased by approximately 2.8 to 3.4 percentage points less than those of their US counterparts (see Table 3b). The estimated coefficients on the variable "Treated post-GDPR" are slightly greater, though the difference is not statistically significant according to the t-test, for firms active in the financial and insurance services sectors than for firms in the information and communications sector. Additionally, large European banks' profit margins grew less during our post-GDPR period than the profit margins of the large US banks. These findings are as expected due to the other regulatory reforms that took place in the banking and other financial services industries in 2018. The post-GDPR profit margin developments among the very large data-intensive European and US companies are not statistically different from one another. This finding is not surprising given that very large US firms are likely to operate in the EU and are, hence, also affected by the GDPR.

The robustness of the estimation results is checked first by estimating the difference-in-difference model for the industrial sectors in which the companies tend to not use data intensively. In 2016, the employment share of data professionals in the manufacturing sector was only approximately 2%, while among firms in the finance and information and communications sector, it was approximately 10% (IDC, 2017). The estimation results are presented in Annex 1. The estimated coefficients on "Treated post-GDPR" are not statistically significant except for the textile industry, where the coefficient is positive, i.e., the sign opposite that expected.

As a second robustness check, we estimate the difference-in-difference model for data-intensive industries using actual pre-treatment years, 2013 to 2015, both as pre-treatment (2013-2014) and (assumed) post-treatment (2015) years. We also estimate this model separately for the financial sector and for firms in the non-bank financial sector and banking firms. The estimation results are presented in Annex 2. We further estimate the model using 2013-2015 data by firm size category for data-intensive industries (see Annex 3). None of these robustness checks yield statistically significant coefficients on the "Treated post-GDPR" variable.

5. Conclusion

This paper contributes to the rather sparse empirical literature exploring firm- and industry-level impacts of the GDPR, which has tightened the privacy and data regulation of organizations dealing with European customer data. Our extensive firm-level data for the years 2014-2018 show that there is no statistically significant difference in the development of the European and US firms' pre- and post-GDPR profit margins in the total sample comprising all (non-agricultural) sectors. This is a rather expected finding because the total sample includes various sectors (e.g., manufacturing) in which companies do not typically use individual-level customer data. Instead, our data suggest that the profit margin developments among European data-intensive firms during the first year of the implementation of the GDPR were inferior to those of their US counterparts. Our empirical findings indicate that the profit margins of the European data-intensive firms increased by approximately 1.7 to 3.4 percentage points less than the profit margins among the data-intensive US firms.

The European data-intensive SMEs were the most disadvantaged group regarding their post-GDPR profit developments compared to their US counterparts, while the large European data-intensive companies' short-term post-GDPR profit margins dropped relatively less. We do not find any statistically significant difference in the profit margin developments between the very large European and US companies. This empirical finding of negligible differences in the development of the pre- and post-GDPR profit margins between the very large data-

intensive companies in the US and Europe does not indicate anything about the order of magnitude of the very large European companies' GDPR compliance costs but shows merely that they were not statistically different from those of their US counterparts. This empirical finding is consistent with the view that the very large, multinational US companies that often have European customers and deal with the personal data of EU citizens also faced substantial costs when the GDPR was implemented.

Our short-term empirical investigation assesses the financial effects of the GDPR during the first year of its implementation. The anecdotal evidence and various surveys undertaken before and after the GDPR came into force in May 2018 indicate that the whole year of 2018 was still a transition time during which many firms began to invest in becoming GDPR compliant. It further seems that there was a relatively large number of firms that were not fully GDPR compliant by the end of 2018. Moreover, the organizations' costs associated with initially achieving GDPR compliance include sunk investments for compliance and likely differ from the annual costs of maintaining compliance over a longer time period. Moreover, investments made concerning the GDPR may have positive impacts on firms' performance over the longer term, e.g., because they may require firms to implement privacy-friendly technologies and practices and make investments in ICT that improve firm efficiency.

We restricted our analysis to the financial burden that tightened the European data privacy regulation applied to European companies compared to their US counterparts. Our firm-level analysis concerned overall differences in the profitability of companies before and after the implementation of the GDPR without explicitly considering the role of and the financial burden arising from different types of costs (e.g., costs from human resources, IT investments, service purchases) that achieving and maintaining GDPR compliance give rise to. A more detailed analysis of the heterogeneous effects of the GDPR on different firms within and across industries would be welcome to shed light on the short- and long-term impacts of the GDPR. These questions remain for future research.

