Master Thesis

The effect of the GDPR on EU imports of data-intensive goods

by

ANKE BODEWES

425708

Supervisor

Dr. A. Erbahar

Second reader

Dr. A.I. Figueiredo

Master of Science

INTERNATIONAL ECONOMICS

at the

ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

JULY 14, 2020

Abstract

This paper studies the effect of the General Data Protection Regulation on imports of data-intensive goods by the European Union. We measure the data intensity of products by their direct and indirect exposure to the GDPR. We use monthly data on nearly all products imported into the EU and the US, between January 2014 and December 2020. We perform a difference-in-difference and a triple difference analysis, while controlling for world demand. This design eliminates almost all possible confounders. We argue that our results can only be affected by shocks that influenced data-intensive relative to non-data-intensive imports by the EU and the US differently. We perform our analysis on the aggregate level and per exporter group. We typically do not find a significant change in the difference between data-intensive and non-data-intensive imports in the EU compared to the US. This indicates that the GDPR does not serve as a non-tariff import barrier.

Contents

1 Introduction

In May 2018, the General Data Protection Regulation (GDPR) went into force. This regulation is created to protect the privacy of all citizens of the European Union (EU). The regulation requires every company that processes personal data from EU citizens to implement numerous data protection and privacy measures. From a firm's perspective, this new regulation is costly (Christensen et al., 2013). All companies that deal with personal data in the EU market, must invest in workforce and facilities to comply with the requirements. Moreover, they risk fines of up to 20 million or 4 percent of their annual turnover (*GDPR* 2018). Due to its global enforceability, the GDPR is one of world's most significant privacy regulations (ITGP, 2017). For firms that are based in the EU, these data protection costs are inevitable. On the contrary, non-EU firms can choose to sell to EU customers or to enter other markets, avoiding potential data protection complications. This causes critics to see the GDPR as a non-tariff import barrier, therefore they accuse the European Commission of protectionism (Mishra, 2015; Ferracane and Van der Marel, 2018; Pasadilla et al., 2020; Bauer, Ferracane, Lee-Makiyama, et al., 2016).

In this context, this study answers the question: *What is the impact of the GDPR on EU imports of data-intensive goods?* To answer this question, we regard all products imported by the EU and the US between the months January 2014 to December 2019. On the level of 234 products (BEA, 2012), we determine how exposed a product is to the GDPR using the average of their direct and indirect data intensity. The direct data intensity of a product is the GDPR exposure of the corresponding goods industry. Indirect data intensity is based on GDPR exposure in upstream goods and services industries. To account for the various other factors that could have affected imports, we perform both a difference-in-difference and a triple difference analysis. In the difference-in-difference approach, we compare the development of data-intensive imports to the development of non-data-intensive imports. This absorbs all factors that were time-fixed and all shocks that affected both product groups similarly. In the triple difference approach, we compare the difference between the two product groups in the EU with the difference between the two product groups in the US. The triple difference design eliminates all confounders that affected the difference between data-intensive and non-data-intensive products in the EU and the US similarly. This deals, for instance, with supply effects. We answer our research question at the aggregate level and per exporter group. In the exporter group-specific analysis, we assess countries with different income levels and countries that have an *adequacy decision*, while controlling for world demand. The adequacy decision is a data safety status that the European Commission assigns to specific countries to enhance data-intensive trade (ITGP, 2017).

We typically do not find a significant effect of the GDPR on EU imports of data-intensive goods. Therefore, we conclude that the GDPR is probably not a non-tariff import barrier. Our findings are highly relevant from a privacy policy perspective. Consumer data plays a major role in society today, giving arise to many economic, legal and ethical questions (Meltzer, 2019). This study shows that we should not be too afraid of the trade deterring effects of the GDPR. It indicates that policy makers can protect their citizens' privacy without suffering substantial trade losses. Whether this is an EU GDPR-specific finding or a more general privacy policy finding is an interesting topic for future research.

In terms of academic relevance, this study is one of the first ex-post empirical studies into the effect of privacy regulation on trade flows. Academics that studied this relation before, derived trade effects from changes in Total Factor Productivity (TFP), neglecting other important mechanisms that could affect trade (such as Bauer, Lee-Makiyama, et al., 2014). Moreover, other authors looked at data restrictive policies in general, of which the impact is not necessarily similar to that of the GDPR in specific (such as Van der Marel et al., 2016). Third, almost all existing GDPR-specific studies regard a particular sector (such as Goldberg et al., 2019). We look at nearly all products. Finally, we are the first in privacy research that use multiple dimensions to control for confounders. Other authors looked merely at an EU versus US difference over time (such as Jia et al., 2018) or a data-intensive versus non-data-intensive difference over time (such as Yuan and J. Li, 2019). We use both types of variation, greatly decreasing the set of possible omitted variables.

In the literature section, Section 2, we start with a discussion on the quantification of non-tariff measures. Then, we go into the legal and economic characteristics of the GDPR and discuss related studies. The data section, Section 3, elaborates on the assignment of direct and indirect data intensity, the collection of trade data and the summary statistics of our data set. In the methodology section, Section 4, we show our difference-in-difference and triple difference specifications and discuss their methodological features. The results section, Section 5, presents the results of our main models and robustness checks, on the aggregate and exporter group-specific level. In Section 6 we conclude, list potential endogeneity issues and offer suggestions for future research.

