PENDING REVIEW – submitted but not yet marked.

Master Thesis

The distributive impact of the GDPR on firm performance

Thi Huyen Vy Vo

Master of Public Policy

Year of Graduation: 2025

Word Count 11,170

Supervisor

Prof. Joanna Bryson, PhD

Executive Summary

This thesis analyses the economic impact of the General Data Protection Regulation (GDPR) on firms in data-driven industries in Germany, with a particular focus on revenues and profit margins across different firm sizes. While prior studies predominantly highlighted the GDPR's compliance costs, especially for small and medium-sized enterprises (SMEs), this study provides new evidence suggesting a more nuanced and positive outcome, particularly for smaller firms.

Using firm-level panel data from the DAFNE database from 2011 to 2023, I apply a Difference-in-Differences (DiD) approach developed by Callaway and Sant'Anna (2020) to estimate dynamic treatment effects. To strengthen causal inference, I further employ entropy balancing and multiple imputation techniques to address issues of missing data and ensure comparability between treated and control groups.

The results indicate that the GDPR did not negatively impact firm performance in data-driven industries. Instead, firms - especially small ones - experienced significant gains in profitability following the regulation's enforcement. Larger firms showed insignificant effects, implying that smaller firms may have leveraged GDPR compliance efforts more effectively to enhance their operational performance, challenging prior indications that data privacy regulation always disproportionately harms smaller firms. Rather, the GDPR's harmonization of privacy standards across EU member states appears to have reduced legal uncertainty and created a more level competitive playing field.

Policy implications drawn from the study emphasize that well-designed regulatory frameworks, like the GDPR, can coexist with firm-level economic resilience and even drive modernization. Harmonization efforts that reduce legal complexity may be especially beneficial for SMEs, enhancing their ability to operate across borders within the EU's Digital Single Market.

Overall, this thesis contributes to a more differentiated understanding of the GDPR's economic impact, demonstrating that privacy regulation, if thoughtfully designed, can support both individual rights and firm-level competitiveness.

List of Abbreviations

AI	Artificial Intelligence
ATT	Average Treatment Effect on the Treated
BDSG	Bundesdatenschutzgesetz (German Federal Data Protection Act)
BvD	Bureau van Dijk
CIS	Community Innovation Survey
COVID-19	Coronavirus Disease 2019
DAFNE	Database for German and Austrian Firms
DiD	Difference-in-Differences
DMA	Digital Markets Act
DSA	Digital Service Act
DSM	Digital Single Market
DPA	Data Protection Authority
GDPR	General Data Protection Regulation
HGB	Handelsgesetzbuch (German Commercial Code)
ICT	Information and Communication Technology
IHS	Inverse Hyperbolic Sine
MAR	Missing at Random
MCAR	Missing Completely at Random
MICE	Multiple Imputation by Chained Equations
OECD	Organisation for Economic Co-operation and Development
PMM	Predictive Mean Matching
R&D	Research and Development
SME	Small and Medium-Sized Enterprises
SUTVA	Stable Unit Treatment Value Assumption
TWFE	Two-Way Fixed Effects

Table of Contents

I. Introduction

The regulation of the digital economy has become a central issue in global economic and political debates. This discussion intensified when the Trump administration pursued a deregulatory agenda aimed at loosening constraints on the tech sector (Kerr and Bhuiyan, 2025). Yet, competing visions of digital governance have existed for some time. In this regard, Anu Bradford (2023) identifies three competing regulatory approaches: the American market-driven approach, the Chinese state-driven model, and the EU's rights-driven regulatory model. While the U.S. emphasizes innovation through minimal government interference, and China aligns technological development with political control, the EU has positioned itself as a normative power, prioritizing democratic values and fundamental rights—particularly data privacy.

This divergence in regulation has led to controversy over the EU's role in shaping the digital economy. Critics argue that the EU's regulatory framework, typified by the General Data Protection Regulation (GDPR), imposes high compliance costs, stifles innovation, and undermines competitiveness, especially compared to U.S. and Chinese tech firms. Europe is often portrayed as lagging in the global race for digital leadership, burdened by bureaucracy and an overly cautious regulatory stance (Bradford, 2024).

At the same time, proponents of the European model propose that regulation can serve as a catalyst for innovation and a safeguard for consumer rights, arguing that trust, transparency, and accountability are essential foundations for sustainable digital growth. The GDPR, implemented in 2018, has become a global benchmark, influencing data protection laws well beyond Europe's borders and reinforcing the EU's ambition to set global standards through what has been termed the "Brussels Effect" (Bradford, 2024). This regulatory assertiveness has sparked intense academic and policy debates about the economic consequences of prioritizing rights-, and in particular privacy-governance in the digital economy. Understanding whether and how such regulation affects firm behavior and economic performance has thus become an important empirical question, particularly in light of broader discussions about the trade-offs between innovation, competition, and fundamental rights in an increasingly digitized world.

I therefore examine the impact of the GDPR enforcement on firms in data-driven industries in Germany, with a focus on profits and revenues different firm sizes. Using firm-level panel data and a difference-in-differences framework that allows for heterogeneous treatment effects - developed by Callaway and Sant'Anna (2020) -, the analysis evaluates whether GDPR compliance hindered or helped firms in data-intensive industries. Unlike more recent EU regulations (e.g., the Digital Markets Act or AI Act), the GDPR offers several years of post-enforcement data, enabling a rigorous empirical investigation.

The results of my empirical analysis reveal that, contrary to some prior indications, firms in data-driven industries - especially small firms - experienced significant profitability gains after the introduction of the GDPR. While the effects on revenue were smaller, they remained positive. These results challenge the prevalent view that privacy regulation inherently harms business performance and highlight the importance of considering different firm sizes in regulatory impact assessments.

The remainder of the thesis is structured as follows: Section II reviews the relevant literature and theoretical mechanisms; Section III describes the data and methodology; Section IV presents the empirical results; Section V discusses the findings and their policy implications; and Section VI outlines the thesis' limitations

II. GDPR and firm performance

The GDPR is a landmark privacy regulation and has been introduced as a milestone of the EU's regulatory framework to govern the digital economy. At its core, the GDPR defines the use and processing of personal data to protect users' fundamental rights over their personal data in the digital age. For example, the GDPR defines that data protection must be included "by design and by default" which means that the processing of personal data should be minimized and pseudonymized as soon as possible (Article 1(d) and Article 78). With its enforcement in 2018, the EU provided a uniform and coherent data privacy framework within the EU, replacing national data protection legislation that have been in place prior to the GDPR and created legal uncertainty. Hence, a further objective of the GDPR is the harmonization of privacy protection standards across the Member States. This objective follows an economic interest: to ensure the free flow of data within the internal market (Article 13) and promote intra-EU trade.[1]

An impact assessment released prior to the introduction of the GDPR reveals that EU policymakers intended to reduce legislative uncertainty produced by the fragmented and divergent national regulations as this was assessed to be very costly for businesses. The impact assessment estimated that due to the legal fragmentation and the resulting administrative burden EU firms faced "close to € 3 billion per year" (p.11) of additional costs. For example, privacy standards are often centered on users "unambiguously giving" consent. The interpretation, however, in which form consent should be given differed very much across Member States. While in some Member States it was sufficient to "expressly" give consent, other Member States required consent to be given in writing. Less strict interpretation also accepted consent given in form of implied consent. The assessment furthermore acknowledges that larger firms are less burdened by this legal fragmentation as they own the necessary resources and capabilities, e.g. human resources, legal aid, technical capabilities , to conform to different privacy standards. Small and medium-size firms (SMEs) lack these resources however and are therefore in a disadvantaged position. Due to the complexity and uncertainty of these divergent and conflicting national legislations firms were, so the assessment concludes, less inclined to expand

[1] See https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng.

their business activities across other Member States. In this regard, it is argued the GDPR can mitigate the problem by harmonizing privacy regulations and facilitating the free flow of data and intra-EU trade.

A similar argument is provided by Bradford (2024). Bradford argues that EU firms face several key challenges in comparison to their U.S. and Chinese competitors which lead the EU to produce less global tech champions. One of these challenges is the absence of a Digital Single Market (DSM) which limits in particular SMEs as they “cannot draw on economies of scale to grow beyond a certain size” (Bradford, 2024, p.423). While there are also cultural and social factors at play, Bradford states that legal barriers such as divergent national laws across member states serve as an additional barrier. She therefore concludes that through the implementation of the GDPR which harmonizes privacy protection standards, firms could benefit by improved scaling by engaging in cross-border activities.

Niebel (2021) also argues in favor of the GDPR, suggesting it could affect firms, especially small ones, through two channels. First, by addressing concentrated market power in the digital economy, where incumbents gain advantage through superior data access. In this regard, the GDPR’s right to data portability could reduce lock-in effects and level the playing field. Second, drawing on the Porter hypothesis,[2] Niebel contends that the GDPR could drive innovation, particularly GDPR-compliant technologies, by pushing firms to develop new processes, products, and business models that both meet regulations and strengthen competitiveness.

Although policymakers intended to reduce complexity and costs for firms with the introduction of the GDPR, most empirical studies reveal the contrary so far. The literature - although still evolving in this field - increasingly shows that GDPR’s impacts firm performance, competition and innovation in often unintended and negative ways as the regulation imposes various responsibilities for firms. Firms may be affected through two channels (Chen et al., 2022; Jia et al., 2020). First, the GDPR raises compliance costs by requiring firms to revise internal processes, invest in IT systems for data management, and appoint data protection officers. Second, by requiring explicit user consent for data processing, the GDPR may reduce data availability, leading to lower sales for firms reliant on data-driven business models. Especially, the first channel has been highlighted as the main driver to induce costs by various studies (Jia et al., 2020; Chen et al., 2022; Demirer et al., 2024).

Goldberg et al. (2021) document declines in web traffic and e-commerce sales. Peukert et al. (2022) and Johnson et al. (2022) study web technology vendors and find that the GDPR increases market concentration. Jia et al. (2021) show a 26% drop in venture capital deals for EU tech firms post-GDPR.

[2] The Porter Hypothesis, initially proposed for environmental context, suggests that well-designed regulation can stimulate innovation and enhance firms' competitiveness by encouraging efficiency improvements and technological advancement. For a more comprehensive review of the Porter hypothesis for the GDPR see Niebel (2021).

More specific research on compliance costs (Demirer et al., 2024) indicates that the privacy regulation acts as a 20% tax on the cost of data storage for EU firms. Hence, EU firms store less data, perform less computation, and reduce their data-intensiveness when compared to US firms.

Studies that are closest to the topic of this thesis attempt to estimate the broader economic cost of the GDPR, by estimating the impact of the GDPR on firm performance. Koski and Valmari (2021) compare the profit margins of EU- and US-based firms from 2018. While they find no statistically significant result for their full sample of EU firms comprising all sectors, they find that EU firms in data-driven industries experience smaller profit gains of approximately 1.7 to 3.4 percentage points less than their US counterfactual. Similarly, Chen et al. (2022) utilize a sample of firms from different industries and countries to demonstrate that GDPR exposure correlates with an 8% reduction in profits over the period of 2011 to 2020. In their analysis, they furthermore confirm that the main burdens on firms are the additional compliance costs as sales in post-GDPR periods only saw a 2% decrease in their estimation. Both studies furthermore highlight that small firms and firms in data-intensive industries are more negatively affected by the GDPR, whereas large, incumbent firms have largely benefited from their capacity to comply and adapt.

Further studies reinforce the assumption that the GDPR disproportionally burdens small firms, especially in data-driven industries (Bessen et al., 2020). As such, GDPR appears to have discouraged market entry and accelerated exit, particularly for smaller players. Janssen et al. (2022) report that the GDPR lead to a sharp decline in app numbers by causing widespread exits, while simultaneously preventing the launch of both successful and unsuccessful apps. Additionally, research on privacy regulation reveals that consumers are more likely to give consent to larger established firms (Campbell et al., 2015).

While most of these empirical findings highlight the negative economic consequences of the GDPR, some studies reveal a more complex relationship.

For example, Maex (2022) studies the impact of the GDPR on US firms that are affected by the GDPR and the impact on their internal information quality. The study finds evidence that the GDPR indeed improves firms' internal information quality which in turn has a positive impact on their operational efficiency. The effect is more pronounced for high technology firms and small firms. However, the authors also conclude that this improvement cannot be compensated by the overall regulatory burden of the GDPR, leaving the overall net effect for firms to be negative.

A study of a large European tech company found that GDPR-compliant consent procedures increased consumer opt-ins while maintaining caution over sensitive data (Godinho de Matos and Adjerid, 2022). The firm experienced gains in sales, marketing effectiveness, and customer retention, suggesting that GDPR’s consent framework can enhance consumer privacy protection while also improving outcomes for European tech firms.

Lefrere et al. (2022) study the market for content providers and capture a broader set of variables for websites, such as cookies and consent mechanisms. The study finds no evidence that EU content providers diverged from their US counterfactuals. The authors argue that EU content providers were able to achieve compliance without “triggering the undesirable outcomes foreseen by the ad-tech industry” (Lefrere et al., 2022, p.4).

Martin et al. (2019) find through interviews with German startups that the GDPR can have both innovation-stimulating as well as constraining effects. Interestingly, they also observe a shift towards EU digital service providers (e.g. cloud storage) among the respondents, driven by the belief that these providers are inherently compliant with EU regulations.

Finally, Blind et al. (2024) exploit panel data from the German part of the community innovation survey (CIS) and quantitively assess the impact of the regulation on innovation activities. Their findings suggest that the GDPR increases firms’ innovation effort but also induces a shift in the type of innovation that firms are pursuing, away from radical and towards incremental innovation. Furthermore, the results indicate that small firms are particularly affected by the GDPR in their innovation activities. Small and young firms mainly report that the GDPR hindered or complicated their innovation efforts, reflected in a 1.2 percentage point drop in the sales share of radical innovations and a 1.6 percentage point rise in incremental innovations. Nevertheless, some firms also found opportunities, using the GDPR to develop new markets and drive incremental innovation, suggesting that the regulation, while challenging, can also foster adaptation and growth.

Building on mixed theoretical and empirical insights from the literature, I investigate the impact of the GDPR on firm performance, with a particular focus on revenues and profit margins across different firm sizes. The primary research question is whether GDPR enforcement in 2018 led to measurable changes in firm outcomes in data-driven industries, and whether these effects were evenly distributed across firms or concentrated among specific groups.

Given the GDPR’s uniform application across the EU but heterogeneous firm exposure depending on data-intensity, I hypothesize that data-driven firms, and particularly surviving SMEs, may have ultimately benefited from GDPR compliance. These benefits could arise through improved internal processes, enhanced consumer trust, and more seamless access to a harmonized DSM despite the initial compliance burden.

To empirically test this, I employ firm-level panel data from Germany obtained from the DAFNE database and apply the Difference-in-Differences approach developed by Callaway and Sant’Anna (2020), which flexibly accounts for treatment heterogeneity and dynamic treatment effects over time. Additionally, I employ entropy balancing to strengthen the covariate balance between treated and control groups and to reinforce the credibility of the parallel trends assumption.

III. Data

For my analysis I exploit firm-level panel data from the DAFNE database from the years 2011 to 2023. The dataset provides annual financial and structural information on approximately 800,000 private German and Austrian companies. The dataset includes key variables such as revenues, profit margins, total assets, and employee counts, as well as firm-specific characteristics like company size, industry classification, and founding year. The panel nature of the data enables the tracking of firms over time, allowing for both cross-sectional and longitudinal analysis.

DAFNE is the German-specific database provided by the Bureau van Dijk (BvD) and contains more detailed information on firms than other cross-country databases (Beuselinck et al., 2021), such as the BvD' Orbis database which have been utilized by Koski and Valmari (2020) and Chen et al. (2022) to study the impact of the GDPR on firm-performance. While Orbis offers excellent data for cross-country comparisons (which have been implemented by Koski and Valmari, 2020 and Chen et al., 2022), it also comes with notable drawbacks. A working paper published by the OECD (Bajgar et al., 2020), for example, identifies that firms in the Orbis database are "disproportionately larger, older and more productive" (p.3). They therefore conclude that Orbis is less suitable for research designs exploring underperforming firms and distributional dynamics, which is an objective of this thesis. The authors also conclude that imputing missing information can "substantially improve the representativeness of the population of firms with at least 10 employees". This will be further discussed in section 3.2.

3.1 Data description and cleaning

For data cleaning and preprocessing, I follow the procedures outlined in Chen et al. (2022). First, all outliers in the continuous variables - profit margin, revenue, total assets, and number of employees - are removed. This is done to ensure that extreme values, which could influence regression estimates, do not distort the empirical analysis. I then perform inverse hyperbolic sine (IHS) transformation to the variables revenue, total assets, and employee count.[3][4]

After preprocessing, the sample is further refined by restricting the dataset to firms with relevant and complete financial information for the period of interest. Specifically, all firms whose most recent balance sheet date precedes 2019 are excluded, ensuring that the analysis captures firms active around and after the implementation of the GDPR in 2018. This step aligns the dataset with the

[3] This transformation is also used by Chen et al. (2022) and offers several advantages over the natural logarithm: unlike the log transformation, which is undefined for non-positive values, the IHS function maps all real numbers—including zeros and negatives—onto a continuous scale. This is particularly suitable for financial data, where losses or zero entries are not uncommon.