Literature

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

Goldberg, S., Johnson, G. and Shriver, S. (2019). Regulating privacy online: The early impact of the GDPR on European web traffic & e-commerce outcomes. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3421731

Iacus, S., King, G. and Porro, G. (2011). Multivariate Matching Methods That are Monotonic Imbalance Bounding. *Journal of the American Statistical Association*, 106(493), 345-361.

Iacus, S., King, G. and Porro, G. (2012). Causal Inference without Balance Checking: Coarsened Exact Matching. *Political Analysis* 20(1), 1-24. doi:10.1093/pan/mpr013

IDC (2017). *European Data Market*. SMART 2013/0063. Final report.

Jia, J., Jin, G.Z. and Wagman, L. (2018). The short-run effects of GDPR on technology venture investmenet. NBER, Working Paper 25248.

Yan, B. and Li, J. (2019). The policy effect of the General Data Protectoin Regulation (GDPR) on the digital public health sector in the European Union: An empirical investigation. *International Journal of Environmental and Public Health* 16, 1-15.

Table 1a: Profit margin

treated	mean	p50	sd
0	18.777	17.87	16.778
1	6.716	3.91	12.139

Table 1b: Operating revenue (thousand US dollars)

treated	mean	p50	sd
0	1240000	11120	9200000
1	82919.13	8701	878000

Table 1c: Size category

	Treated		
Size category	0	1	Total
1 (≤ 249 empl.)	40923	976878	1017801
	72.24	87.10	86.38
2 (250 … 999)	5925	107136	113061
	10.46	9.55	9.60
3 (≥ 1000 empl.)	9803	37561	47364
	17.30	3.35	4.02
Total	56651	1121575	1178226
	100.00	100.00	100.00

The first row has *frequencies,* and the second row has *column percentages*

Table 1d: Age category

	Treated		
Age category	0	1	Total
1 (≤ 5 years)	2131	134060	136191
	3.76	11.95	11.56
2 (6 … 9)	2070	120976	123046
	3.65	10.79	10.44
3 (10 … 14)	3285	156531	159816
	5.80	13.96	13.56
4 (15 … 24)	5549	299612	305161
	9.80	26.71	25.90
5 (25 … 49)	9415	325682	335097
	16.62	29.04	28.44
6 (≥ 50 years)	34201	84714	118915
	60.37	7.55	10.09
Total	56651	1121575	1178226
	100.00	100.00	100.00

The first row has *frequencies,* and the second row has *column percentages*

Table 1e: Industry, 2-digit NACE Rev. 2

Industry	Treated 0	1	Total
1	33	12477	12510
	0.06	1.11	1.06
2	5	2391	2396
	0.01	0.21	0.20
3	0	2611	2611
	0.00	0.23	0.22
5	60	56	116
	0.11	0.00	0.01
6	280	495	775
	0.49	0.04	0.07
7	19	112	131
	0.03	0.01	0.01
8	40	3678	3718
	0.07	0.33	0.32
9	139	840	979
	0.25	0.07	0.08
10	235	30546	30781
	0.41	2.72	2.61
11	81	5462	5543
	0.14	0.49	0.47
12	24	228	252
	0.04	0.02	0.02
13	38	6861	6899
	0.07	0.61	0.59
14	57	5784	5841
	0.10	0.52	0.50
15	41	5238	5279
	0.07	0.47	0.45
16	43	8182	8225
	0.08	0.73	0.70
17	89	6131	6220
	0.16	0.55	0.53
18	103	7081	7184
	0.18	0.63	0.61
19	103	782	885
	0.18	0.07	0.08
20	422	13196	13618
	0.74	1.18	1.16
21	372	3666	4038
	0.66	0.33	0.34
22	89	15242	15331
	0.16	1.36	1.30
23	58	11237	11295
	0.10	1.00	0.96
24	189	7140	7329
	0.33	0.64	0.62
25	158	43524	43682
	0.28	3.88	3.71
26	1255	8492	9747
	2.22	0.76	0.83
27	138	9037	9175
	0.24	0.81	0.78
28	444	26105	26549
	0.78	2.33	2.25
29	234	7531	7765
	0.41	0.67	0.66
30	137	3544	3681
	0.24	0.32	0.31
31	97	6385	6482
	0.17	0.57	0.55
32	312	8478	8790
	0.55	0.76	0.75
33	6	10522	10528
	0.01	0.94	0.89
35	593	10698	11291
	1.05	0.95	0.96
36	46	2232	2278
	0.08	0.20	0.19
37	4	1084	1088
	0.01	0.10	0.09
38	48	8491	8539
	0.08	0.76	0.72
39	5	612	617
	0.01	0.05	0.05
41	126	35465	35591
	0.22	3.16	3.02
42	89	11033	11122
	0.16	0.98	0.94
43	42	66161	66203
	0.07	5.90	5.62
45	81	41649	41730
	0.14	3.71	3.54
46	474	145143	145617