2 Literature

2.1 Non-tariff measures

2.1.1 GDPR as non-tariff measure

Before we discuss the literature on data protection regulations, we regard a more general field of study. The relation between international trade and non-tariff measures (NTMs), more specifically technical (TBTs) and sanitary/phytosanitary (SPSs) measures. It is intuitive to see the GDPR as a TBT, because it sets technical regulations and standards that may in effect create obstacles to international trade (*TBT Agreement* 1995, art. 2). Moreover, several economic authors argue that data protection regulations are *de facto* non-tariff trade barriers (Mishra, 2015; Ferracane and Van der Marel, 2018; Pasadilla et al., 2020; Bauer, Ferracane, Lee-Makiyama, et al., 2016). A question that arises is whether the GDPR is a TBT from a legal perspective. We conclude that this is not yet the case. Firstly, because we find no notification of any data protection related measure in the World Trade Organization's TBT notification database (WTO, 2020). Secondly, only two specific trade concerns were raised on this topic, both on a particular Chinese cyber security law (WTO, 2020). As many countries have data protection laws, we would expect many more specific trade concerns if data protection measures were TBTs. Thirdly, legal and governmental authors write that WTO trade law has potential to regulate digital trade in the future, but that the legal frameworks are currently not ready to fulfill this role (Stone et al., 2015; Meltzer, 2019; Aaronson and Leblond, 2018).

2.1.2 Quantification of non-tariff measures

The literature on NTMs, in particular on TBTs and SPSs, developed rapidly since the 1950s (Ronen, 2017). We discuss the relevant elements of ex-post empirical research in this field. The literature identifies two potential effects of those measures on trade. An import enhancing effect due to the quality assurance that they provide and an import deterring effect due to their impact on the intensive and extensive margins of trade. Empirical literature suggests that therefore the overall trade effect depends on the type of measure, the product or sector, the size of the exporting firms and the affected countries. The estimated effects also appear to depend heavily on the methodological approach and the assumptions of the researcher (Ronen, 2017).

The main methods for quantifying NTMs are the inventory, price-comparison and quantity-impact approach (Fugazza, 2013). The most common goal of those quantification measures is to translate their estimates into price-effect estimates and ad valorem changes in production costs. Among the inventory measures, two NTM indicators exist: the frequency index and the coverage ratio. The frequency index summarizes the percentage of a country's import products on which an NTM applies. The disadvantage of this approach is that it only regards the incidence and not the importance of NTMs. The coverage ratio measures the percentage of a country's imports subject to an NTM. Although this ratio does assess the importance of the NTMs, it disregards the information that can be obtained from bilateral trade flows (Fugazza, 2013; Crivelli and Groeschl, 2015). The price-comparison approach looks at the gap between the domestic and international price of a specific good. Major caveats of this measure are that it assumes that domestic and international goods are perfect substitutes and that the gap is not affected by other market dynamics than the NTM (Ronen, 2017).

The methodology of this study is partly based on the recently most prevalent method: the quantity-impact approach. In this method, researchers predominantly apply gravity type models to estimate the impact of policy measures on trade. Some studies combine the different approaches and employ frequency and coverage measures in their gravity models (Korinek et al., 2008). Important methodological takeouts are the attention for heteroskedasticity, zero trade flows (Silva and Tenreyro, 2006) and multilateral resistance (Bratt, 2017; Silva and Tenreyro, 2006; Y. Li and Beghin, 2017).

2.2 Privacy regulation

2.2.1 Scope of the GDPR

In order to estimate economic consequences of the GDPR, we explore the legal framework of the regulation. Before the GDPR, privacy rules in EU member states were subject to the Data Protection Directive 1995 (*DPD* 1995). The most important difference between the GDPR and the DPD, is that the former is a regulation and the latter a directive. In the European legislative system, a directive only provides minimum results to be achieved, giving member states the flexibility to adopt and enforce their own rules. In contrast, a regulation is directly legally binding in all member states, leaving no room for national flexibility (Seo et al., 2018).

The DPD set relatively low minimum privacy standards, and member states implemented a variety of different levels of privacy protection (Bauer, Ferracane, and Van der Marel, 2016). The unification of European privacy law serves two key goals: i) protecting the rights, privacy and freedoms of natural persons in the EU and ii) reducing barriers to business by facilitating the free movement of data throughout the EU (ITGP, 2017). The European Commission estimated ex-ante that this latter goal would be worth 2.3 billion euro per year (Wigand and Voin, 2018). The GDPR is the world's most significant data protection law because its application reaches far beyond EU member states. It applies to every company in the world that deals with personal data of EU citizens, meaning that non-EU companies must either comply or not do business in the EU at all (ITGP, 2017).