[4] While variables like revenue, assets, and employment benefit from transformation due to their scale and skewness, profit margin is already in interpretable shape and relatively symmetrically distributed around zero. As such, it remains in its original form without transformation.

temporal focus of the empirical strategy and enhances the reliability of outcome comparisons across pre- and post-GDPR periods.

3.2 Missing Data and Diagnostic Analysis

Missing data is a common challenge in empirical research, particularly in financial panel data. Broadly, two main strategies exist to manage missing values in causal inference settings: listwise deletion or imputation. Listwise deletion involves removing any observation with missing values in relevant variables. This method - although very common in empirical research - can produce problematic cases. It reduces the analytical sample size and might produce unprecise estimates. In its worst case, this approach could introduce selection bias, especially if firms that report their financial data systematically differ from firms that do not report (see Bryzgalova et al. 2022). In financial datasets, however, missing values can also occur mechanically, for example, for young firms that lack historic information due to firms being recently established or not legally obligated to report certain information (see Bryzgalova et al. 2022).

In the present dataset, missingness is substantial, particularly for the key outcome variables. Specifically, approximately 94% of profit margin observations and 90% of revenue observations are missing in the raw dataset. Table 1 presents an overview of missingness across all core variables. Although this level of missingness may appear extreme, it is not uncommon in firm-level panel data. For instance, Bryzgalova et al. (2022) document similarly high rates of missingness in the Compustat database, a standard source for corporate financial data. The authors also show that variables are cross-sectionally correlated as they could observe clusters of correlations. Their findings emphasize the importance of carefully analyzing the mechanism behind missingness before choosing an appropriate strategy.

Given the focus of this study on the distributional effects of the GDPR, it is particularly important to examine whether missingness correlates with firm characteristics, such as firm size. Bryzgalova et al. (2022) find that for their database, Compustat, missingness is not systematically associated with firm size, suggesting it may not bias analyses of distributional outcomes. However, this pattern might not hold for the DAFNE database.

Table 1: *Percentage of Missing Values in the key variables*

Percentage of Missing Values by Variable

	Variable	Missing_Percentage
profit_margin	Profit Margin	93.95
revenue_IHS	Revenue	89.59
employees_IHS	Employees	71.52
total_assets_IHS	Total Assets	44.40
firm_age	Firm Age	4.36

To investigate the nature of missingness in this dataset, I follow the approach of Bryzgalova et al. (2022) and estimate logistic regression models, using missingness in the key variables - profit margin, revenue, total assets, and employee count - as binary dependent variables. Each model regresses the missingness indicator on a set of observable firm-level characteristics, such as firm size, firm age, industry, total assets, number of employees, and year fixed effects. The aim is to assess whether missingness is consistent with a Missing Completely at Random (MCAR) assumption, or more realistically, with a Missing at Random (MAR) pattern, in which the probability of missingness depends on observed variables. This diagnosis is also important to determine if imputation can be performed on the dataset.

Table 1 shows the results for the logistic regression models. The results partially confirm my initial speculation that firm size is a significant predictor of missing data. Specifically, small firms exhibit a higher likelihood for missing data on profit margins than larger firms. Medium-sized firms are also more likely to have missing profit data, but less so than small firms. Furthermore, older firms are slightly more likely to have missing profits and less likely to have missing employee data. This pattern suggests that missingness is not random but systematically related to firm characteristics, particularly firm size. Hence, the results reject the MCAR assumption across all four variables and supports the assumption of MAR and therefore, the use of imputation.

Table 2: *Results of the Logistic Regression Model for missing values*

	Profit Margins	Revenues	Employees	Total Assets
Medium Firm	-1.350***	-0.818***	0.417***	-0.362*
	(0.020)	(0.017)	(0.040)	(0.154)
Small Firm	1.281***	0.007	0.251***	-0.167
	(0.013)	(0.009)	(0.035)	(0.103)
Old Firm	0.184***	0.029***	-0.364***	-0.173+
	(0.014)	(0.007)	(0.035)	(0.089)
Young Firm	-0.753***	-0.754***	-0.091**	-0.154
	(0.014)	(0.006)	(0.034)	(0.109)
Employees (IHS)	-0.499***	-0.441***		-0.017
	(0.004)	(0.002)		(0.039)
Total Assets (IHS)	-0.466***	-0.073***	0.085***	
	(0.004)	(0.002)	(0.012)	
Revenue (IHS)			-0.676***	0.071+
			(0.014)	(0.037)
Profit Margin			0.011***	-0.020***
			(0.001)	(0.004)
Year FE	0.045***	-0.008***	-0.308***	-0.614***
	(0.002)	(0.001)	(0.005)	(0.020)
Num.Obs.	845710	845710	68814	59787
AIC	267651.3	916766.8	38786.4	7306.2
BIC	267954.2	917069.7	39024.0	7540.1
Log.Lik.	-133799.662	-458357.404	-19367.176	-3627.092
F	4615.635	3307.744	437.032	42.686
RMSE	0.21	0.43	0.29	0.12

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Logistic Regression: Predictors of Missingness

3.3 Imputing missing firm information

In the following section I impute missing firm information. As mentioned in section 3.2, imputing can enhance firm representativeness in financial panel datasets (see Bajgar et al., 2020).

For imputation I adopt a hybrid approach: I keep observations in my analytical sample where at least one dependent variable (either revenue or profit margin) is available, while dropping all cases where both are missing. This strategy preserves partial outcome information, while avoiding reliance on imputed dependent variables. As a result, approximately 86% of the inital observations are excluded. The resulting estimation sample consists of 483,327 firm-year observations representing 127,967 unique firms.

As shown in Table 1 in the appendix, this new generated dataset exhibits substantially improved data quality compared to the previous dataset. While profit margin has the highest missingness at 55.7% of cases, revenue missingness drops to 27.3%. The explanatory variables - total assets and employee count - still exhibit moderate levels of missingness (15–40%), but are now suitable for imputation under MAR. These diagnostics justify the application of multiple imputation methods to fill in missing covariate data and ensure more efficient use of the available sample in the final analysis.

Following the earlier diagnostics that supported the MAR assumption, I perform multiple imputation by chained equations (MICE) using the *mice* package in R (van Buuren & Groothuis-Oudshoorn, 2011)[5], a widely adopted and flexible method well-suited for panel data. During this process, however, I explicitly exclude missing values in the dependent variables, as imputing outcomes could lead to biased estimates of treatment effects, particularly in causal inference settings. The MICE algorithm iteratively fills in missing values using a fully conditional specification, whereby each missing value is imputed conditionally on all other variables in the dataset.[6]

Imputing missing values, which turn out to be wrong, can also introduce imputation bias. Therefore, I perform diagnostic checks to assess the quality of imputing by comparing the distribution of the imputed data with the initial dataset. Summary statistics in Table 2 in the appendix show that the distributions of imputed values closely mirrored those of the observed data in terms of central tendency and dispersion. For example, the mean of *total_assets* in the imputed dataset was 8.34 compared to 8.25 in the observed data, while the medians were even closer (7.94 vs. 7.87).

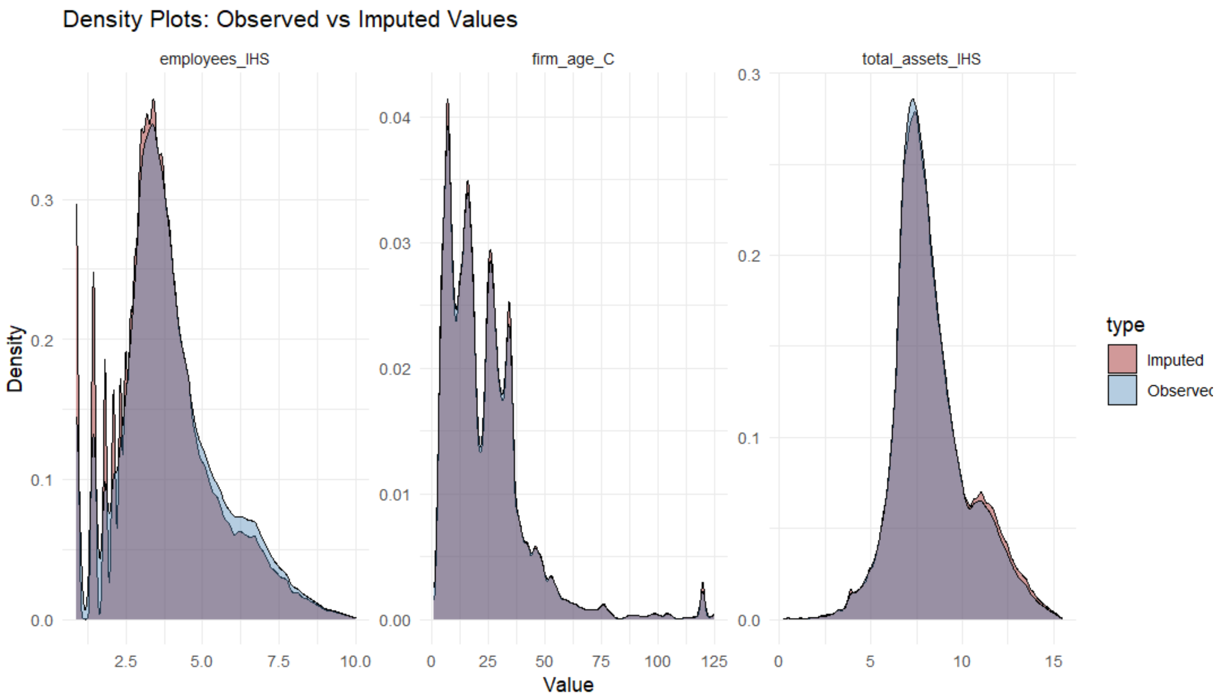
Further validation is provided by density plots in Figure 1, which compare the distributions of observed and imputed values. For most variables the imputed distributions align closely with observed patterns, suggesting that the model successfully captured the underlying data structure. For the employees variable, some deviation is observed in the lower range of the distribution,

[5] I furthermore used the code provided by van Buuren, Stef (2018). *Flexible Imputation of Missing Data*. https://stefvanbuuren.name/fimd/, and adapt it to my research design.

[6] For continuous variables predictive mean matching (PMM) was chosen as the imputation method. PMM is particularly effective in preserving the observed distribution of variables and avoiding unrealistic values, which is important when dealing with skewed financial data. For categorical variables, such as company size, a polytomous logistic regression (polyreg) method was employed. A custom predictor matrix was created using the *quickpred()* function, which selects only those predictors with at least moderate correlation (r ≥ 0.1) with the variable being imputed, balancing model complexity with predictive power. The final imputation settings included five imputed datasets (m = 5), each generated with ten iterations (maxit = 10), consistent with best practices in the literature (see van Buuren & Groothuis-Oudshoorn, 2011).

indicating minor shape distortion introduced by imputation. However, no extreme or implausible values were detected.

Figure 1: *Density plots of imputed and observed values*



Overall, the diagnostics confirm that the multiple imputation procedure preserved the distributional characteristics of the original data and produced plausible imputations, thereby supporting its further analytical use.

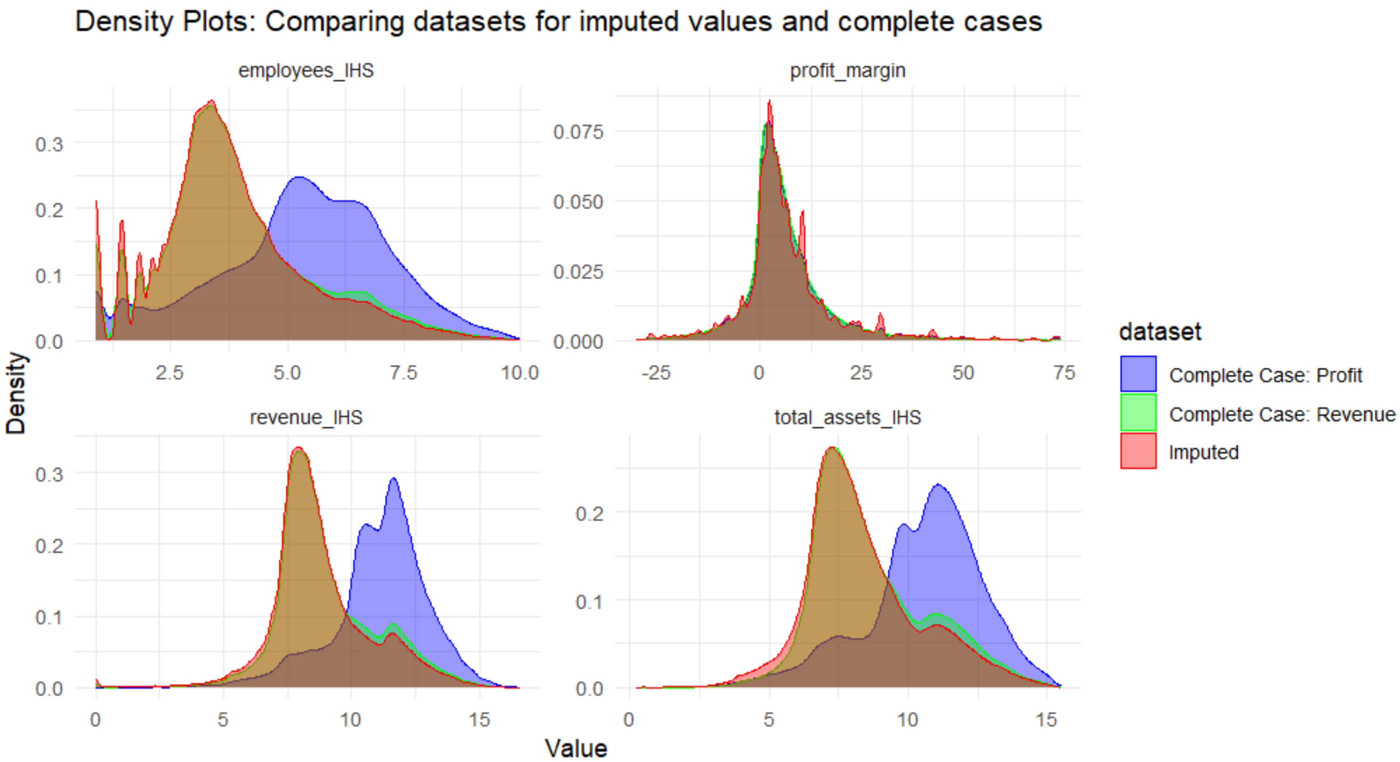
3.4 Comparing imputed and complete case dataset

To assess the potential impact of missing data on sample representativeness and statistical inference, I compare results based on the imputed dataset to those from complete-case analyses. Rather than constructing a unified complete-case sample that requires full data across all variables, I build separate complete-case datasets for the profit and revenue models. This strategy maximizes sample size in each model by keeping observations with partial outcome availability, thereby preserving statistical power.

As outlined in Section 3.2, missingness in profit data is particularly high (~94%). After removing all observations with missing profit margin, total assets, firm age, or employee count, the resulting complete-case sample for profits includes close to 70,000 observations, representing about 19,000 unique firms. In contrast, missingness is less severe for revenues (~90% in the raw data, but substantially lower after initial filtering). Consequently, the revenue-complete sample is larger, containing approximately 245,000 observations and about 80,000 unique firms (see Table 3 in the appendix).

Figure 2 illustrates the density plots for the continuous variables for imputed and complete-case datasets. The figure reveal that the profit-complete sample is skewed toward larger firms. On average, these firms report higher revenues, employ more workers, and hold more total assets than firms in both the imputed dataset and the revenue-complete sample. This suggests that restricting the analysis to firms with complete-case data for profitability would have introduced selection bias and limited the generalizability of the findings. By contrast, the imputed dataset captures a broader and diverse cross-section of the firm population. Descriptive comparisons of the summary statistics for the imputed and the complete-case dataset in Table 4 in the appendix confirm this pattern.

Figure 2: *Comparing density plots across datasets*



These findings suggest that relying exclusively on complete-case samples could bias estimates of GDPR's distributive effects, especially with respect to firm size.

In contrast, the imputed dataset provides a more balanced and comprehensive foundation for analysis. By recovering missing covariate values under the MAR assumption, the imputation process allows for a more inclusive study of GDPR's economic consequences - particularly in identifying distributional impacts across the size spectrum of firms.

To account for any uncertainty and ensure transparency, I also run the regression analysis on the complete cases as this serves as a important benchmark to assess the robustness and sensitivity of the results to different approaches of missing data.