	0.84	12.94	12.36
47	733	58226	58959
	1.29	5.19	5.00
49	214	41227	41441
	0.38	3.68	3.52
50	30	2940	2970
	0.05	0.26	0.25
51	87	1162	1249
	0.15	0.10	0.11
52	39	20760	20799
	0.07	1.85	1.77
53	5	874	879
	0.01	0.08	0.07
55	112	22842	22954
	0.20	2.04	1.95
56	263	23722	23985
	0.46	2.12	2.04
58	571	7633	8204
	1.01	0.68	0.70
59	37	4651	4688
	0.07	0.41	0.40
60	151	1146	1297
	0.27	0.10	0.11
61	208	4655	4863
	0.37	0.42	0.41
62	436	28078	28514
	0.77	2.50	2.42
63	112	4587	4699
	0.20	0.41	0.40
64	44283	34367	78650
	78.17	3.06	6.68
65	12	1478	1490
	0.02	0.13	0.13
66	285	9865	10150
	0.50	0.88	0.86
68	264	32274	32538
	0.47	2.88	2.76
69	28	11866	11894
	0.05	1.06	1.01
70	104	36210	36314
	0.18	3.23	3.08
71	64	22655	22719
	0.11	2.02	1.93
72	93	3646	3739
	0.16	0.33	0.32
73	61	9547	9608
	0.11	0.85	0.82
74	231	7523	7754
	0.41	0.67	0.66
75	1	624	625
	0.00	0.06	0.05
77	94	9363	9457
	0.17	0.83	0.80
78	91	10149	10240
	0.16	0.90	0.87
79	56	4894	4950
	0.10	0.44	0.42
80	4	3275	3279
	0.01	0.29	0.28
81	25	12988	13013
	0.04	1.16	1.10
82	57	20705	20762
	0.10	1.85	1.76
84	0	869	869
	0.00	0.08	0.07
85	83	11569	11652
	0.15	1.03	0.99
86	244	18700	18944
	0.43	1.67	1.61
87	10	9877	9887
	0.02	0.88	0.84
88	13	8710	8723
	0.02	0.78	0.74
90	0	2169	2169
	0.00	0.19	0.18
91	0	737	737
	0.00	0.07	0.06
92	4	2724	2728
	0.01	0.24	0.23
93	104	6859	6963
	0.18	0.61	0.59
94	583	1453	2036
	1.03	0.13	0.17
95	0	1250	1250
	0.00	0.11	0.11
96	44	6991	7035
	0.08	0.62	0.60
97	0	51	51
	0.00	0.00	0.00
98	0	32	32
	0.00	0.00	0.00
99	6	30	36
	0.01	0.00	0.00
Total	56651	112157	117822
		5	6
	100.00	100.00	100.00

The first row has *frequencies,* and the second row has *column percentages*

Table 2a. Firms in all industries

	(1) Profit margin (%)	(2) Profit margin (%)
Post-GDPR	0.118	0.390
	(0.11)	(1.54)
Treated post-GDPR	0.531	
	(0.50)	
Treated post-GDPR * data intensity weights		0.768
		(1.84)
Observations	794886	764464
R^2	0.005	0.005
Industry fixed effects	No	No
Country fixed effects	No	No
Year fixed effects	Yes	Yes

z statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2b. Firms in all industries; by size category

	(1) SME	(2) Large	(3) Very large
Post-GDPR	-0.316	-0.959	0.288
	(-0.18)	(-0.72)	(0.57)
Treated post-GDPR	0.853	1.533	-0.100
	(0.50)	(1.04)	(-0.19)
Observations	681971	78108	34807
R^2	0.001	0.007	0.001
Industry fixed effects	No	No	No
Country fixed effects	No	No	No
Year fixed effects	Yes	Yes	Yes

z statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3a. Data-intensive firms