The GDPR establishes a set of rights for individuals, which naturally results in a set of obligations for organizations that collect, store and process personal data. The regulation distinguishes between two roles that an organization can have in processing personal data. The *controller* is the party that determines the purpose and the type of data processing. The *processor* is the party that processes personal information on behalf of the controller (*GDPR* 2018, art. 4). In practice, organizations often fulfill both roles, depending on the type of activity. Both data processors and controllers are expected to fully understand the rights set out in the GDPR and adapt their practices in accordance. The expansion of the GDPR in comparison to the preceding DPD is characterised by the following items:

- *Sanctions.* Companies risk fines of up to the greater of 4 percent of *global* annual turnover or 20 million euro (*GDPR* 2018, art. 83).[1] In addition, individuals have the right to request financial compensation if an organization breaches their privacy rights (*GDPR* 2018, art. 82).

- *Definition personal data.* In the new regulation, personal data is *anything* that could be linked in *any way* to an individual (*GDPR* 2018, art. 4). This is a much broader definition than in the DPD, causing more economic activities to fall within the scope of the regulation (Seo et al., 2018).

- *Fair processing.* Organizations should be transparent and concise about what personal data is processed, the legal ground of the processing and whether it is processed by a third party processor. An important aspect of the legal ground criteria, is that many types of processing may only take place after specific, freely given and informed consent by the individual (*GDPR* 2018, art. 12).

- *Requests of individuals.* Data controllers must respond without undue delay to a request of an individual to access, rectify or remove all his or her personal data (*GDPR* 2018, art. 15, 16 and 17). Related to this are the right to request copies of personal data in easily accessible formats and the right to object to personal data processing.

- *Technical requirements.* While the GDPR does not explicitly prohibit data breaches, cyber security is a major focus of the regulation. Organizations must take numerous physical

[1] Until now, 200 fines were imposed worth 144 million in total. The largest fine was for Google Inc., being 50 million euro (Privacy Affairs, 2020).

and administrative measures to secure all personal data against loss and damage (ITGP, 2017).

- *Data localization.* Organizations may only process personal data outside of the EEA if one of the following conditions is applicable: i) the destination country is subject to an adequacy decision of the European Commission, meaning that it is one of the currently 13 countries that are regarded sufficiently privacy friendly, or ii) the firm in question made sure that the data transfer is subject to one of the seven possible legal safeguards (*GDPR* 2018, art. 45 and 49).[2]

- *Data protection officer.* To ensure monitoring of above obligations, firms must appoint a data protection officer that is an expert in the GDPR (*GDPR* 2018, art 37).

2.2.2 Impact on firms and consumers

The changes in the European privacy law are expected to have substantial impacts on the both firms and consumers.[3] The literature identifies several types of costs for firms related to the new obligations. First, human resource costs for a data protection officer and training for employees to deal with requests from individuals and cyber security related tasks.[4] This cost is relatively independent from output, and can in most cases be viewed as a fixed production cost. Second, IT capital costs to achieve a data infrastructure that complies with the technical requirements and data localization rules of the GDPR. This is also likely to affect fixed costs. Third, the costs of external legal and technical services, simply because many companies do not have the required expertise in-house. This mainly affects fixed costs as well. Fourth, TFP goes down if fair processing obligations and technical requirements disrupt or slow down normal operations, increasing marginal costs (Koski and Valmari, 2020; Bauer, Ferracane, and Van der Marel, 2016). Fifth, an expected cost for the risk to be fined based on the revenue of the company (Lee et al., 2019).

Those various types of costs may have two consequences for supply. The increased fixed costs may prevent EU and non-EU firms from entering the market (affecting the extensive margin) and the increased marginal costs may increase prices and therefore lower demand (affecting the intensive margin) (UNCTAD, 2013; Ronen, 2017). Those two mechanisms both lead to a fall in trade in data-intensive goods and services. Important to note is that there could also be beneficial effects for supply, as the unified regulation for all member states is easier to work with than the various national rules. This is mainly the case for intra-EU trade. After all, the GDPR is partly introduced to promote the free flow of data within EU member states. However, literature suggests that the economic costs are more prevalent than the benefits, for both EU and non-EU firms (Bauer, Ferracane, and Van der Marel, 2016; Ferracane and Van der Marel, 2018; Koski and Valmari, 2020).

Applying the theory from NTM literature on the GDPR, we suggest that the regulation may

[2]The US and the EU have a privacy shield that is supposed to facilitate cross-border data flows. However, the legal power of the privacy shield is much more limited than the adequacy decision (ITGP, 2017).

[3]Christensen et al. (2013) estimated ex-ante that the average SME in Europe would spend between 16 and 40 percent of annual SME IT budgets on compliance.

[4]The average annual salary of a data protection officer is around 80.000 euro in Europe (Koski and Valmari, 2020).

affect consumer behavior as well. On the one hand, the regulation can convince downstream companies and individuals of the adequate privacy level of suppliers, increasing demand. On the other hand, demand may decrease if purchasing companies and individuals are more aware of the risks involved in personal data processing (Goldberg et al., 2019). Another plausible dynamic is that consumers trust European suppliers more in terms of privacy than non-European suppliers, causing trade to shift from extra- to intra-EU trade. The demand effect of the GDPR is therefore expected to be ambiguous.