IV. Methodological approach

4.1 Exposure to GDPR

Most existing empirical studies on the economic consequences of the GDPR (see Koski and Valmari 2020; Jia et al. 2021; Aridor et al. 2022) adopt a cross-country research design, comparing outcomes between EU-based firms and firms located in other regions not legally subjected to the GDPR (e.g., the United States or other non-EU countries). While this approach enables identification via geographic variation, it risks introducing endogeneity concerns by conflating the effects of the GDPR with other country-specific factors or economic shocks unrelated to the GDPR (see Frey & Presidente 2024).

Furthermore, there are spillover effects of the GDPR outside of the EU as the regulation also applies to non-EU firms if these target EU citizens. The literature also shows that some non-EU firms even show greater compliance than EU firms (Johnson 2022). In addition, firms that operate globally might also experience additional spillover effect as they adapt the standards of the GDPR for their global processes due to cost-efficiencies, often described by the "Brussels effect" (Bradford, 2012). As such, Peukert et al. (2022) find evidence that the GDPR serves as a global privacy standard for entities, which are legally unaffected by the GDPR.[7] Firms may adapt the GDPR as a global standard for all of their processes due to cost-efficiencies. Therefore, the broad scale and global scope of the GDPR makes the identification of a suitable control group, which is completely unaffected by the privacy regulation, particularly challenging for empirical researchers (see Johnson 2022 for a more detailed review of the challenges to empirically assess the economic consequences of the GDPR).

However, there also exist challenges in identifying relevant units for the treatment group as firms' compliance with the GDPR varies across Member States. Although the GDPR harmonizes privacy protection standards across Member States, the enforcement was split between each EU country and its data protection authority (DPA). Therefore, firms in different member states show divergent compliance behaviors towards the GDPR, depending on the regulatory strictness of the respective DPA (see Jia et al., 2020; Johnson et al., 2020; Goldberg et al., 2021; Demirer et al., 2024).

As mentioned in section 2, compliance with the GDPR also varies significantly between industries, with data-driven industries being most affected by the regulation. Koski and Valmari (2021), for instance, find no statistically significant result for their full sample of EU firms comprising all sectors, only after running their regression on firms in data-intensive industries (ICT and financial sector). Similarly, Blind et al. (2024) exploit survey answers from the

[7] Microsoft and Apple have also introduced the GDPR as their global privacy standard. For Microsoft see https://blogs.microsoft.com/on-the-issues/2018/05/21/microsofts-commitment-to-gdpr-privacy-and-putting-customers-in-control-of-their-own-data/ and for Apple https://www.euronews.com/business/2020/11/17/europe-s-gdpr-provides-our-privacy-model-around-the-world-apple-vp-says.

German part of the CIS and notice that among others firms in ICT and financial services report to be the most affected (both positive and negative) by the GDPR, while firms located in non-R&D intensive manufacturing industries (such as agriculture, textiles, wood/paper, metals) and in construction and transport services report to be not affected.

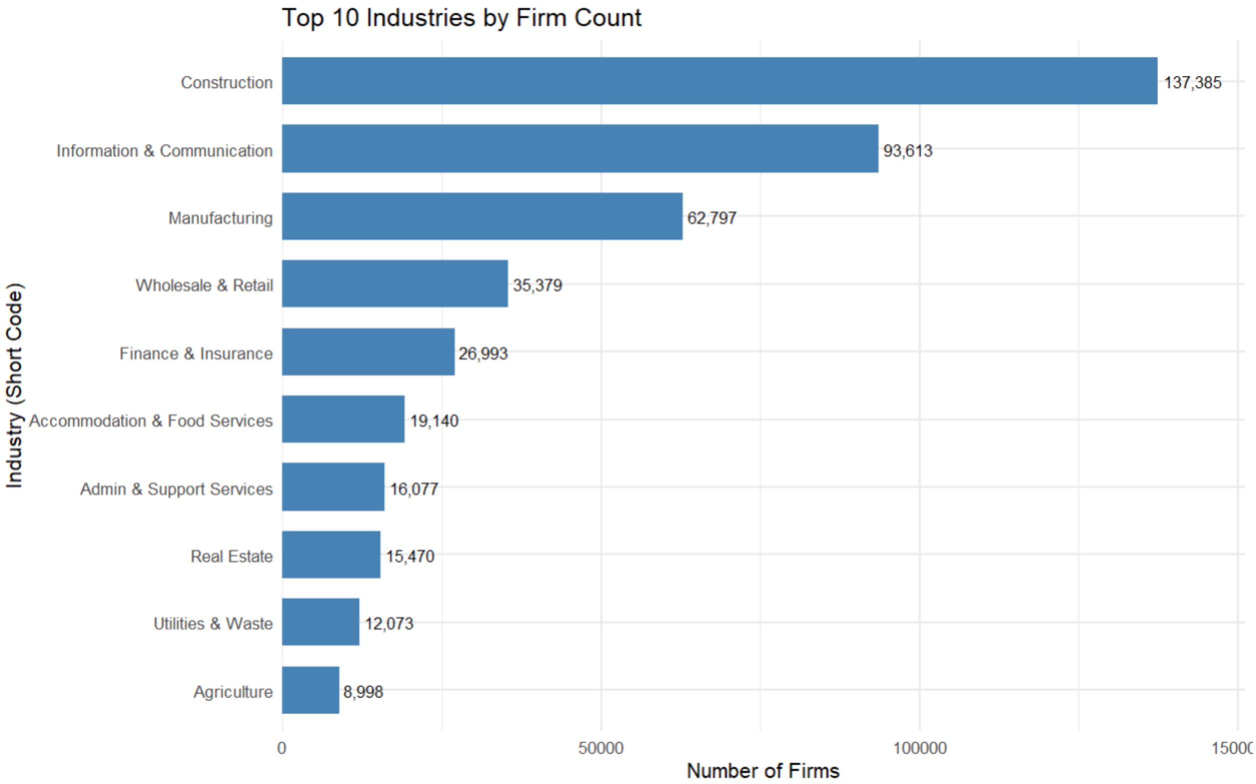
Drawing from these insights and to circumvent the mentioned limitations, I adopt an alternative within-country strategy, exploiting variation in data-intensity across industries to define treated and control groups.[8] Specifically, I use the two-digit NACE Rev. 2 industry classification[9], the official EU system for categorizing economic activities, to identify the sector in which each firm operates. As such, firms in the sectors Information and Communication Sector and Financial and Insurance Services are identified as the most data-intensive and are therefore assigned to the treatment group as they are likely to be most directly and intensively affected by the GDPR. [10] Firms in all other sectors form the control group, under the assumption that their business operations involve significantly less personal data processing, and thus, the regulation imposes relatively lower compliance burdens. Furthermore, I exclude firms from the sector "Professional, Scientific and Technical Activities" from the analysis as they might show an ambiguous effect to the privacy regulation. Figure 3 provides an overview of the distribution of industries in my analytical sample.

[8] Blind et al. (2024) also adopted within-country strategy. In their analysis, the exploited answers from the CIS in which respondent firms could indicate whether they were affected by the GDPR. On this basis, firms were assigned to treatment and control groups.

[9] There has been a recent update of the NACE classification from Rev. 2 to Rev. 2.1. I still use the older version, NACE Rev. 2.

[10] Industries in which firms have been assigned to treated group are: Section J – 58-63: Information and Communication (58-63) and Section K – 64-66: Financial and insurance activities.

Figure 3: *Occurrence of industries in the sample*



This industry-based treatment assignment approach mitigates endogeneity concerns often raised in the literature (see Johnson 2022 for a detailed analysis). By focusing on intra-country, cross-industry variation, it reduces the risk that macroeconomic shocks, country-specific events, or policy spillovers unrelated to the GDPR confound the estimated effects.

Although the GDPR was formally adopted in April 2016, it was not enforced until May 2018, allowing firms a two-year transition period for compliance. Johnson (2022) identifies this policy design as a challenge to empirical research as researchers need to consider the appropriate timing for their identification strategy. For the baseline estimation I therefore focus on the enforcement date in 2018, which is also consistent with most studies (see Johnson 2022). Additionally, to test for anticipatory behavior, I perform robustness tests using the two pre-enforcement years, 2017 and 2016, as placebo. This helps assess whether firms in data-intensive sectors began adapting their operations in expectation of the GDPR, which would bias post-2018 estimates if not properly accounted for. If significant effects are observed prior to the enforcement date, this would suggest pre-treatment trends, raising questions about the robustness of the research design and causal inference.

4.2 Empirical Strategy

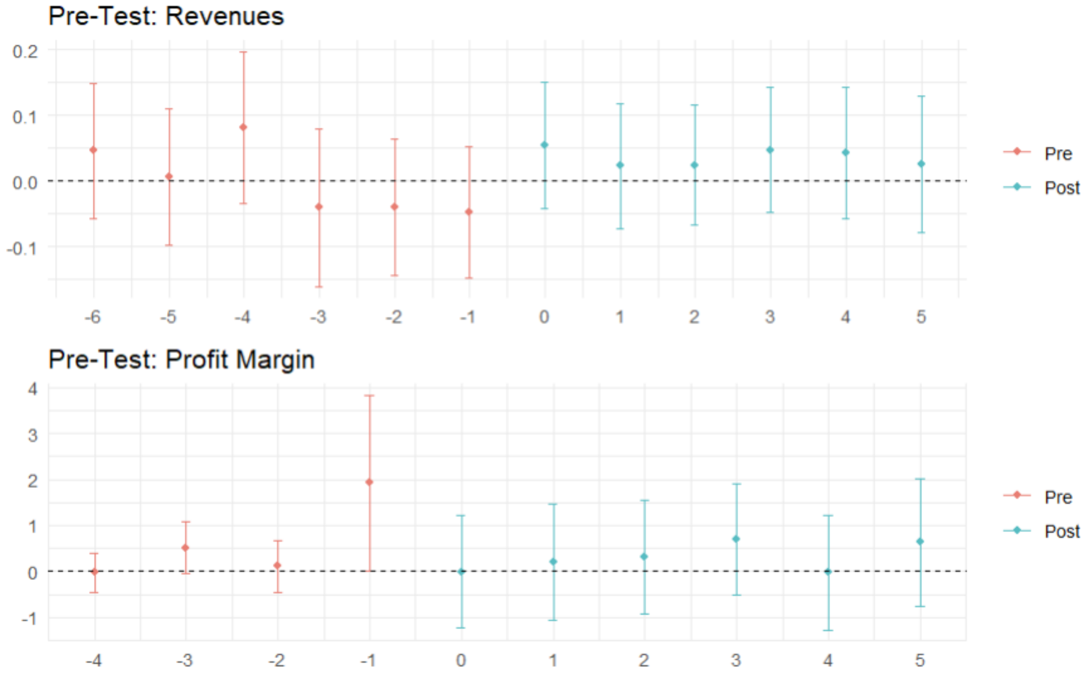
To evaluate the impact of the GDPR on firm performance, I employ a DiD estimator with staggered treatment timing, following the approach of Callaway and Sant'Anna (2020).[11] Recent literature has shown that using the conventional two-way fixed effects regressions (TWFE) as a DiD identification strategy could produce poor estimates as TWFE requires strong assumptions to hold. Any deviations from these assumptions, it is argued, can lead to poor estimates which would severely bias the estimation results. This bias is particular prominent in settings with multiple periods and treatment effect heterogeneity which is the case in this study (for a more comprehensive review of the drawbacks of the use of TWFE for policy evaluation see Callaway, 2022). Instead, Callaway and Sant'Anna (2020) offer a unified framework for estimating average treatment effects for the treated (ATT) for DiD setups with multiple time periods, and when conditional parallel trends assumption is given. Furthermore, the approach proposed by Callaway & Sant'Anna (2020) allows for dynamic effects and treatment effect heterogeneity. This approach enables treatment effect heterogeneity as it allows researchers to estimate varying treatment effects across periods, unlike the canonical DiD set-up which only reflects the aggregate treatment effect on the treated. In my case, this allows me to understand how the impact of the GDPR evolves over time.

Firms in the treated group are assumed to be "treated" starting in 2018, the year the GDPR came into effect. ATTs are estimated using the *att_gt()* function, specifying 2018 as the treatment onset year for treated firms and 0 for never-treated firms. To control for confounding factors, the regression includes covariates capturing firm age, total assets and employment size.

A key assumption of the DiD research design is the parallel trends assumption. This assumption states that, in the absence of treatment, the treated and control groups would have followed similar trends in the outcome variable over time. If this condition is violated, observed post-treatment differences may reflect underlying structural divergences rather than the causal effect of the policy intervention. To assess if the parallel trends assumption holds, I perform event studies as a form of pre-tests. In this context, the parallel trends assumption holds, if the coefficients of the ATT are equal or close to 0 (see Callaway, 2022).

[11] Most conveniently, their approach also comes along with a R package, called *did*, which facilitated the implementation of this approach. I have used the code outlined in their user guidelines and adapted it to my research design: https://bcallaway11.github.io/did/articles/pre-testing.html.

Figure 4: *Results for the Pre-Tests before entropy balancing*



For revenues, the pre-test displayed in Figure 4 suggests that the parallel trends assumption holds at conventional significance levels, as the confidence intervals include 0. In contrast, for profit margins, the pre-test rejects the null hypothesis of parallel trends. Figure 4 shows that for the year 2017 (time period -1) treated and control group deviates from zero, indicating a potential pre-treatment divergence between treated and control groups.

Based on the results of the pre-tests, I employ entropy balancing to improve the comparability between treated and control groups and enhance the credibility of the parallel trends assumption.[12] Entropy balancing is a data preprocessing method that aims at achieving covariate balance with binary treatments by reweighting the control group so that the distribution of covariates (means and variances) matches that of the treated group. In contrast to other methods, such as matching and propensity scores, entropy balancing always improves the balance of covariates as the reweighting scheme is directly incorporated. In addition, entropy balancing keeps all observations, thereby preserving statistical power while improving covariate comparability (Hainmueller 2011). With this approach I follow Blind et al. (2024) that examine the impact of the GDPR on innovation.

I implement the procedure using the *WeightIt* package in R and balancing is based on values in the pre-treatment periods. The balancing covariates are firm age, total assets, and employee counts and then I apply the estimated weights to the full dataset. Table 4 in the appendix illustrates the balance diagnostics before and after implementing entropy balancing. It shows that entropy balancing has significantly improved the balance of treated and control group.

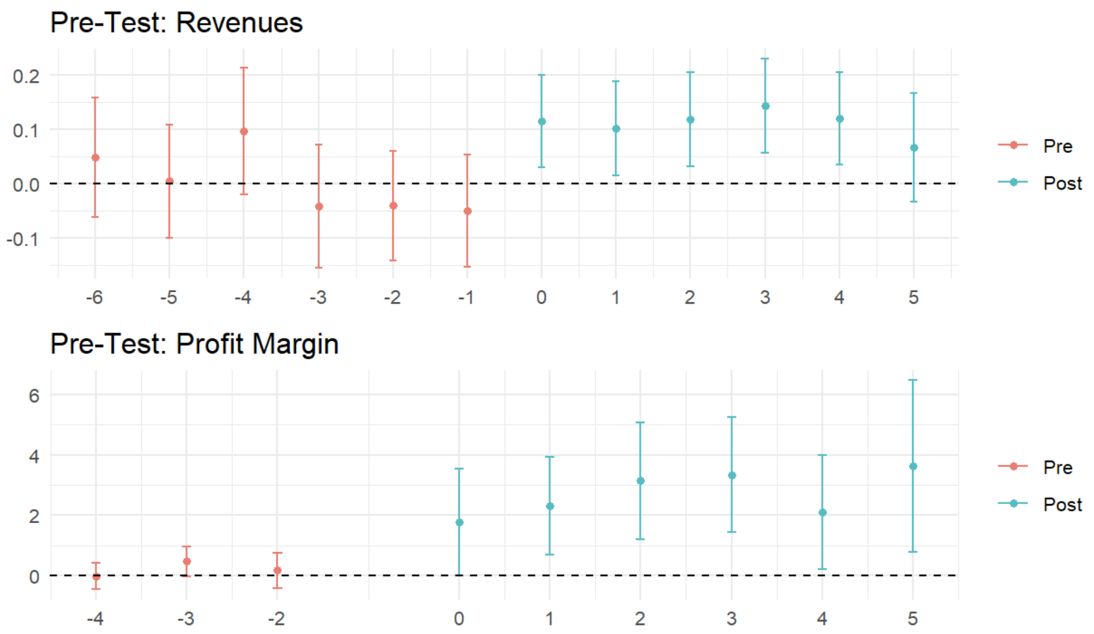
[12] I use the package *WeightIt* (Greifer) in R. The code is adapted from https://ngreifer.github.io/WeightIt/reference/method_ebal.html.

Prior to weighting, notable differences exist between the treated and control groups. For example, prior to balancing treated firms are younger on average (21.15 vs. 24.04 years) and after balancing the average firm age narrows to average firm age narrows to 23.11 (control) and 22.50 (treated).

I re-estimated dynamic treatment effects to assess the pre-treatment periods as shown in Figure 1 in the appendix. For revenues, visual inspection confirms stable pre-treatment differences and confidence intervals consistently covering zero, justifying the parallel trends assumption without further adjustments. For profit margins, although the pre-treatment estimates are generally centered around zero, the significant deviation in year -1 which was previously detected remains.