	(1) Profit margin (%)	(2) Profit margin (%)
Post-GDPR	3.014***	3.014***
	(4.18)	(4.18)
Treated post-GDPR	-1.876**	
	(-2.78)	
Treated post-GDPR * info&communication		-1.699*
		(-2.50)
Treated post-GDPR * bank		-2.054**
		(-2.90)
Treated post-GDPR * other financial inst.		-2.134**
		(-2.76)
Observations	107705	107705
R^2	0.006	0.006
Industry fixed effects	No	No
Country fixed effects	No	No
Year fixed effects	Yes	Yes

z statistics in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3b. Data-intensive firms; by size category

	(1) SME	(2) Large	(3) Very large
Post-GDPR	4.186***	3.131***	1.678*
	(3.87)	(4.25)	(2.45)
Treated post-GDPR x info&communic.	-2.823**	-1.467	-0.829
	(-2.85)	(-1.71)	(-0.89)
Treated post-GDPR * bank	-3.430***	-1.893*	-0.897
	(-3.30)	(-2.35)	(-1.10)
Treated post-GDPR * other financial inst.	-3.343**	-1.644	-1.909
	(-3.13)	(-1.32)	(-0.68)
Observations	87601	13427	6677
R^2	0.007	0.011	0.006
Industry fixed effects	No	No	No
Country fixed effects	No	No	No
Year fixed effects	Yes	Yes	Yes

z statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Annex 1. Robustness test: Manufacturing industries

	(1) food	(2) textile	(3) wood&paper	(4) chemicals	(5) metal	(6) construction
Post-GDPR	3.190	-2.411*	1.793	0.477	0.170	-2.625
	(0.71)	(-2.15)	(0.43)	(0.06)	(0.06)	(-0.94)
Treated post-GDPR	-2.923	2.280*	-0.689	-0.022	0.892	3.107
	(-0.64)	(2.03)	(-0.16)	(-0.00)	(0.33)	(1.16)
Observations	27706	13817	11040	14121	38969	60935
R^2	0.011	0.006	0.019	0.014	0.007	0.011
Industry fixed effects	No	No	No	No	No	No
Country fixed effects	No	No	No	No	No	No
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

z statistics in parentheses
$^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

The industry groups used in the estimations were defined as follows (2-digit NACE 2 Rev. codes in parentheses): food (10-11), textile (13-15), wood&paper (16-17), chemicals (19-21), metal (24-25) and construction (41-43).

Annex 2. Robustness test: Pre-GDRR years: 2013-2014, Post-GDPR: 2015

Data intensive firms

	(1) Profit margin (%)	(2) Profit margin (%)
Post-GDPR	-0.322	-0.657
	(-0.19)	(-0.35)
Treated post-GDPR * info&communic.		-0.102
		(-0.36)
Treated post-GDPR * bank		1.053
		(0.56)
Treated post-GDPR * other financial instit.		1.252
		(0.66)
Treated post-GDPR	1.389	
	(0.78)	
Observations	82623	57302
R^2	0.004	0.004
Industry fixed effects	No	No
Country fixed effects	No	No
Year fixed effects	Yes	Yes

z statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Annex 3. Robustness test: Pre-GDRR years: 2013-2014, Post-GDPR: 2015

Data intensive firms; by size category

	(1) SME	(2) Large	(3) Very large
Post-GDPR	-0.605	1.488*	-0.611
	(-0.24)	(2.13)	(-0.69)
Treated post-GDPR	1.719	-0.446	1.601
	(0.67)	(-0.68)	(1.81)
Observations	67774	9994	4855
R^2	0.004	0.006	0.008
Industry fixed effects	No	No	No
Country fixed effects	No	No	No
Year fixed effects	Yes	Yes	Yes

z statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Elinkeinoelämän tutkimuslaitos

The Research Institute of the Finnish Economy

ISSN-L 2323-2420
ISSN 2323-2420 (print)
ISSN 2323-2439 (pdf)

Publisher: Taloustieto Oy

Tel. +358-9-609 900
www.etla.fi
firstname.lastname@etla.fi

Arkadiankatu 23 B
FIN-00100 Helsinki