2.2.3 Literature on data protection

We discuss three types of economic data protection literature: i) the effects of privacy regulation in general, ii) ex-ante GDPR studies and iii) ex-post GDPR studies. Economic privacy literature started with mainly theoretical studies, modelling both market inefficiencies (Posner, 1981; Stigler, 1980) and welfare improving consequences (Hirshleifer, 1971; Spence, 1973). More recently, empirical studies suggest that privacy restrictions can have negative economic impacts on advertising and e-commerce industries (Swire and Litan, 1998; Goldfarb and Tucker, 2011), health industries (Miller and Tucker, 2009) and credit markets (Acquisti et al., 2016). Several studies regard the impact of data protection on cross-border trade. Bauer, Ferracane, Lee-Makiyama, et al. (2016) use a general equilibrium model to estimate that the EU would gain 8 billion euro per year GDP gain in intra-EU trade if the data localization measures that were in force before the GDPR were removed. In contrast, Goldfarb and Trefler (2018) argue that international trade can increase due to privacy regulations. They propose that young companies in the field of artificial intelligence have more opportunity to grow in less data restrictive countries. After their domestic growth, they have the resources to export to countries with stricter policies. Several papers measure the economic costs of restrictions on the free flow of data (Bauer, Lee-Makiyama, et al., 2014; Bauer, Ferracane, and Van der Marel, 2016; Van der Marel et al., 2016). These papers all follow the same methodology. First, they identify observable regulatory barriers to cross-border data flows and link them to downstream TFP on the industry level. Then they use their estimated change in TFP in a general equilibrium model to calculate losses in GDP, production, imports and exports. These academics conclude that the effect on trade is negative. In a later paper, Ferracane and Van der Marel (2018) designed the DTRI Trade Restrictiveness Index. The index weighs data policies of countries by the downstream data intensity in services sectors and distinguishes between policies targeting cross-border movement of data and policies focusing on domestic data use.

In terms of ex-ante GDPR research, Lee et al. (2019) use a computable general equilibrium model to estimate the welfare effects of the GDPR on the EU and on South Korea. They expect substantial welfare losses due to additional trade cost that may serve as trade barriers. In contrast, Ciriani (2015) suggests that the GDPR will promote trade because it contributes to a level playing field between EU suppliers and US competitors in the European market. Christensen et al. (2013) use expected direct costs of the GDPR for SMEs, being between 16 and 40 percent of current annual SME IT budgets, to simulate impacts on business and job creation. They show a substantial negative impact that is most severe in those sectors where compliance with the GDPR implies higher fixed costs for firms. Bauer, Erixon, et al. (2013)

take on a similar approach and estimate a decrease in EU services exports by 6.7 percent and a decrease in EU goods that rely on data-intensive services of 11 percent. Sobolewski and Paliński (2017), assess the potential welfare gain from privacy intervention with a survey among 143 respondents. They estimate a welfare gain from the GDPR of 6.50 euro per capita per month and interpret this as a proof of the existence of demand for privacy.

Recently, a small set of ex-post GDPR studies has emerged. Jia et al. (2018) compare venture capital investments in new and emerging technology firms in the EU and the US in a difference-in-difference approach. They find that the GDPR affected EU investments negatively. Yuan and J. Li (2019), also use a difference-in-difference model and estimate a negative impact on the financial performance of hospitals that provide digital health services as their primary business. Goldberg et al. (2019) find a 10 percent drop in web traffic and e-commerce sales after the out-roll of the GDPR. Koski and Valmari (2020) use firm level data in a difference-in-difference approach to explore short term profitability effects of GDPR compliance. They find that the profitability of European data-intensive SMEs decreased more than the profitability of data-intensive US SMEs. They did not find this difference for very large data-intensive firms. Finally, Jia et al. (2019) demonstrate that the GDPR has a greater negative effect on EU ventures financed by non-EU investors than those financed by EU investors.

Our study adds to the ex-post empirical GDPR literature in three ways. Most importantly, to our knowledge we are the first to empirically study the relation to trade flows in a direct sense. Other authors merely derived trade effects from changes in TFP, possibly neglecting other important mechanisms that affect trade (Bauer, Lee-Makiyama, et al., 2014; Bauer, Ferracane, and Van der Marel, 2016; Van der Marel et al., 2016; Ferracane and Van der Marel, 2018). Moreover, they looked at data restrictive policies in general, of which the impact can be very different than that of the GDPR in specific. Second, almost all GDPR-specific studies so far (except Ferracane and Van der Marel (2018)), focused on one particular industry. We look at all data-intensive industries. Finally, we are the only ones that use multiple dimensions to control for large sets of confounders. We do not merely look at an EU versus US difference or a data-intensive versus non-data-intensive difference, but we use both types of variation.

3 Data

3.1 Data intensity of products

To measure the effect of the GDPR on international trade, we are interested in the extent to which products are exposed to the regulation. Because all companies that process personal data must comply with the GDPR, we compute a measure for the *data intensity* of products. Our measure is the average of two types of data intensity: direct and indirect data intensity. The direct measure for data intensity is the industry level data intensity that is available on the NACE Rev-2 level (BEA, 2012). We discuss this in detail in Paragraph 3.1.1.