To ensure robustness, I therefore exclude year -1 from the pre-treatment period. This is a common approach in econometrics and has been also applied by Koski and Valmari (2020). I then re-estimate the pre-tests, shown in Figure 5. The coefficient for both dependent variables are now stable and confidence intervals consistently covering zero, justifying the parallel trends assumption without further adjustments.

Figure 5: *Event Study: Pre-Tests after entropy balancing (on revenues and profit) and excluding period -1 (for profit margin only)*



In Table 3, I furthermore provide the separate summary statistics for the pre-GDPR and post-GDPR periods to provide a clearer picture of firm characteristics before and after GDPR enforcement. The summary statistics indicate that the pre- and post-GDPR periods are broadly comparable across key firm characteristics, such as firm age, profit margin, revenue, and total assets. The mean values and standard deviations remain stable before and after the introduction of GDPR, suggesting no major shifts in the composition of the sample. This pattern holds both when including and excluding the transition year 2017.

Table 3: *Summary statistics for Pre-GDPR and Post-GDPR*

Summary statistics (with year 2017)

Table: Summary Statistics Before and After GDPR

Variable	Period	N	Mean	Standard Deviation	Minimum	Maximum
0	Number of Employees (IHS-transformed)	213422	3.93	1.68	0.88	10.01
0	Firm Age (Years)	213422	24.98	18.49	1.00	125.00
0	Profit Margin (%)	147252	6.85	12.85	-30.26	73.88
0	Revenue (IHS-transformed)	94999	9.10	2.18	0.00	16.27
0	Total Assets (IHS-transformed)	213422	8.44	2.18	0.28	15.50
1	Number of Employees (IHS-transformed)	236851	3.64	1.46	0.88	10.01
1	Firm Age (Years)	236851	21.72	17.28	1.00	125.00
1	Profit Margin (%)	45516	6.75	12.00	-30.28	73.86
1	Revenue (IHS-transformed)	236833	8.59	1.81	0.00	16.61
1	Total Assets (IHS-transformed)	236851	8.25	1.87	0.24	15.50

Summary statistics (without year 2017)

Table: Summary Statistics Before and After GDPR

Variable	Period	N	Mean	Standard Deviation	Minimum	Maximum
0	Number of Employees (IHS-transformed)	199803	3.91	1.66	0.88	10.01
0	Firm Age (Years)	199803	24.98	18.46	1.00	125.00
0	Profit Margin (%)	141173	6.86	12.94	-30.26	73.88
0	Revenue (IHS-transformed)	81380	9.09	2.15	0.00	16.19
0	Total Assets (IHS-transformed)	199803	8.41	2.16	0.28	15.50
1	Number of Employees (IHS-transformed)	236851	3.64	1.46	0.88	10.01
1	Firm Age (Years)	236851	21.72	17.28	1.00	125.00
1	Profit Margin (%)	45516	6.75	12.00	-30.28	73.86
1	Revenue (IHS-transformed)	236833	8.59	1.81	0.00	16.61
1	Total Assets (IHS-transformed)	236851	8.25	1.87	0.24	15.50

V. Empirical results

5.1 Base model

I run the baseline regression on the imputed dataset[13] to assess dynamic treatment effects of the GDPR using the DiD approach by Callaway and Sant'Anna (2020) which also demonstrates how the effect evolves over time.

Revenues

Table 4 shows the full results of the event study on revenues. Figure 6 illustrates the dynamic treatment effect over time. Following the enforcement of the GDPR (year 0), revenues of treated firms increase modestly relative to control firms. The post-treatment dynamic effects are consistently positive and statistically significant in years 0 to 4, with point estimates ranging from approximately 0.10 to 0.14. The overall ATT is estimated at 0.11 (standard error = 0.03), statistically significant at the 5% level. These results suggest that GDPR enforcement was associated with a small but statistically meaningful increase in firm revenues, potentially reflecting enhanced consumer trust or

[13] It is important to note that through the previous transformation the imputed dataset repeated cross-sectional data, rather than a fully balanced panel. This reflects the nature of the underlying data, where many firms only report financials intermittently.

access improved data management practices for firms in data-driven industries.

Table 4: *Results for Event Study on Revenues*

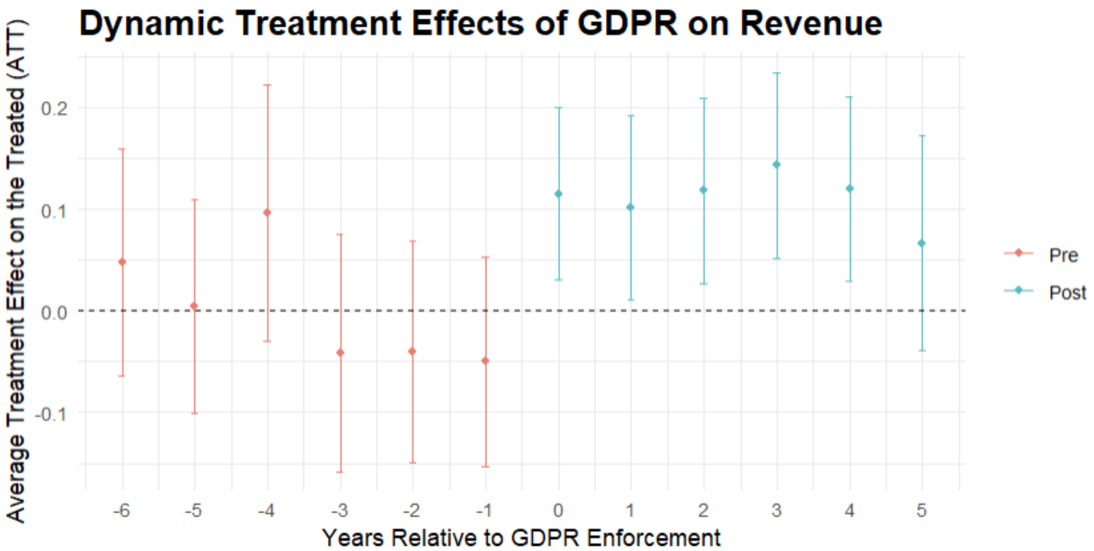
Results for Event Study: Revenues (IHS-transformed)

```
Overall summary of ATT's based on event-study/dynamic aggregation:
    ATT     Std. Error     [ 95%  Conf. Int.]
 0.1105         0.0289     0.0539      0.1671 *


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -6   0.0476     0.0394      -0.0644      0.1597
         -5   0.0043     0.0371      -0.1012      0.1098
         -4   0.0962     0.0444      -0.0299      0.2224
         -3  -0.0412     0.0411      -0.1579      0.0755
         -2  -0.0404     0.0383      -0.1493      0.0686
         -1  -0.0502     0.0363      -0.1533      0.0530
          0   0.1146     0.0298       0.0299      0.1993 *
          1   0.1013     0.0319       0.0107      0.1919 *
          2   0.1180     0.0320       0.0272      0.2089 *
          3   0.1428     0.0321       0.0516      0.2341 *
          4   0.1196     0.0318       0.0293      0.2099 *
          5   0.0666     0.0370      -0.0385      0.1717
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Figure 6: *Results for the Event Study for Revenues*



Profits

In contrast to the revenue results, the analysis of profit margins reveals a more pronounced effect. Table 5 reports the full results of the event study with dynamic effects. Figure 7 plots the dynamic treatment effects over time. Following the enforcement of the GDPR (year 0), profit margins of treated firms increase significantly and consistently relative to control firms. The post-treatment effects are statistically significant across all years, with point estimates ranging from approximately 2.1 to 3.6 percentage points. The overall ATT is estimated at 2.71 percentage points (standard error = 0.34), statistically significant at the 5% level, indicating that, on average, firms in data-intensive industries experienced an increase in profit margins of approximately 2.71 percentage points following the enforcement of the GDPR, relative to firms in less data-driven industries. These results suggest that compliance with GDPR did not hinder firm performance but may have improved operational efficiencies, or changes in market dynamics favored GDPR-compliant companies by changing market dynamics.

Table 5: *Results for the Event Study on Profit Margin*

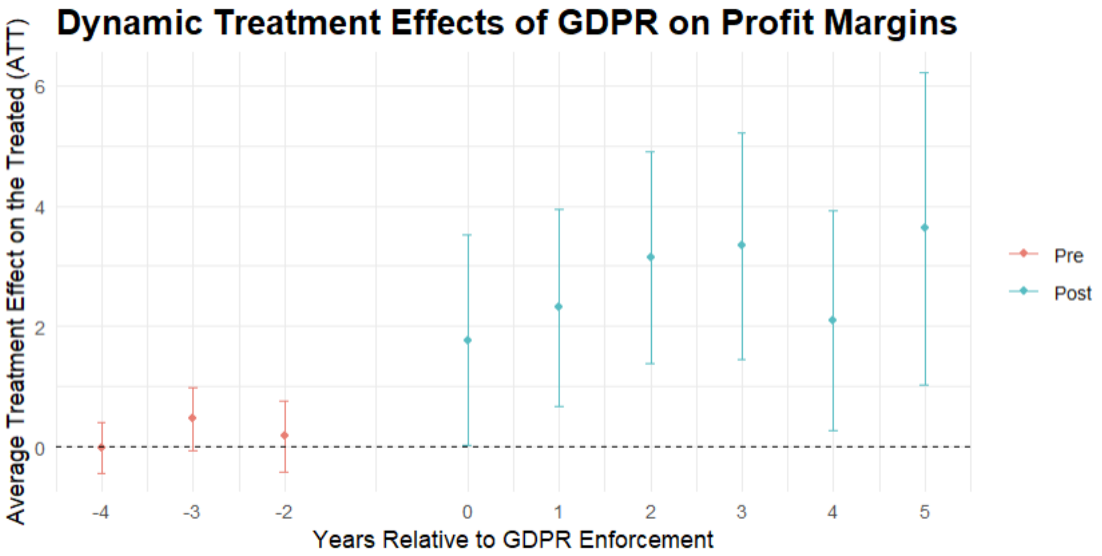
Results for Event Study: Profit Margins

```
Overall summary of ATT's based on event-study/dynamic aggregation:
    ATT    Std. Error     [ 95%  Conf. Int.]
 2.7132        0.3475     2.0322      3.3943 *


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -4  -0.0143     0.1579       -0.4343      0.4058
         -3   0.4671     0.1960       -0.0542      0.9885
         -2   0.1727     0.2206       -0.4141      0.7595
          0   1.7671     0.6557        0.0229      3.5113 *
          1   2.3084     0.6150        0.6724      3.9445 *
          2   3.1405     0.6610        1.3822      4.8988 *
          3   3.3365     0.7094        1.4494      5.2236 *
          4   2.1020     0.6872        0.2739      3.9300 *
          5   3.6250     0.9748        1.0319      6.2181 *
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Figure 7: *Results for the Event Study for Profit Margin*



In sum, the evidence reveals important differences in how GDPR enforcement affected firm profits and revenues. While GDPR had only a modest effect on revenue growth, it may have had a more substantial and positive impact on firm profitability. This has important implications for interpreting the impact of the GDPR on firms which are discussed in section 6. Overall, the findings imply that the economic gains following GDPR enforcement were driven more by improved operational profitability rather than by expanding revenues. The results imply that compliance costs - often cited as a potential burden on firm performance - were not as severe as anticipated. Instead, GDPR may have conferred a strategic advantage to firms already engaged in data-intensive operations, possibly by improving governance structures.

5.2 Placebo test: Treatment in 2016 and 2017

While the GDPR was formally adopted in 2016, its enforcement only began in May 2018. It is plausible that firms began adjusting their internal processes earlier in anticipation of the upcoming requirements. These adjustments could include strengthening data protection policies and restructuring internal IT systems. As the pre-tests have revealed, control and treated groups already show divergent outcomes for profit one year prior to the official enforcement date of the GDPR in 2018.

To further assess the potential for anticipation effects following the announcement of GDPR in 2016, I conduct robustness checks using placebo treatment years in 2016 and 2017. I replicate the DiD models using alternative cutoff years for GDPR implementation—2017 and 2016, instead of the main specification's 2018.

All robustness checks were conducted using the imputed dataset, which ensures a more comprehensive use of the available information and reduces bias from listwise deletion. Given the substantial amount of missing data in the complete case sample, especially for pre-treatment years, the imputed dataset provides a more reliable basis for evaluating the stability of the results across alternative specifications. Figure 8 illustrates the results for revenues, while Figure 9 represents results for profit margins. The full results are in the appendix, see Table 5 and 6 for revenues and Table 7 and 8 for profit margins.

Figure 8: *Results for Placebo tests on Revenues*

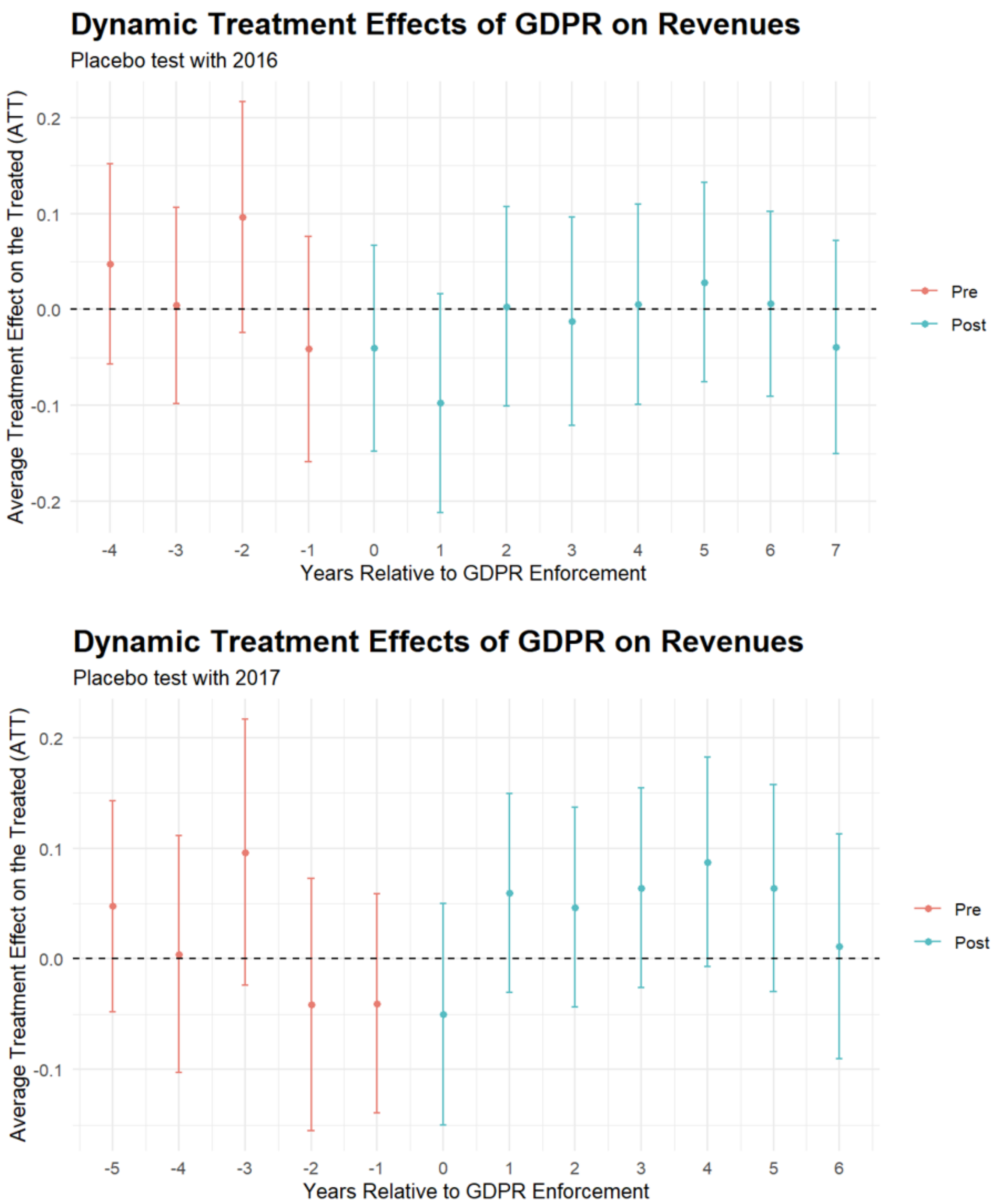
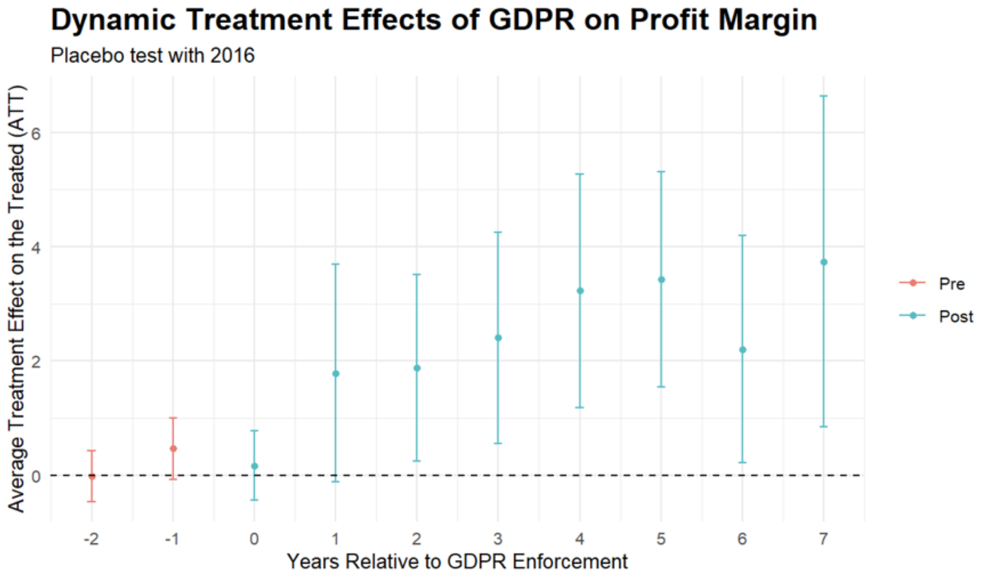
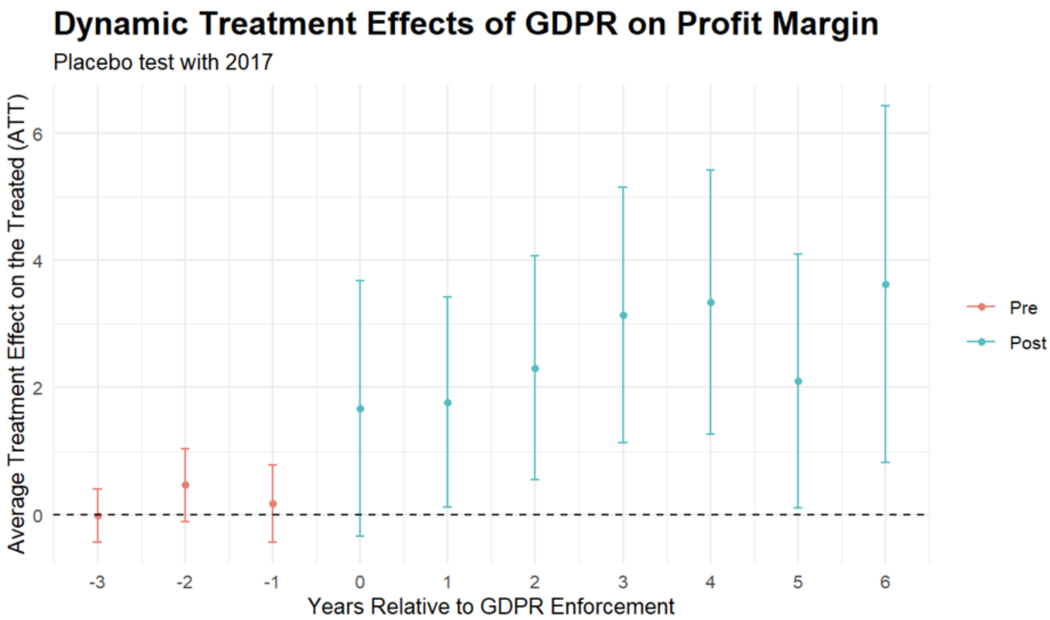


Figure 9: *Results for Placebo tests on Profit Margin*





For revenues, the placebo tests yield small and statistically insignificant overall ATT estimates in both years (0.04 in 2017 and -0.02 in 2016), with dynamic treatment effects fluctuating closely around zero and confidence intervals consistently covering zero. This indicates that revenue trends between treated and control firms remained parallel prior to the actual GDPR enforcement, supporting the validity of the main findings.

For profit margins, however, the overall ATT estimates are statistically significant in both placebo years (2.56 in 2017 and 2.36 in 2016). While the dynamic effects suggest no significant differences before the placebo treatment dates, significant positive effects emerge post-placebo. This pattern points to some early divergence in profit margins, possibly reflecting market adjustments in anticipation of GDPR or other sector-specific dynamics, but does not fundamentally challenge the conclusion that GDPR enforcement contributed to the observed profit margin increases. Importantly, for all models the pre-GDPR periods the confidence interval include 0, indicating the validity of the parallel trends assumption.

5.3 Heterogeneous effect: Firm size

To analyze the distributional effects of the GDPR by firm size, I estimate treatment effects separately for small, medium-sized, and large firms. This heterogeneity analysis allows me to capture potential differences in how firms were affected by the regulatory shock. To define firm size I use two classification schemes based on thresholds outlined in §267 of the *German Commercial Code* (HGB). The first scheme categorizes firms by their average number of employees, while the second scheme relies on average total assets. I create two firm size categories as employment alone might not accurately predict data-driven firms' size.[14] Both size measures are calculated using pre-treatment values (prior to 2018), and firm-level averages are used to assign size categories to avoid post-treatment bias.[15]

One of the key limitations in prior research on the distributional effects of the GDPR - such as in Chen et al. (2022) - was the underrepresentation of small firms in the analyzed datasets. This is particularly important given the hypothesis that regulatory burdens may disproportionately affect firms based on their size. In contrast, my analytical design offers improved coverage across firm sizes by using a country-specific database and imputing missing values. The summary statistics in Table 7 indicate that the samples provides strong coverage across different firm size groups, particularly for small firms. In both asset- and employee-based classifications, small firms represent a substantial share of the observations, with total observations of about 160,000 for size measured based on total assets and about 180,000 observations based on employees.[16] This broad coverage addresses a common limitation in previous studies, where analyses of GDPR and related policies often focused disproportionately on larger firms due to data constraints. By ensuring a robust representation of small firms, the split model design allows for more nuanced insights into heterogeneous treatment effects across firm sizes, improving the external validity of the findings and offering a more comprehensive understanding of GDPR's broader economic impact.

[14] This approach was also conducted by Koski & Valmari (2020) and Chen et al. (2022) which run their regression on different firm size categories.

[15] Before classification, I back-transform values from their inverse hyperbolic sine (IHS) transformation to return them to their original scale.

[16] Note that the number of observations for the regression model is lower as some observations are dropped during this process due to missing values.

Table 6: *Summary Statistics for split models based on firm size*

Summary statistics for split models (by size)

Table: Summary Statistics Across Firm Size Groups (Selected Variables)

Variable	Dataset	N	Mean	SD	Min	Max
Number of Employees (IHS-transformed)	Small (Assets)	165763	3.27	1.14	0.88	9.98
Firm Age (Years)	Small (Assets)	165763	22.58	16.66	2.00	125.00
Profit Margin (%)	Small (Assets)	86724	6.89	13.30	-30.21	73.88
Revenue (IHS-transformed)	Small (Assets)	91617	8.00	1.43	0.00	15.67
Total Assets (IHS-transformed)	Small (Assets)	165763	7.38	1.28	0.24	15.23
Number of Employees (IHS-transformed)	Medium (Assets)	48762	4.33	1.50	0.88	10.00
Firm Age (Years)	Medium (Assets)	48762	23.90	19.41	1.00	125.00
Profit Margin (%)	Medium (Assets)	36173	6.70	12.23	-30.09	73.88
Revenue (IHS-transformed)	Medium (Assets)	28711	9.83	1.54	0.00	16.39
Total Assets (IHS-transformed)	Medium (Assets)	48762	9.33	1.47	0.30	15.10
Number of Employees (IHS-transformed)	Large (Assets)	70845	5.64	1.83	0.88	10.01
Firm Age (Years)	Large (Assets)	70845	26.76	22.83	1.00	125.00
Profit Margin (%)	Large (Assets)	59027	6.74	11.44	-30.28	73.88
Revenue (IHS-transformed)	Large (Assets)	46616	11.61	1.76	0.00	16.61
Total Assets (IHS-transformed)	Large (Assets)	70845	11.16	2.13	0.28	15.50
Number of Employees (IHS-transformed)	Small (Employees)	181210	3.22	1.08	0.88	9.94
Firm Age (Years)	Small (Employees)	181210	22.89	17.57	1.00	125.00
Profit Margin (%)	Small (Employees)	99341	7.13	13.34	-30.21	73.88
Revenue (IHS-transformed)	Small (Employees)	100179	8.14	1.57	0.00	16.52
Total Assets (IHS-transformed)	Small (Employees)	181210	7.63	1.58	0.24	15.23
Number of Employees (IHS-transformed)	Medium (Employees)	64878	4.87	1.31	0.88	10.00
Firm Age (Years)	Medium (Employees)	64878	25.73	21.24	2.00	125.00
Profit Margin (%)	Medium (Employees)	49380	6.82	11.89	-30.26	73.88
Revenue (IHS-transformed)	Medium (Employees)	40782	10.40	1.56	0.00	16.46
Total Assets (IHS-transformed)	Medium (Employees)	64878	9.83	1.84	0.30	15.35
Number of Employees (IHS-transformed)	Large (Employees)	39282	6.47	1.77	0.88	10.01
Firm Age (Years)	Large (Employees)	39282	25.12	20.46	1.00	125.00
Profit Margin (%)	Large (Employees)	33203	5.83	10.65	-30.28	73.88
Revenue (IHS-transformed)	Large (Employees)	25983	12.14	1.67	0.00	16.61
Total Assets (IHS-transformed)	Large (Employees)	39282	11.43	2.14	0.28	15.50

To ensure that treated and control firms are comparable within each subgroup, I again apply entropy balancing within each category size.[17] Finally, similar to the baseline estimation I exclude pre-treatment periods that violate the parallel trends assumption from estimation. This ensures that the dynamic effects are only interpreted for periods where the identifying assumptions of the DiD framework are met. The results for the dynamic effects are presented in Figure 10 for large firms, Figure 11 for medium-sizes firms, and Figure 12 for small firms. The full results are shown in the appendix in Tables 5 and 6 for large firms, Tables 7 and 8 for medium-sized firms, and Tables 9 and 10 for small firms.

[17] I use the *WeightIt* package again in R. The balancing procedure is performed on pre-treatment data and reweighted across the full dataset by carrying forward weights to post-treatment years within each firm. Covariates used for balancing include firm age, total assets, and the number of employees.

Figure 10: *Event Study Results for Large Firms*

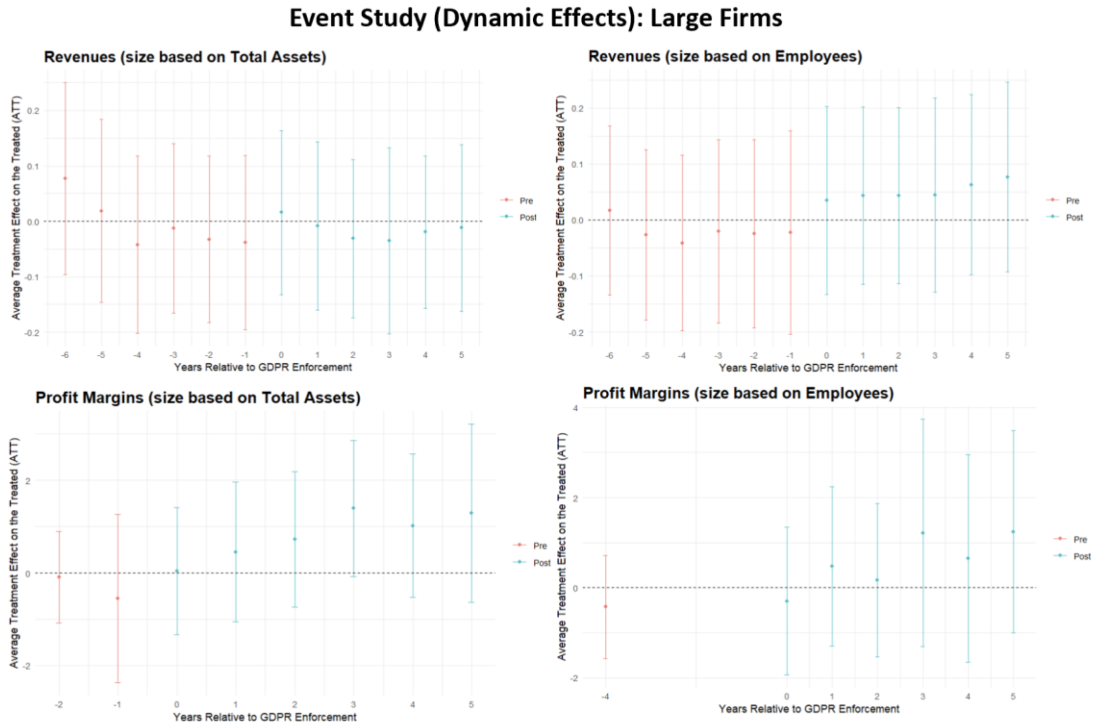


Figure 11: *Event Study Results for Medium-sized Firms*

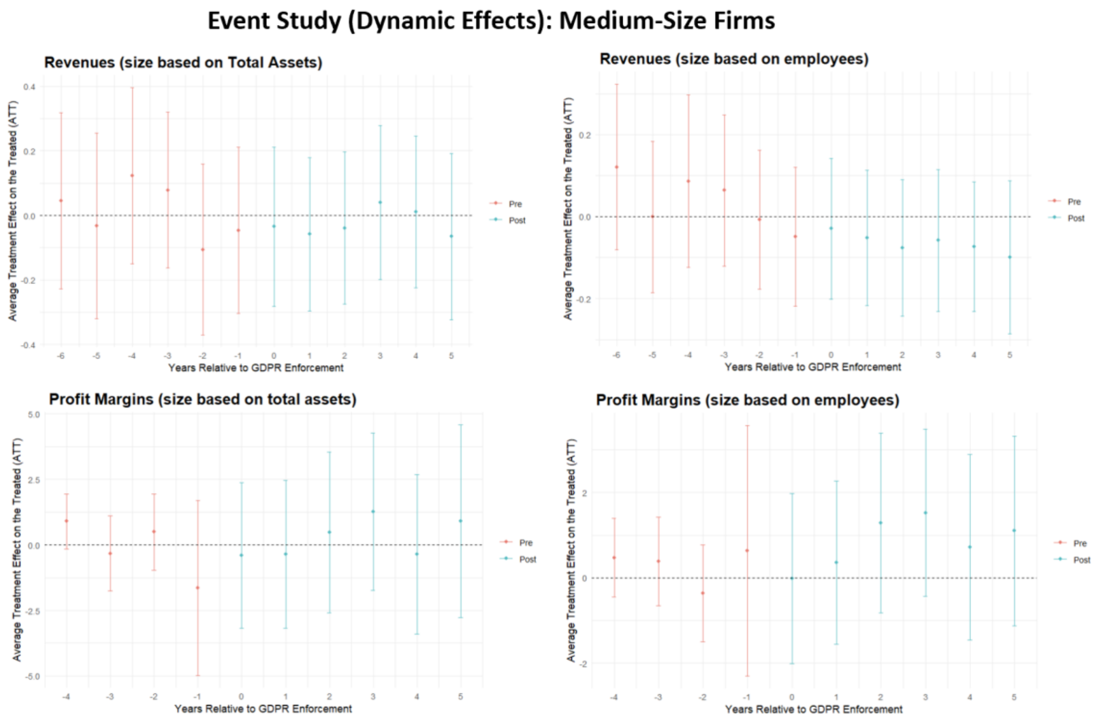
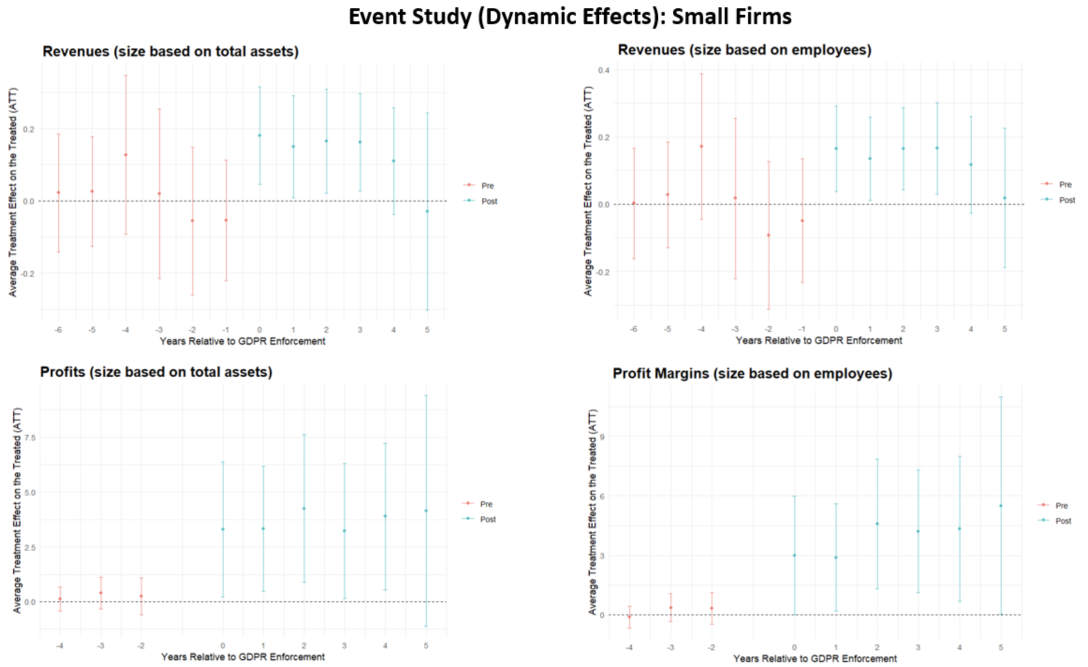


Figure 12: *Event Study Results for Small Firms*



Overall, the results indicate that the GDPR's impact is particularly pronounced among small firms, with relatively muted effects among larger firms. For small firms, both revenues and profit margins increased significantly after GDPR enforcement. The dynamic effects show positive and statistically significant treatment effects starting from the enforcement year (2018) onward, irrespective of whether size was measured by assets or employees.