Our indirect measure is based on Bauer, Lee-Makiyama, et al. (2014). We measure the indirect data intensity of a downstream product by the weighted sum of the direct data intensity of all its upstream inputs as in Eq. (1).

$$\text{Indirect Data Intensity}_p = \sum_{k=1}^{K} \text{Data Intensity}_k \times \text{Input}_{kp} \tag{1}$$

In Eq. (1), Indirect Data Intensity$_p$ is the data intensity of a downstream product p, Data Intensity$_k$ is the data intensity of an upstream industry k and $Input_{kp}$ is the input share of industry k in product p. This method relies on the principle that downstream industries are affected by trade policy in the corresponding upstream industries (Arnold et al., 2012). From a data protection perspective, this means that downstream products which use more data-intensive upstream goods or services, are more affected by data regulations. An illustration is a software-product that depends strongly on financial and advertising services that process personal data. Higher costs and slower processes in those upstream services may be passed on to downstream products. Other examples are products that depend on personal data processing because a large share of their sales comes from e-commerce and consumer analytics (Goldfarb and Tucker, 2011). We assume that these types of products are more affected by the GDPR than products that do not rely on upstream data-intensive industries. With this approach, we account for the fact that the GDPR is likely to affect trade in goods through upstream services. Moreover, this method can systematically assign sector-level upstream data intensity information to detailed downstream product data.

3.1.1 Direct data intensity

We measure direct data intensity by averaging two indicators: *percentage of enterprises that have documents on measures, practices or procedures on ICT security* (Eurostat, 2019b) and *percentage of enterprises selling at least 1 percent of their turnover online* (Eurostat, 2019a). The first indicator is directly linked to the GDPR, because the regulation requires companies to document many measures and procedures related to data protection. Thus, this indicator measures the percentage of companies within a group of NACE Rev-2 industries that perceives itself as subject to the GDPR. The second indicator gives an image of the percentage of companies per industry that processes personal data due to online sales. This includes sales via online marketplaces or outsourced payments services. In most cases, online marketplace or payment services are data processors that process data on behalf of the supplying firm (ITGP, 2017). As shown in Table A in the appendix, the indicators vary substantially per NACE industry group. Highly data-intensive industries in terms of security measures are computer related services, professional, scientific and technical activities and telecommunications. The e-commerce indicator is particularly high for accommodation and publishing activities. For NACE industries in mining, fishing and hunting (NACE 1-9) the indicators are not available. We discuss this possible measurement error in Section 6.

3.1.2 Downstream assignment to products

We assign upstream data intensity to downstream products via an industry-to-commodity input-output table. BEA (2012) provides the only available input-output matrix that contains the industry-level inputs used in 405 commodities, of which 234 are products. First, we assign our NACE level data intensity to the 71 upstream BEA goods and services industries. Then we assign the NACE industry level upstream data intensity to the 234 BEA products. The assignment is based on industry input shares: the share of production that is directly or indirectly required to produce 1 dollar of a commodity for final use (BEA, 2012). An important note is that the goods that have data-intensive inputs may also be inputs for other final products. This indirect production is accounted for in the input-output table.

The BEA table is the only input-output table that is detailed enough for our purposes. Still, two remarks must be made. First, this input-output matrix is based on US input-output. We are also interested in input-output levels in the EU, which might not be similar to the US. However, just like Bauer, Lee-Makiyama, et al. (2014) we use this information merely as a benchmark. It reflects a typical economy, not a country or year in specific. A second concern is that the latest available year of this data is 2012. Some products that did not use data-intensive inputs in 2012 may have started to use those inputs by now. We discuss this further in Section 6. The advantage of using 2012 instead of more a more recent year, is that this helps eliminating reverse causality problems that could arise if the scope of the GDPR was focused on highly traded products.

3.2 Imports

We obtain import data on the six-digit Harmonized System (HS) product level over the months January 2014 to December 2019 from the Comtrade database (United Nations, 2020). For our first part, we collect aggregate imports from the rest of the world by the EU[5] and the US. For our second part, we extract exporter-specific import data from the EU and the US, covering all 211 countries that exported to the EU or the US according to the United Nations (2020). We request trade data from 5336 six-digit HS products. This corresponds with the around 5300 existing six-digit HS product codes United Nations (2020). Of those products, Comtrade provides data on 4924 six-digit HS products that were imported by the US or the EU in our time span. We aggregate the imports to the BEA product level.

In order to deal with implicit zero trade values, we complete the aggregate data with a zero trade value for each missing month-product combination (we add 441 zeros). The exporter-specific data we complete with zero trade values for each month-product combination (we add 529354 zeros). Finally, we add an income class for 184 exporting countries (WorldBank, 2020).[6] We aggregate the exporter-specific imports per income group. We do this as well for the exporters with an adequacy decision (European Commission, 2020).