In contrast, medium-sized firms exhibit no consistently significant effects: the ATT estimates remain statistically insignificant across most post-treatment periods for both revenues and profit margins, suggesting a more limited or heterogeneous adjustment to GDPR compliance costs. Among large firms, there is no statistically significant dynamic effect, with confidence intervals consistently covering zero. This pattern suggests that larger firms may possess better resources and capabilities to absorb GDPR-related adjustments without significant disruptions to their financial outcomes and highlights that GDPR compliance imposed a relatively greater adjustment burden on small firms compared to their larger counterparts. However, these adjustment efforts have a positive impact on small firms' performance. Moreover, the consistency across the two different size measures (total assets and employees) reinforces the robustness of the size heterogeneity pattern observed.

5.4 Robustness test: Complete case analysis

To assess the robustness of the results, I re-estimated the dynamic treatment effects using only complete case observations.[18] Figure 2 in the Appendix shows the full results for revenues and Figure 13 displays the dynamic effects. For revenues, the complete case analysis yields an overall ATT of 0.04, with a 95% confidence interval spanning from -0.01 to 0.01. None of the dynamic treatment effects are statistically significant throughout the event window. Compared to the imputed dataset, where the ATT was higher and statistically significant, the complete case results show a weaker and more uncertain association between GDPR enforcement and revenue developments.

For profit margins the full results are displayed in Figure 3 in the appendix and Figure 14 illustrates the dynamic effects. The complete case analysis reports an overall ATT of 0.73, which remains statistically significant at the 5% level. However, similar to revenues, the individual dynamic effects are not statistically significant across periods. In comparison, the imputed results indicated a stronger and more consistently significant positive effect on profit margins.

Overall, the complete case analysis exhibits wider confidence intervals and larger standard errors, indicating reduced statistical power relative to the imputed analysis. This is likely due to the substantial loss of observations when restricting to complete cases, which reduces efficiency and precision. Hence, the models based on imputed data are preferred for the main analysis, as they allow for a larger sample, greater statistical power, and more reliable inference, while still yielding qualitatively consistent results in the robustness checks.[19]

The comparison of complete case and imputed values highlight valuable information about the sensitivity of missing data and its consequences on model estimation. As I have shown in the heterogeneity analysis in 5.3, the increase in profits and revenues is mainly driven by small firms. As small firms exhibit more missing values, particularly in profit data, they are underrepresented in the complete case analysis. Therefore, the estimation similarly yields more pessimistic results than for the imputed case. An estimation solely based on complete cases would therefore introduce selection bias into the estimation.

---

[18] For the complete case datasets, I employed the same data preprocessing methods (i.e. entropy balancing) as for the imputed dataset for consistency.

[19] Due to extensive missing data in the complete-case sample—particularly for firm size variables— I can not perform the heterogeneity analysis on the complete-case sample as this would not yield any meaningful and robust results. For instance, more than 60% of employee-based size values are missing in the complete-case data, making it unsuitable for size-based segmentation.

Figure 13: *Results for the Event Study for Revenues (Complete-Case)*

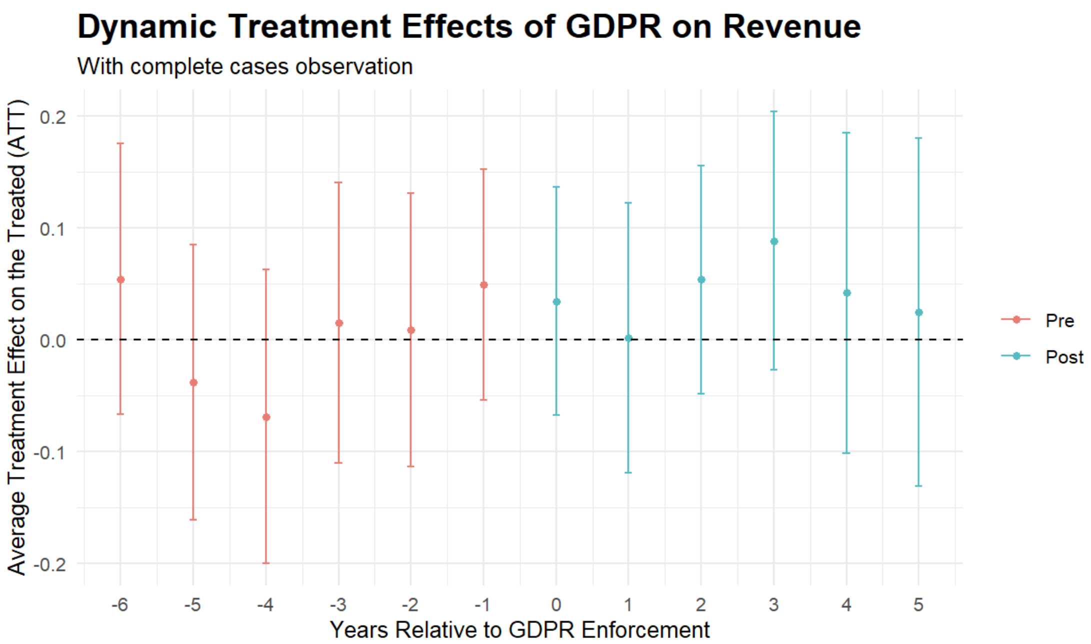
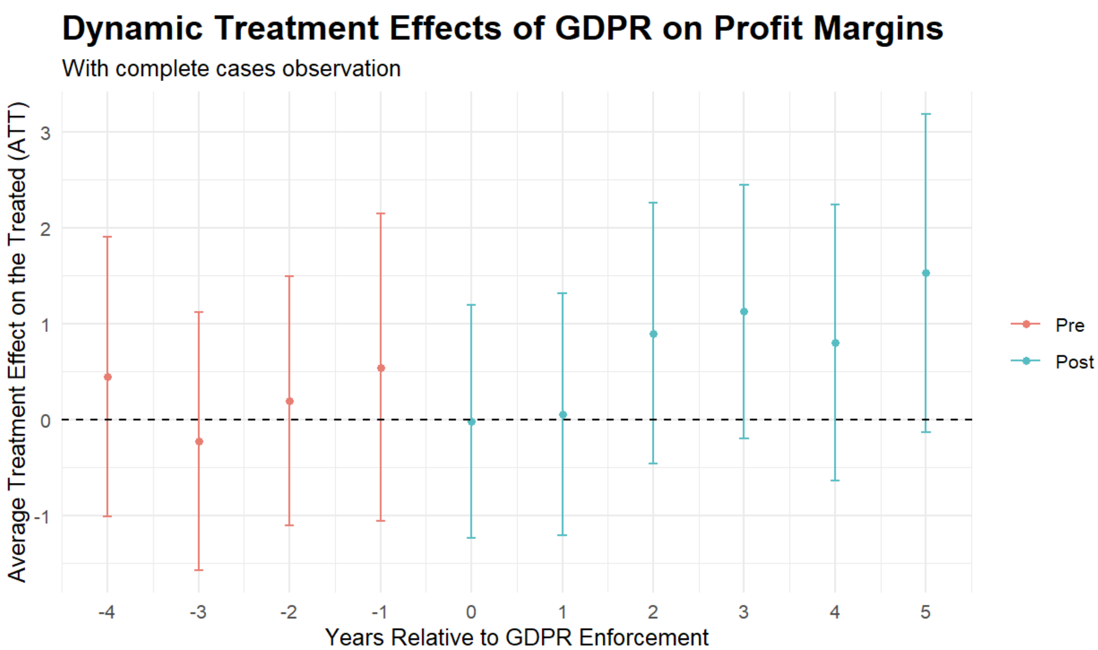


Figure 14: *Results for the Event Study for Profit Margin (Complete-Case)*



VI. Discussion

6.1 Possible Mechanisms

The results reveal interesting dynamics regarding the effects of the GDPR on firms, particularly for small firms. While much of the literature expected the GDPR to impose high compliance costs, especially for SMEs, the results of this study indicate that compliance costs may have been lower than anticipated.

The GDPR did not harm firm performance in data-driven industries in Germany. On the contrary, the findings suggest that firms may have benefited, particularly along the profit channel, rather than the revenue channel. While revenues remained relatively stable or saw only modest increases in magnitude, the treatment effects on profit margins were substantial and statistically significant. The results further show that this effect is mainly driven by small firms while larger firms seem to be unaffected in their financial performance by the GDPR. Medium-sized firms show conflicting and not significant results.

Analyzing the dynamic treatment effects separately for revenues and profit margins might hint at differing adjustment and benefit patterns over time. Revenues for small firms show a relatively immediate increase following the GDPR's enforcement, peaking around 2018, shortly after the regulation came into force. This early response could be interpreted as small firms benefiting from an expanded and harmonized internal market as argued by Bradford (2024), with lowered barriers to cross-border business activities within the EU.

In contrast, the profit margins exhibit a delayed but stronger increase, with the largest effects materializing several years after GDPR enforcement. This lagged response implies that the longer-term benefits to small firms may stem from internal organizational improvements induced by GDPR compliance, such as better data management, customer trust, and operational efficiencies. These internal gains likely took more time to fully develop and translate into higher profitability.

This difference between the immediate revenue gains and the delayed profit supports the interpretation that small firms initially benefited externally through increased market access and later internally through improved business processes. It also highlights the strategic importance of compliance investments: although costly upfront, they can yield substantial efficiency and profitability gains over time. This would be consistent with literature studying the impact of the GDPR in a shorter time period.[20]

The results suggest that small firms benefitted both from immediate market access and from longer-term organizational gains, although the latter appears to have played a stronger role in enhancing firm performance. The larger and more persistent effects observed for profit margins compared to revenues suggest that GDPR-induced organizational changes had a meaningful impact on firms' cost structures or customer relationships, beyond simply expanding market opportunities.

These findings align with the evidence provided by Blind et al. (2024) and other authors (Maex, 2022). The GDPR may have induced firms to improve data management and reduce redundancy in processing, as part of compliance to the GDPR. Firms may have perceived the GDPR as an opportunity to review

[20] See Koski and Valmari (2020) that study the short-time effect of the GDPR on firm performance from 2014-2018.

their internal IT routines and identify and improve weaknesses, leading to better performance.

Furthermore, firms may have realized their untold potential by “improving existing products and better aligning them to customer needs through a more effective use of data” (Blind et al., 2024, p.330). Blind et al. (2024) document that firms affected by the GDPR in their innovation activities also exhibit significantly higher IT and data capabilities than firms that were not affected. Similarly, in their analysis small firms benefitted by increased sales of incremental innovation, incentivized by the GDPR. They argue that this is because small firms tend to have less established internal processes and can therefore benefit more from improving IT operations. The evidence suggests that the GDPR forced firms to adjust their data management to ensure compliance. This created the opportunity for firms to update existing products "through a more thoroughly or smarter use of data" (Blind et al., 2024, p.313)

Overall, these findings underline the dual role of regulation: not only as a market-shaping instrument but also as a catalyst for internal firm transformation, especially for smaller, more adaptable firms.

Robustness checks, including placebo tests using 2016 and 2017 as false treatment years, further support the causal identification of the effects tied to the actual enforcement year, 2018, although the results for profit margins indicate that firms may have adapted their process prior to the official enforcement year in 2018.

In addition to the main specification, this thesis incorporates models estimated using both complete case data and imputed datasets to strengthen the robustness of the results. A considerable share of observations contained missing values, particularly in firm-level variables such as total assets and employment, which if excluded would have led to a substantial reduction in sample size and potentially biased estimates due to non-random missingness. To address this issue, I implemented multiple imputation techniques that preserve data variability and better reflect the underlying distribution of missing observations. Importantly, the results obtained from the imputed dataset remained qualitatively consistent with those from the complete data analysis, especially regarding the direction and timing of treatment effects. For instance, the positive effects on profit margins for post-GDPR enforcement periods remained significant across both specifications, while the patterns for revenue remained relatively muted in both datasets. This consistency enhances the internal validity of the findings and increases confidence that the observed treatment effects. Moreover, imputing allowed the use of a more comprehensive and representative panel, particularly benefiting the analysis of small and medium-sized firms, which are often more prone to missingness due to incomplete administrative or financial reporting.

The results are in line with recent studies by Blind et al. (2024) or Lefrere et al. (2022). However, they stand in stark contrast to other studies that outline the negative effects of the GDPR on firm performance (Chen et al., 2022; Koski &

Valmari, 2020), market concentration (Peukert et al. 2022) or venture investments (Jia et al., 2021).

While the findings suggest a positive effect of the GDPR on firm profit margins, data limitations prevent fully accounting for all potential confounders. Furthermore, I cannot rule out that the analysis inherently captures only the surviving firms, meaning the sample excludes those that may have exited the market due to GDPR-induced burdens which was documented by the literature (Janßen et al., 2022). This selection bias - and specifically survivor bias - may distort the true distributional consequences and is a common threat to validity in econometrics.

6.2 Policy Implications

The results of this study carry several important implications for policymakers. The findings suggest that the GDPR did not harm firm performance in data-driven industries by adding compliance costs. Rather, firms - particularly smaller ones - experienced gains in profitability following its enforcement in 2018. This challenges the prevailing narrative that privacy regulations disproportionately burden small businesses and indicates that well-designed regulatory frameworks can be compatible with firm-level economic resilience. Moreover, the results underline that the GDPR's profit channel effects - rather than revenue increases - likely stem from efficiency gains, which suggests that regulations may also serve as catalysts for internal modernization and process optimization.

Additionally, the evidence that smaller firms benefited from the implementation of the GDPR disproves concerns that data protection rules inevitably widen competitive gaps between large and small enterprises. Instead, the GDPR may have leveled the playing field by setting common compliance standards, enabling smaller firms to compete more effectively with their larger counterparts. Finally, the study supports the idea that regulatory harmonization, such as the GDPR's replacement of fragmented national laws, can reduce compliance complexity and legal uncertainty, particularly for firms operating across EU member states. These findings collectively point toward the possibility that wecoherent regulatory designs not only protect consumer rights but also foster innovation, operational improvements, and cross-border business scalability within the internal market.

VII. Conclusion

This thesis analyzed the impact of the GDPR on firm performance, focusing on revenues and profit margins across different firm sizes in Germany. Using firm-level panel data and employing the Callaway and Sant'Anna (2020) DiD framework to account for heterogeneous treatment effects, the study offers new insights into the distributional impacts of the GDPR by firm size.

The findings challenge earlier concerns about the economic consequences of privacy regulation. Notably, small firms - often perceived as the most vulnerable

- exhibited the strongest positive effects in profit margins. This suggests that previous studies may have underestimated the resilience or adaptive capacity of smaller firms, particularly in digital sectors. Contrary to widespread fears, the GDPR did not negatively affect firm performance in data-intensive sectors and instead, treated firms experienced stable or improved outcomes after 2018. While revenue effects were modest, profit margins increased significantly, particularly among small firms. These results imply that GDPR compliance may have fostered internal cost efficiencies, possibly through improved data governance, streamlined operations, or reduced redundancies introduced during compliance efforts.

This study contributes to the literature by highlighting the distributional effects of the GDPR. It shows that small firms, among those that survived, demonstrated some of the strongest profitability gains, emphasizing the need to consider heterogeneity in regulatory impact studies.

Overall, the GDPR did not impose the severe performance costs that critics and much of the literature anticipated. Instead, for firms that adapted successfully - especially smaller firms in data-driven industries - compliance may have supported resilience and operational advantages. These insights are valuable for policymakers, underlining the importance of thoughtful regulatory design, targeted SME support, and attention to market dynamics during large-scale policy transitions.

However, these optimistic results must be interpreted with caution. First, the findings are only valid for the German context, where a national privacy law - the *Bundesdatenschutzgesetz* (BDSG) - already existed before the GDPR. While Martin et al. (2019) document that the BDSG was often loosely enforced, the GDPR's introduction of higher fines substantially increased the cost of non-compliance.

Furthermore, there are several limitations to this study. First, treatment assignment is based on whether a firm operates in a data-intensive sector (ICT, finance and insurance), simplifying variation in GDPR exposure. This may lead to misclassification, as firms outside these sectors could also be affected by the GDPR. Thus, the estimated effects reflect differential impacts across sectors rather than a pure treated-versus-untreated contrast, partially violating the Stable Unit Treatment Value Assumption (SUTVA) assumption underlying DiD estimation.

Second, the strong positive effects among small firms might reflect survivor bias. Firms that survived post-GDPR implementation periods are likely the most resilient or efficient and may not represent the broader population of small firms. This is underlined by studies on start-ups which found high exits rates after the introduction of the GDPR (see Janssen et al., 2022). This selection and more specifically survivor bias is a common threat in econometric studies dealing with firm-performance data. The result can thus be too optimistic and show a distorted version of reality (see Ball and Ross, 1979; Shermer, 2014).

Third, the GDPR restricts data availability, complicating empirical analysis. Financial data gaps, as discussed in Section 3, further constrain the ability to fully assess GDPR's impacts.

Finally, the post-treatment window includes 2020 and 2021, during which the COVID-19 pandemic caused major economic disruptions. These external shocks, especially sector-specific effects and additional regulatory reforms (e.g., in the financial sector), may confound the estimated impacts. Although entropy balancing and dynamic DiD mitigate some concerns, it remains possible that the results capture broader economic trends beyond GDPR effects.

While this thesis provides evidence on the distributional impact of the GDPR on firm performance across different firm sizes, future research could investigate further topics. First, it would be valuable to extend the analysis beyond Germany to a cross-country setting within the European Union to assess whether the observed effects generalize across different legal, economic, and institutional contexts. Second, future studies could explore more granular mechanisms behind the observed profit improvements, such as investments in IT security, changes in customer trust, or operational restructuring, by incorporating additional firm-level survey or investment data. Third, distinguishing between firms that successfully adapted to the GDPR and those that exited the market would allow researchers to more precisely estimate the true selection effects and better account for potential survivor bias. Finally, as the GDPR represents only one of the first comprehensive tech regulation, it would be important to investigate its interaction with subsequent regulatory initiatives such as the Digital Services Act (DSA) and the Digital Markets Act (DMA), in order to better understand the cumulative regulatory burden — or synergy — on European firms' competitiveness and innovation outcomes over the longer term.

Bibliography

Aridor, G., Che, Y.-K., & Salz, T. (2022). The effect of privacy regulation on the data industry: Empirical evidence from GDPR. *The RAND Journal of Economics*, *54*(4), 695–730. https://doi.org/10.1111/1756-2171.12455.

Bajgar, M., Berlingieri, G., Calligaris, S., Criscuolo, C., & Timmis, J. (2020). *OECD Science, Technology and Industry Working Papers 2020/06 Coverage Coverage and representativeness of Orbis data* (OECD Science, Technology and Industry Working Papers, 2020/06). OECD. https://www.oecd.org/content/dam/oecd/en/publications/reports/2020/05/coverage-and-representativeness-of-orbis-data_9628c322/c7bdaa03-en.pdf https://doi.org/10.1787/c7bdaa03-en.

Ball, R., & Watts, R. (1979). Some Additional Evidence on Survival Biases. *The Journal of Finance*, *34*(1), 197. https://doi.org/10.2307/2327153.

Bessen, J. E., Impink, S., Reichensperger, L., & Seamans, R. (2020). *GDPR and the Importance of Data to AI Startups.* https://doi.org/10.2139/ssrn.3576714.

Beuselinck, C., Elfers, F., GASSEN, J., & Pierk, J. (2021). Private firm accounting: the European reporting environment, data and research perspectives. *Accounting and Business Research*, *53*(1), 38–82. https://doi.org/10.1080/00014788.2021.1982670.

Blind, K., Niebel, C., & Rammer, C. (2024). The impact of the EU General data protection regulation on product innovation. *Industry and Innovation*, *31*(3), 311–351. https://doi.org/10.1080/13662716.2023.2271858.

Bradford, A. (2012). *The Brussels Effect.*

Bradford, A. (2023). Europe's Digital Constitution. *Va. J. Int'l L.*, *64*, 1. https://scholarship.law.columbia.edu/faculty_scholarship/4168.

Bradford, A. (2024). *The False Choice Between Digital Regulation and Innovation.* https://doi.org/10.2139/ssrn.4753107.

Bryzgalova, S., Lerner, S., Lettau, M., & Pelger, M. (2022). *Missing Financial Data.* https://doi.org/10.2139/ssrn.4106794.

Callaway, B. (2022). Difference-in-Differences for Policy Evaluation. *Papers*. https://ideas.repec.org/p/arx/papers/2203.15646.html.

Callaway, B., & Sant'Anna, P. H. (2020). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001.

Campbell, J., Goldfarb, A., & Tucker, C. (2015). Privacy Regulation and Market Structure. *Journal of Economics & Management Strategy*, *24*(1), 47–73. https://doi.org/10.1111/jems.12079.

Chen, C., Carl Benedikt, & Presidente, G. (2022). Privacy regulation and firm performance: Estimating the GDPR effect globally. *Economic Inquiry*, *62*(3), 1074–1089. https://doi.org/10.1111/ecin.13213.

Demirer, M., Hernández, D. J. J., Li, D., & Peng, S. (2024). Data, Privacy Laws and Firm Production: Evidence from the GDPR. *National Bureau of Economic Research.* Advance online publication. https://doi.org/10.3386/w32146.

European Commission. (2012). *Commission staff working paper: Impact Assessment Accompanying the document - General Data Protection Regulation and Directive on processing of personal data (SEC(2012) 72 final*. https://www.europarl.europa.eu/cmsdata/59702/att_20130508ATT65856-1873079025799224642.pdf.

European Commission. (2025, April 27). *Regulation - 2016/679 - EN - gdpr - EUR-Lex*. https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng.

GASSEN, J., & MUHN, M. (2025). Financial Transparency of Private Firms: Evidence from a Randomized Field Experiment. *Journal of Accounting Research*, *63*(1), 413–460. https://doi.org/10.1111/1475-679X.12568.

Goddard, M. (2017). The EU General Data Protection Regulation (GDPR): European Regulation that has a Global Impact. *International Journal of Market Research*, *59*(6), 703–705. https://doi.org/10.2501/IJMR-2017-050.

Godinho de Matos, M., & Adjerid, I. (2019). *Consumer Consent and Firm Targeting after GDPR: The Case of a Large Telecom Provider.* https://doi.org/10.2139/ssrn.3777417.

Goldberg, S. G., Johnson, G. A., & Shriver, S. K. (2021). Regulating Privacy Online: An Economic Evaluation of the GDPR. *American Economic Journal: Economic Policy*, *16*(1), 325–358. https://doi.org/10.1257/pol.20210309.

Goldfarb, A., & Tucker, C. (2012). Shifts in Privacy Concerns. *American Economic Review*, *102*(3), 349–353. https://doi.org/10.1257/aer.102.3.349.

Hainmueller, J. (2012). Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies. *Political Analysis*, *20*(1), 25–46. https://doi.org/10.1093/pan/mpr025.

Janßen, R., Kesler, R., Kummer, M., & Waldfogel, J. (2022). GDPR and the Lost Generation of Innovative Apps. *National Bureau of Economic Research.* Advance online publication. https://doi.org/10.3386/w30028.

Jia, J., Jin, G. Z., & Wagman, L. (2020). *The Short-Run Effects of GDPR on Technology Venture Investment.* https://doi.org/10.2139/ssrn.3278912.

Johnson, G. (2022). Economic Research on Privacy Regulation: Lessons from the GDPR and Beyond. *National Bureau of Economic Research.* Advance online publication. https://doi.org/10.3386/w30705.

Johnson, G., & Shriver, S. (2022). *Privacy & Market Concentration: Intended & Unintended Consequences of the GDPR.* https://doi.org/10.2139/ssrn.3477686.

Kerr, D., & Bhuiyan, J. (2025, February 22). Crypto and big tech's backing pays off as Trump makes tech-friendly moves. *The Guardian*. https://www.theguardian.com/technology/2025/feb/22/crypto-big-tech-trump.

Koski, H., & Valmari, N. (2020). *Short-term Impacts of the GDPR on Firm Performance* (ETLA Working Papers No. 77). Helsinki: The Research Institute of the Finnish Economy (ETLA). https://www.econstor.eu/handle/10419/237362.

Lefrere, V., Warberg, L., Cheyre, C., Marotta, V., & Acquisti, A. (2022). *Does Privacy Regulation Harm Content Providers? A Longitudinal Analysis of the Impact of the GDPR.* https://doi.org/10.2139/ssrn.4239013.

Maex, S. A. (2022). *Modern PRIVACY REGULATION, INTERNAL INFORMATION QUALITY, AND OPERATING EFFICIENCY: Evidence FROM THE GENERAL DATA PROTECTION REGULATION.* https://doi.org/10.34944/DSPACE/8025.

Martin, N., Matt, C., Niebel, C., & Blind, K. (2019). How Data Protection Regulation Affects Startup Innovation. *Information Systems Frontiers*, *21*(6), 1307–1324. https://doi.org/10.1007/s10796-019-09974-2.

Niebel, C. (2021). The impact of the general data protection regulation on innovation and the global political economy. *Computer Law & Security Review*, *40*, 105523. https://doi.org/10.1016/j.clsr.2020.105523.

Peukert, C., Bechtold, S., Batikas, M., & Kretschmer, T. (2022). Regulatory Spillovers and Data Governance: Evidence from the GDPR. *Marketing Science*, *41*(4), 746–768. https://doi.org/10.1287/mksc.2021.1339.

Shermer, M. (2014, January 9). How the Survivor Bias Distorts Reality. *Scientific American*. https://www.scientificamerican.com/article/how-the-survivor-bias-distorts-reality/.

van Buuren, S. (2018). *Flexible Imputation of Missing Data*. https://stefvanbuuren.name/fimd/.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice : Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03.

Appendix

The full code can be accessed at:
https://github.com/Vyhuyen/gdpr_firm_analysis

Access to the data is restricted and can therefore not be provided with the code.

Furthermore, analyses were conducted in R version 4.3.2 (R Core Team, 2024), using among others the following packages: dplyr (Wickham et al.), did (Callaway and Sant'Anna), WeightIt (Greifer), and ggplot2 (Wickham).

Table 7: Percentage of Missing Values in the key variables

Percentage of Missing Values by Variable

	Variable	Missing_Percentage
profit_margin	Profit Margin	57.19
employees_IHS	Employees	39.45
revenue_IHS	Revenue	26.30
total_assets_IHS	Total Assets	14.24
firm_age	Firm Age	4.79

Table 8: *Summary Statistics of imputed and observed values*

Summary of Observed vs. Imputed Values

	Mean			Median		SD	
Variable	Mean_Observed	Mean_Imputed	Mean_Diff_Perc	Median_Observed	Median_Imputed	SD_Observed	SD_Imputed
total_assets_IHS	8.25	8.34	1.10	7.87	7.94	1.97	2.02
employees_IHS	3.89	3.78	-2.93	3.64	3.58	1.61	1.57
firm_age_C	23.25	23.26	0.06	19.00	19.00	17.94	17.94

Table 9: *Summary Statistics of the imputed and complete-case datasets*

Summary Statistics in Long Format

dataset	Variable	Mean	SD	N
Complete Case: Profit	Profit Margin	6.34	11.24	68325
Complete Case: Profit	Revenue (IHS)	11.04	1.88	59543
Complete Case: Profit	Employees (IHS)	5.21	1.86	68325
Complete Case: Profit	Total Assets (IHS)	10.49	2.12	68325
Complete Case: Profit	Firm Age	23.54	22.04	68325
Complete Case: Revenue	Profit Margin	6.15	10.64	59543
Complete Case: Revenue	Revenue (IHS)	8.94	1.93	245606
Complete Case: Revenue	Employees (IHS)	3.87	1.61	245606
Complete Case: Revenue	Total Assets (IHS)	8.58	2.03	245606
Complete Case: Revenue	Firm Age	22.77	18.37	245606
Imputed	Profit Margin	6.83	12.65	192768
Imputed	Revenue (IHS)	8.74	1.94	331832
Imputed	Employees (IHS)	3.78	1.57	450273
Imputed	Total Assets (IHS)	8.34	2.02	450273
Imputed	Firm Age	23.26	17.94	450273

Table 10*: Balance Diagnostics before and after employing entropy balancing*

Balance Diagnostics Before and After Entropy Balancing

	Covariate Means			
Treatment Group	Firm Age (C)	Total Assets (IHS)	Employees (IHS)	Type of Mean
0	24.04	8.26	3.71	Unweighted Mean
1	21.15	8.55	3.96	Unweighted Mean
0	23.11	8.68	4.03	Weighted Mean
1	22.50	8.58	4.02	Weighted Mean

Figure 15: *Event Study: Pre-Tests after entropy balancing*

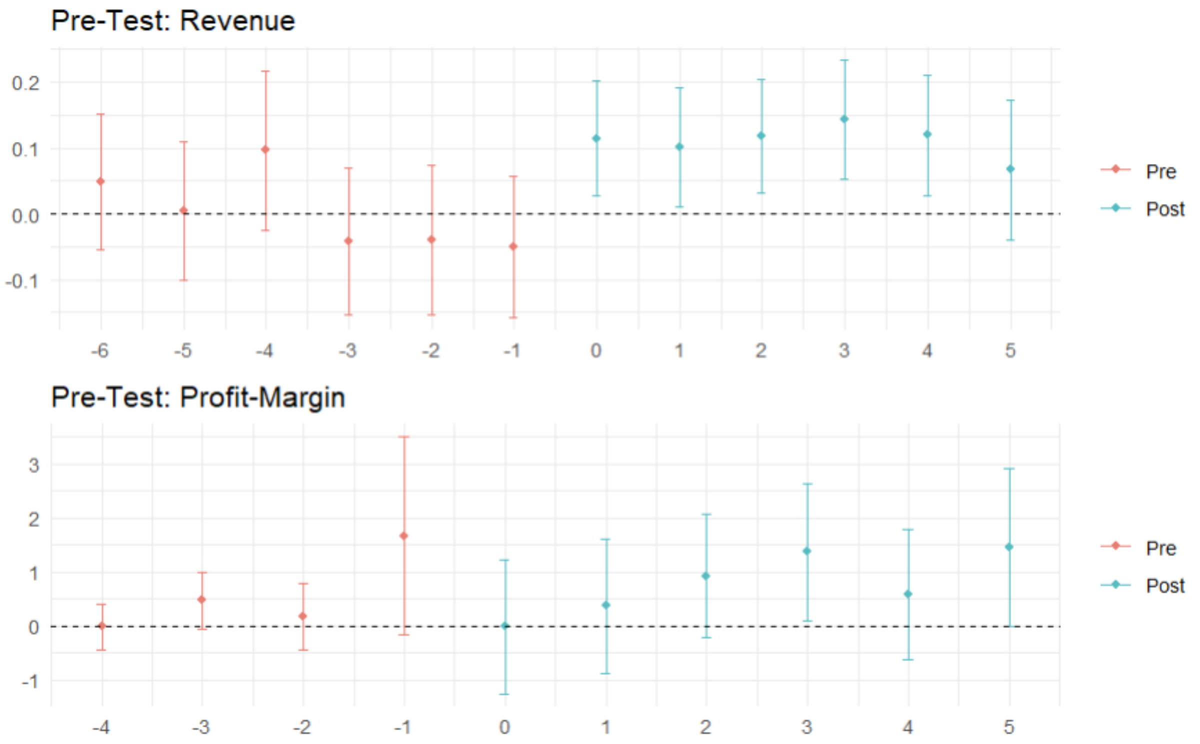


Table 5: *Results for the Robustness test on Revenues (Placebo = 2016)*

Results for the Placebo test (2016) on Revenues

```
Overall summary of ATT's based on event-study/dynamic aggregation:
     ATT    Std. Error     [ 95%  Conf. Int.]
 -0.0182        0.0314    -0.0797      0.0432


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -4   0.0476     0.0361      -0.0499      0.1451
         -3   0.0043     0.0364      -0.0940      0.1027
         -2   0.0962     0.0440      -0.0225      0.2150
         -1  -0.0412     0.0408      -0.1513      0.0689
          0  -0.0404     0.0379      -0.1427      0.0620
          1  -0.0972     0.0411      -0.2084      0.0140
          2   0.0031     0.0367      -0.0961      0.1024
          3  -0.0122     0.0376      -0.1138      0.0893
          4   0.0055     0.0369      -0.0942      0.1051
          5   0.0286     0.0369      -0.0712      0.1283
          6   0.0060     0.0352      -0.0891      0.1011
          7  -0.0392     0.0390      -0.1446      0.0661
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Table 6: *Results for the Robustness test on Revenues (Placebo = 2017)*

Results for the Placebo test (2017) on Revenues