3.3 Summary statistics

Table 3.3 shows the summary statistics of our data. We show the aggregate data, the data per exporter income group and the data for the exporters with an adequacy decision. Each group consists of 33,696 observations: imports of 234 BEA product groups in 72 months by the US

[5] We use the EU-28, as this is the union that was together in our full time frame.

[6] This means we lose 27 small exporters. The corresponding measurement error is not expected to be large, as the trade values of those exporters are low.

and the EU. Because Eurostat (2019a) does not provide the data intensity of NACE industries 1-9, direct data intensity is not available for 20 BEA product groups. The import means are interpreted as the average imports by the EU and the US, of a BEA product group per month from an exporter group in US dollar. The highest maximum trade value comes from middle high income countries, which might be explained by the presence of several oil countries in this group.[7] We have 32 low income, 48 middle low income, 58 middle high income and 46 high income countries. Our sample contains all nine larger countries with an adequacy decision.[8] The data intensity variables are distributed equally over the groups because we added implicit zero trade values, resulting in a balanced data set.

Table 1: Summary statistics

	N	Mean	St. dev.	Min	Max
Aggregate					
Imports (in US$)	33,696	694,642,712	1,834,548,515	0	41,266,141,916
Indirect data intensity	33,696	24.1	5.3	6.5	33.8
Direct data intensity	30,816	27.6	5.0	19.5	36.0
Low income					
Imports (in US$)	33,696	953,909.7	7,371,691	0	287,902,559
Indirect data intensity	33,696	24.1	5.3	6.5	33.8
Direct data intensity	30,816	27.6	5.0	19.5	36.0
Middle low income					
Imports (in US$)	33,696	17,508,765	150,391,591	0	6,193,110,762
Indirect data intensity	33,696	24.1	5.3	6.5	33.8
Direct data intensity	30,816	27.6	5.0	19.5	36.0
Middle high income					
Imports (in US$)	33,696	99,935,136	724,286,896	0	21,302,336,833.
Indirect data intensity	33,696	24.1	5.3	6.5	33.8
Direct data intensity	30,816	27.6	5.0	19.5	36.0
High income					
Imports (in US$)	33,696	55,910,951	445,620,514	0	14,155,298,218
Indirect data intensity	33,696	24.1	5.3	6.5	33.8
Direct data intensity	30,816	27.6	5.0	19.5	36.0
Adequacy					
Imports (in US$)	33,696	8,203,027	52,008,581	0	2,453,424,280
Indirect data intensity	33,696	24.1	5.3	6.5	33.8
Direct data intensity	30,816	27.6	5.0	19.5	36.0

Note : Import values are the average imports of a BEA product, aggregated per group.

[7] Such as Iraq, Iran, Russia and Venezuela.

[8] We regard Japan, Canada, Israel, Argentina, New Zealand, Switzerland, Uruguay, Andorra and the Faeroe Islands. We do not regard Isle of Man and Guernsey because they are not in our Comtrade data.

4 Methodology

Our analysis consists of two parts: a study on aggregate imports by the EU from the rest of the world and a study on imports per exporter group. We perform both a difference-in-difference and a triple difference estimation.

4.1 Aggregate

4.1.1 Difference-in-difference

We first compare the change in EU imports of data-intensive products to that of non-data-intensive products. This difference-in-difference approach controls for all product group-specific, time-fixed variables and all time-specific, product group-fixed variables (Wooldridge, 2007). An example of the former is that the EU always imports a larger value of laptops than books. An example of the latter, is that an economic shock increases demand for both laptops and books equally. Our specification is shown in Eq. (2).

$$lnM_{tp} = \beta_0 + \beta_1 T_t + \beta_2 DI_p + \beta_3 T_t \times DI_p + e_{tp} \tag{2}$$

In Eq. (2), lnM_{tp} is BEA level product p imports by the EU from the rest of the world in month t. Imports are log-linearized, so that we can interpret the coefficients in percentage point import differences. T_t is the post treatment dummy. We do our estimations separately using the GDPR announcement in May 2016 and the coming into force of the GDPR in May 2018. DI_p is average of the direct and indirect data intensity of a product. β_3 is our coefficient of interest, measuring the difference between the trade development of data-intensive versus non-data-intensive products. e_{tp} is the product and time-specific error term.

4.1.2 Triple difference

An important issue with the causal interpretation of Eq. (2), is the assumption that trade in data-intensive and non-data-intensive products would have developed similarly in absence of the GDPR. This assumption is not plausible, as we can expect that data-intensive trade increased in the past decade in relation to the global spread of technology. Additionally, there could be other factors that influenced relative data-intensive imports. Therefore, we continue with a triple difference approach. In addition to the treatment variation per product, we add a non-treated region: the US. We use the US as the control group, because this is another large and developed region with a comparable commercial environment. Although the US did implement a new data protection law in 2018, this law is much less stringent than the GDPR (ITGP, 2017). Next to the characteristics of the difference-in-difference, the triple difference absorbs time-varying effects that are specific to trade in data-intensive goods but similar for the US and the EU. An important example is a supply effect from exporting countries.

$$\begin{aligned} lnM_{itp} =& \beta_0 + \beta_1 T_t + \beta_2 DI_p + \beta_3 EU_i + \beta_4 T_t \times DI_p \\ &+ \beta_5 T_t \times EU_i + \beta_6 DI_p \times EU_i + \beta_7 T_t \times DI_p \times EU_i + e_{itp} \end{aligned} \tag{3}$$

In Eq. (3), we add a dummy that has the value one for the EU as importer i and zero for the US as importer. The dependent variable and the error term become importer-, time- and product-specific. The coefficient of interest is β_7, measuring the difference in the development of data-intensive compared to non-data-intensive products in the EU versus the US.

4.2 Exporter-specific

4.2.1 Difference-in-difference

It is not necessarily the case that trade from all exporters responds similarly to the GDPR. Literature suggests that developing countries could be more vulnerable to non-tariff measures (UNCTAD, 2013). Firms in developing countries could be less able to comply with the regulation due to resource constraints. On the other hand, firms in those countries could be more ignorant towards the laws than those in developed countries. Especially famous companies from the US might be more afraid of reputational damage. Similarly, trade enhancing effects might differ across exporters. Compliance with the GDPR can increase trust in privacy behaviour of foreign firms from developed countries, increasing imports from those firms. We therefore continue our analysis with an exporter group-specific approach. We also use this approach to assess whether the adequacy decision of the European Commission has the intended trade enhancing effects.

Our specification for this approach looks similar to Eq. (2). The difference is that the dependent variable and the error term are exporter-, product- and time-specific. Additionally, we control for demand effects by including the world demand for a product by the EU in a month. We aggregate the imports over the exporter groups: low income, middle low income, middle high income and high income countries and exporters with an adequacy decision.

4.2.2 Triple difference

In our exporter-specific difference-in-difference model, we deal with the same interpretation issue as in the aggregated estimation. We do not know whether we observe a change in trade due to the GDPR or due to other product group-specific shocks. Important factors in this respect are supply shifts. If the production and use of data-intensive technologies shifts from developed to emerging countries, we observe a negative change in the former and a positive change in the latter. This is problematic for our interpretation, as this supply effect has not much to do with the GDPR. Therefore, we will again compare EU imports to US imports in a triple difference approach. This is Eq. (3), but with importer-, exporter-, product- and time-specific imports and error terms. We also control for importer-, product- and time-specific world demand.

4.3 Estimation features

The NTM literature addresses several issues in estimating the effect of policy measures on trade. Firstly, in its log-linear form, our estimations may be biased due to zero trade values. A zero trade value could imply that the costs to enter a market are prohibitively large. To deal with this issue, we follow the commonly used Poisson pseudo-maximum-likelihood (PPML) method (Silva and Tenreyro, 2006). This method is also robust for heteroskedasticity of the error term (Y. Li and Beghin, 2017). A second issue in international trade research is the *multilateral resistance* between trading pairs. Multilateral resistance are importer specific trade costs distortions faced by the importer towards all exporters and the exporter-specific trade costs distortions faced by exporter i in all import markets (Y. Li and Beghin, 2017). Researchers control for those terms by including several types of importer or exporter fixed effects. Our triple difference method automatically controls for most of the possible multilateral resistance, as exporter- and importer-specific characteristics that do not vary between the EU and the US cancel out.

We discuss several assumptions of our approach. In order to obtain an unbiased estimation of the treatment effect, β_7 in Eq. (3), the correlation between error term e_{ijtp} and our triple interaction term must be zero. This condition is relaxed by the elimination of product group-fixed effects, time-fixed effects and time-varying product group-varying effects that are equal in the US and the EU. Related to this, is the common trends assumption (Wing et al., 2018). In our research design, the common trends assumption implies that without the GDPR, the difference in imports of data-intensive and non-data-intensive products should have developed similar in the EU and the US. This assumption is quite plausible. A violating factor must affect the difference between imports from data-intensive products in the EU differently than in the US. We attempt to test the common trend assumption with a graphical assessments of the development of the two product groups in the two countries and fake treatment tests around 2015. The second difference-in-difference assumption is strict exogeneity, meaning that the timing of the applicability of the GDPR on a product must be independent of the imports of this product. This assumption could be violated, because high imports from data-intensive goods could have encouraged the European Commission to establish a data protection regulation. By regressing our independent variable on several lags of lnM_{ijtp}, we see whether imports predict the out-roll of the GDPR. If that is the case, strict exogeneity is violated (Wing et al., 2018). Additionally, we perform a Granger causality test by estimating the triple difference version of Eq. (3) with several leads of $GDPR_{ijp}$. If the corresponding coefficients are insignificant, this means that future GDPR applicability is not associated with current imports (Wing et al., 2018).