```
Overall summary of ATT's based on event-study/dynamic aggregation:
   ATT     Std. Error     [ 95%  Conf. Int.]
 0.0404        0.0309    -0.0202       0.1011


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -5   0.0476     0.0352       -0.0479      0.1432
         -4   0.0043     0.0365       -0.0947      0.1033
         -3   0.0962     0.0448       -0.0253      0.2178
         -2  -0.0412     0.0402       -0.1502      0.0679
         -1  -0.0404     0.0402       -0.1494      0.0687
          0  -0.0502     0.0380       -0.1532      0.0529
          1   0.0595     0.0333       -0.0308      0.1498
          2   0.0467     0.0329       -0.0426      0.1360
          3   0.0641     0.0318       -0.0223      0.1504
          4   0.0875     0.0353       -0.0082      0.1832
          5   0.0641     0.0340       -0.0282      0.1564
          6   0.0114     0.0390       -0.0946      0.1173
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Table 7: *Results for the Robustness test on Profit Margin (Placebo = 2016)*

Results for the Placebo test (2016) on Profit Margin

```
Overall summary of ATT's based on event-study/dynamic aggregation:
   ATT     Std. Error     [ 95%  Conf. Int.]
 2.3559         0.262     1.8425       2.8693 *


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -2  -0.0143     0.1542       -0.4435      0.4149
         -1   0.4671     0.1896       -0.0605      0.9948
          0   0.1727     0.2292       -0.4651      0.8104
          1   1.7872     0.6697       -0.0764      3.6508
          2   1.8812     0.6888       -0.0354      3.7978
          3   2.4067     0.6401        0.6256      4.1878 *
          4   3.2264     0.6758        1.3460      5.1067 *
          5   3.4278     0.6675        1.5704      5.2851 *
          6   2.2071     0.7047        0.2463      4.1680 *
          7   3.7380     0.9674        1.0460      6.4300 *
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Table 8: *Results for the Robustness test on Profit Margin (Placebo = 2017)*

Results for the Placebo test (2017) on Profit Margin

```
Overall summary of ATT's based on event-study/dynamic aggregation:
    ATT    Std. Error     [ 95%  Conf. Int.]
 2.5637        0.3285       1.92      3.2075 *


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -3  -0.0143     0.1470       -0.4369      0.4084
         -2   0.4671     0.1868       -0.0699      1.0042
         -1   0.1727     0.2282       -0.4835      0.8289
          0   1.6668     0.6806       -0.2900      3.6235
          1   1.7671     0.6682       -0.1539      3.6880
          2   2.3084     0.6159        0.5378      4.0791 *
          3   3.1405     0.6959        1.1397      5.1413 *
          4   3.3365     0.6377        1.5033      5.1697 *
          5   2.1020     0.7247        0.0186      4.1853 *
          6   3.6250     1.0007        0.7481      6.5019 *
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Table 11: *Results for Large Firms: Revenues*

Large Firms (Revenues): Results for Event Study

Revenues (size based on Total Assets)

```
Overall summary of ATT's based on event-study/dynamic aggregation:
     ATT    Std. Error     [ 95%  Conf. Int.]
 -0.0151         0.045    -0.1034      0.0732


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -6   0.0771     0.0619       -0.0953      0.2494
         -5   0.0187     0.0539       -0.1312      0.1686
         -4  -0.0422     0.0635       -0.2190      0.1347
         -3  -0.0127     0.0541       -0.1632      0.1378
         -2  -0.0329     0.0560       -0.1889      0.1230
         -1  -0.0386     0.0556       -0.1934      0.1163
          0   0.0161     0.0567       -0.1417      0.1738
          1  -0.0084     0.0507       -0.1497      0.1328
          2  -0.0310     0.0516       -0.1746      0.1126
          3  -0.0356     0.0562       -0.1920      0.1208
          4  -0.0194     0.0529       -0.1666      0.1278
          5  -0.0121     0.0601       -0.1795      0.1554
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Revenues (size based on Employees)

```
Overall summary of ATT's based on event-study/dynamic aggregation:
    ATT    Std. Error     [ 95%  Conf. Int.]
 0.0509        0.0529    -0.0527      0.1546


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -6   0.0167     0.0557       -0.1338      0.1672
         -5  -0.0266     0.0561       -0.1783      0.1251
         -4  -0.0412     0.0579       -0.1977      0.1154
         -3  -0.0203     0.0604       -0.1835      0.1429
         -2  -0.0244     0.0622       -0.1925      0.1438
         -1  -0.0227     0.0673       -0.2045      0.1592
          0   0.0346     0.0620       -0.1331      0.2023
          1   0.0434     0.0586       -0.1151      0.2019
          2   0.0437     0.0582       -0.1135      0.2009
          3   0.0446     0.0642       -0.1290      0.2181
          4   0.0628     0.0596       -0.0984      0.2240
          5   0.0766     0.0627       -0.0930      0.2461
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Table 12: *Results for Large Firms: Profits*

Large Firms (Revenues): Results for Event Study

Profit Margins (size based on Total Assets)

```
Overall summary of ATT's based on event-study/dynamic aggregation:
   ATT    Std. Error     [ 95%  Conf. Int.]
 0.8192        0.4271    -0.018      1.6564


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -2  -0.0884     0.4046       -1.1624      0.9856
         -1  -0.5547     0.6654       -2.3209      1.2116
          0   0.0387     0.5717       -1.4789      1.5562
          1   0.4512     0.5792       -1.0863      1.9886
          2   0.7228     0.5469       -0.7288      2.1743
          3   1.3929     0.5311       -0.0170      2.8027
          4   1.0194     0.6045       -0.5851      2.6239
          5   1.2904     0.7171       -0.6131      3.1939
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Profit Margins (size based on Employees)

```
Overall summary of ATT's based on event-study/dynamic aggregation:
   ATT    Std. Error     [ 95%  Conf. Int.]
 0.5746        0.4248   -0.2581      1.4072


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -4  -0.4295     0.4431       -1.5722      0.7131
          0  -0.2994     0.6365       -1.9407      1.3419
          1   0.4742     0.6834       -1.2881      2.2366
          2   0.1655     0.6578       -1.5309      1.8619
          3   1.2160     0.9787       -1.3078      3.7398
          4   0.6488     0.8951       -1.6596      2.9572
          5   1.2424     0.8709       -1.0036      3.4884
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Table 13: *Results for Medium-size firms: Revenues*

Medium-size Firms (Revenues): Results for Event Study

Revenues (size based on Total Assets)

```
Overall summary of ATT's based on event-study/dynamic aggregation:
     ATT    Std. Error     [ 95%  Conf. Int.]
 -0.0245        0.0797    -0.1807      0.1318


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -6   0.0454     0.1012        -0.2275      0.3184
         -5  -0.0329     0.1065        -0.3201      0.2543
         -4   0.1232     0.1010        -0.1492      0.3956
         -3   0.0787     0.0897        -0.1631      0.3205
         -2  -0.1059     0.0983        -0.3710      0.1592
         -1  -0.0461     0.0954        -0.3032      0.2110
          0  -0.0344     0.0915        -0.2812      0.2124
          1  -0.0583     0.0883        -0.2963      0.1797
          2  -0.0388     0.0876        -0.2749      0.1972
          3   0.0401     0.0884        -0.1983      0.2784
          4   0.0103     0.0872        -0.2249      0.2455
          5  -0.0657     0.0953        -0.3228      0.1913
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Revenues (size based on Employees)

```
Overall summary of ATT's based on event-study/dynamic aggregation
     ATT    Std. Error     [ 95%  Conf. Int.]
 -0.0649        0.0477    -0.1583      0.0285


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -6   0.1210     0.0724        -0.0807      0.3226
         -5  -0.0007     0.0662        -0.1851      0.1837
         -4   0.0866     0.0756        -0.1239      0.2972
         -3   0.0637     0.0663        -0.1208      0.2482
         -2  -0.0074     0.0609        -0.1769      0.1622
         -1  -0.0490     0.0609        -0.2185      0.1205
          0  -0.0296     0.0618        -0.2017      0.1425
          1  -0.0520     0.0592        -0.2168      0.1128
          2  -0.0763     0.0597        -0.2425      0.0899
          3  -0.0583     0.0620        -0.2309      0.1143
          4  -0.0739     0.0567        -0.2317      0.0839
          5  -0.0994     0.0672        -0.2864      0.0877
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Table 14: *Results for Medium-size firms: Profits*

Medium-size Firms (Profits): Results for Event Study

Profit Margins (size based on Total Assets)

```
Overall summary of ATT's based on event-study/dynamic aggregation:
   ATT    Std. Error     [ 95%  Conf. Int.]
 0.2565        0.8773    -1.4629      1.9759


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -4   0.8996     0.3822        -0.1536      1.9527
         -3  -0.3271     0.5192        -1.7579      1.1036
         -2   0.4901     0.5272        -0.9628      1.9429
         -1  -1.6484     1.2113        -4.9864      1.6895
          0  -0.4076     1.0058        -3.1794      2.3641
          1  -0.3566     1.0215        -3.1717      2.4585
          2   0.4839     1.1120        -2.5806      3.5483
          3   1.2667     1.0889        -1.7341      4.2676
          4  -0.3526     1.1069        -3.4028      2.6976
          5   0.9054     1.3325        -2.7667      4.5776
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Profit Margins (size based on Employees)

```
Overall summary of ATT's based on event-study/dynamic aggregation:
   ATT    Std. Error     [ 95%  Conf. Int.]
 0.8284        0.5619    -0.2729      1.9297


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -4   0.4760     0.3360        -0.4438      1.3959
         -3   0.3846     0.3779        -0.6499      1.4191
         -2  -0.3632     0.4157        -1.5011      0.7748
         -1   0.6320     1.0722        -2.3031      3.5671
          0  -0.0180     0.7285        -2.0121      1.9761
          1   0.3571     0.6998        -1.5584      2.2727
          2   1.2836     0.7670        -0.8161      3.3833
          3   1.5269     0.7145        -0.4291      3.4828
          4   0.7187     0.7953        -1.4583      2.8958
          5   1.1021     0.8124        -1.1218      3.3261
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Table 15: *Results for small firms: Revenues*

Small Firms (Revenues): Results for Event Study

Revenues (size based on Total Assets)

```
Overall summary of ATT's based on event-study/dynamic aggregation:
   ATT    Std. Error     [ 95%  Conf. Int.]
 0.1231       0.0464     0.0322      0.2141 *


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -6   0.0226     0.0582       -0.1404      0.1856
         -5   0.0258     0.0540       -0.1253      0.1770
         -4   0.1271     0.0786       -0.0928      0.3469
         -3   0.0198     0.0837       -0.2145      0.2540
         -2  -0.0559     0.0731       -0.2605      0.1487
         -1  -0.0545     0.0596       -0.2212      0.1123
          0   0.1809     0.0481        0.0463      0.3155 *
          1   0.1494     0.0503        0.0087      0.2900 *
          2   0.1657     0.0514        0.0219      0.3095 *
          3   0.1624     0.0482        0.0277      0.2972 *
          4   0.1099     0.0525       -0.0369      0.2567
          5  -0.0294     0.0976       -0.3024      0.2437
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Revenues (size based on Employees)

```
Overall summary of ATT's based on event-study/dynamic aggregation:
   ATT    Std. Error     [ 95%  Conf. Int.]
 0.1276       0.0463     0.0368      0.2185 *


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -6   0.0028     0.0610       -0.1618      0.1674
         -5   0.0279     0.0580       -0.1288      0.1846
         -4   0.1715     0.0800       -0.0445      0.3874
         -3   0.0176     0.0883       -0.2208      0.2560
         -2  -0.0930     0.0813       -0.3126      0.1266
         -1  -0.0491     0.0683       -0.2335      0.1354
          0   0.1646     0.0472        0.0373      0.2919 *
          1   0.1350     0.0457        0.0116      0.2584 *
          2   0.1649     0.0452        0.0428      0.2871 *
          3   0.1659     0.0506        0.0294      0.3025 *
          4   0.1166     0.0533       -0.0273      0.2605
          5   0.0187     0.0766       -0.1882      0.2255
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Table 16: *Results for small firms: Profits*

Small Firms (Profits): Results for Event Study

Profit Margins (size based on Total Assets)

Overall summary of ATT's based on event-study/dynamic aggregation:

ATT	Std. Error	[95%	Conf. Int.]	
3.6848	0.5976	2.5135	4.8561	*

Dynamic Effects:

Event time	Estimate	Std. Error	[95% Simult.	Conf. Band]	
-4	0.1202	0.2001	-0.4268	0.6672	
-3	0.3821	0.2633	-0.3374	1.1017	
-2	0.2505	0.3103	-0.5976	1.0985	
0	3.3002	1.1273	0.2192	6.3812	*
1	3.3201	1.0473	0.4577	6.1824	*
2	4.2483	1.2313	0.8828	7.6138	*
3	3.2195	1.1289	0.1341	6.3048	*
4	3.8841	1.2194	0.5512	7.2170	*
5	4.1366	1.9262	-1.1279	9.4011	

Signif. codes: `*' confidence band does not cover 0

Control Group: Never Treated, Anticipation Periods: 0
Estimation Method: Doubly Robust

Profit Margins (size based on Employees)

Overall summary of ATT's based on event-study/dynamic aggregation:

ATT	Std. Error	[95%	Conf. Int.]	
4.0817	0.5696	2.9654	5.198	*

Dynamic Effects:

Event time	Estimate	Std. Error	[95% Simult.	Conf. Band]	
-4	-0.1228	0.1992	-0.6726	0.4269	
-3	0.3577	0.2532	-0.3413	1.0567	
-2	0.3263	0.2899	-0.4738	1.1265	
0	2.9837	1.0873	-0.0176	5.9850	
1	2.8925	0.9796	0.1887	5.5964	*
2	4.5752	1.1854	1.3031	7.8473	*
3	4.2103	1.1164	1.1288	7.2918	*
4	4.3314	1.3224	0.6812	7.9817	*
5	5.4968	1.9851	0.0172	10.9763	*

Signif. codes: `*' confidence band does not cover 0

Control Group: Never Treated, Anticipation Periods: 0
Estimation Method: Doubly Robust

Figure 16: *Results for the Complete Case Dataset: Revenues*

Results for Complete Case Data: Revenues

Event Study

```
Overall summary of ATT's based on event-study/dynamic aggregation:
    ATT    Std. Error     [ 95%  Conf. Int.]
 0.0405        0.0295    -0.0173       0.0984


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -6   0.0539     0.0432      -0.0682      0.1760
         -5  -0.0383     0.0438      -0.1620      0.0854
         -4  -0.0689     0.0505      -0.2115      0.0738
         -3   0.0149     0.0448      -0.1116      0.1415
         -2   0.0084     0.0417      -0.1094      0.1262
         -1   0.0491     0.0388      -0.0604      0.1586
          0   0.0343     0.0352      -0.0651      0.1337
          1   0.0013     0.0428      -0.1196      0.1221
          2   0.0534     0.0363      -0.0492      0.1560
          3   0.0882     0.0423      -0.0313      0.2078
          4   0.0416     0.0485      -0.0954      0.1785
          5   0.0245     0.0551      -0.1312      0.1802
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

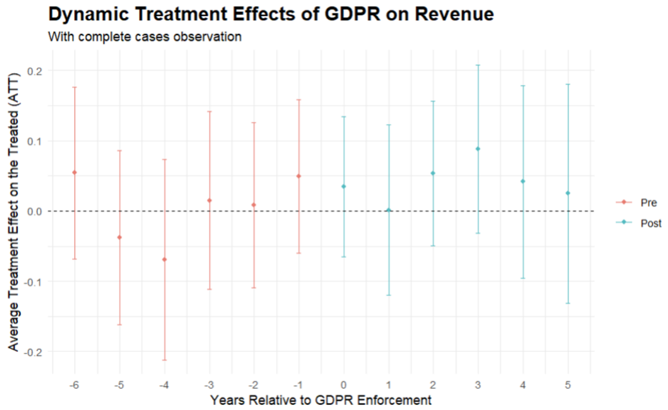


Figure 17: *Results for the Complete Case Dataset: Profit Margins*

Results for Complete Case Data: Profit Margins

Event Study

```
Overall summary of ATT's based on event-study/dynamic aggregation:
    ATT    Std. Error     [ 95%  Conf. Int.]
 0.7303        0.3585     0.0277      1.4329 *


Dynamic Effects:
 Event time Estimate Std. Error [95% Simult.  Conf. Band]
         -4   0.4469     0.4911      -0.9018      1.7957
         -3  -0.2260     0.4622      -1.4955      1.0435
         -2   0.1934     0.4659      -1.0861      1.4729
         -1   0.5425     0.5443      -0.9525      2.0375
          0  -0.0203     0.4859      -1.3548      1.3141
          1   0.0518     0.4564      -1.2016      1.3052
          2   0.8997     0.5038      -0.4840      2.2833
          3   1.1251     0.5002      -0.2488      2.4990
          4   0.7991     0.5205      -0.6303      2.2285
          5   1.5264     0.5998      -0.1209      3.1736
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```