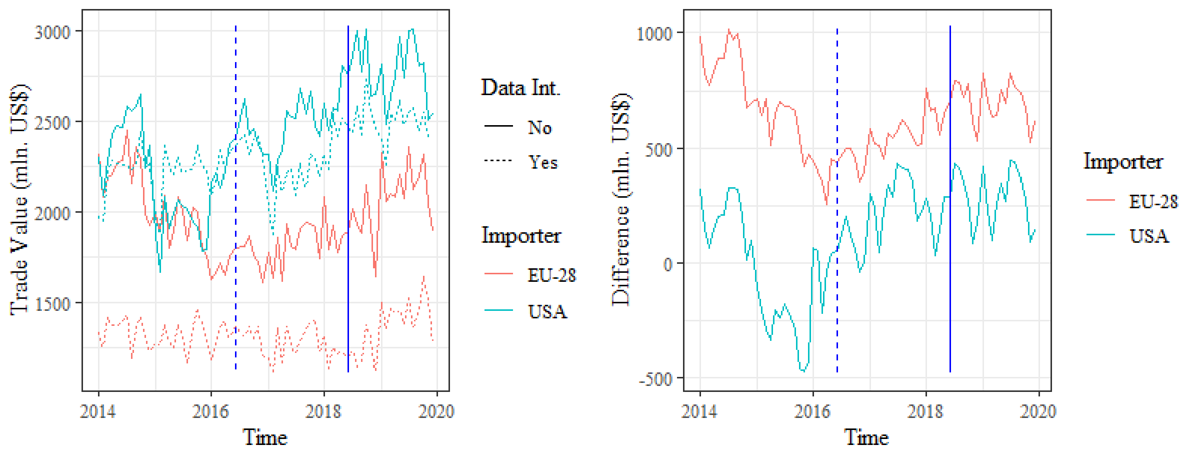
5 Results

5.1 Aggregate

5.1.1 Difference-in-difference

Figure 1 Panel (a) shows the development of average monthly imports per BEA product for the EU and the US. For the purpose of this graph, we divide the products in a data-intensive and a non-data-intensive group, based on above or below median data intensity. Although we cannot show this with continuous data intensity, this gives an image of what we are measuring in the difference-in-difference model. The vertical solid line is the coming into force of the GDPR, the vertical dotted line the announcement. Our coefficient of interest represents the difference between the pink and light blue solid and dotted lines before and after the treatment. We do this estimation separately for the EU and the US. The figure shows that overall, the US on average imported more of both product groups than the EU. The US also experienced a stronger increase in this product group. Figure 1 Panel (a) indicates that we cannot assume a common trend for the difference between data-intensive and non-data-intensive products in the absence of the GDPR. The two product groups did not seem to develop similarly before the treatment.

Figure 1: Aggregate



(a) Difference-in-difference

(b) Triple difference

Note : Panel (a) is the development of above and below median data-intensive and non-data-intensive products. Panel (b) is the development of the difference between above and below median data-intensive and non-data-intensive products. The dotted vertical line is the announcement and the solid vertical line the coming into force of the GDPR. Data intensity is the average of direct and indirect data intensity.

Table 2 shows our main difference-in-difference PPML regression results. All four models regard the months January 2014 to December 2019. Model 1 and 3 use the announcement month of the GDPR as the treatment month. Model 1 and 2 look at the EU and Model 3 and 4 assess the US. All models have robust standard errors to account for hetroskedasticity. None of the models finds a significant change in the difference between the two product groups. This indicates that the imports of the two product groups grew at a similar pace in both importing regions. We find this somewhat surprising, as we would expect an increasing trend in relative data-intensive imports.

5.1.2 Triple difference

Figure 1 Panel (b) represents the triple difference research design. Per importer, we show the difference between non-data-intensive and data-intensive imports. The triple difference coefficient of interest is the difference between the pink and the light blue line before the treatment minus this difference after the treatment. The EU always imported more non-data-intensive products than data-intensive products. The US's difference moves around zero, meaning that relative data-intensive imports were first growing and later falling.

Table 3 shows the aggregate triple difference results. The number of observations is twice that of the difference-in-difference analysis, because we now look at the EU and the US together. Model 1 uses May 2016 as the treatment month, Model 2 uses May 2018 as the treatment month. Model 3 and 4 do the same, but use a below or above median data intensity dummy. None of

Table 2: Aggregate: difference-in-difference

	Dependent variable: log Imports (in US $)			
	Model 1	Model 2	Model 3	Model 4
Intercept	17.919***	17.780***	17.650***	17.607***
	(0.164)	(0.125)	(0.178)	(0.132)
Post	-0.271	-0.088	0.001	0.154
	(0.219)	(0.254)	(0.226)	(0.238)
Data Int.	0.077***	0.081***	0.096***	0.098***
	(0.006)	(0.005)	(0.007)	(0.005)
Post × Data Int.	0.010	0.004	0.003	-0.000
	(0.008)	(0.009)	(0.009)	(0.009)
Num. obs.	15408	15408	15408	15408
Importer	EU	EU	US	US
Treatment month	May 2016	May 2018	May 2016	May 2018
Start sample	Jan 2014	June 2017	Jan 2014	June 2017
End sample	Dec 2019	May 2019	Dec 2019	May 2019

Note : $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$. Robust standard errors.

the models gives a significant result for the triple interaction term.

Although the level coefficients are not too helpful in answering our research question, we discuss them briefly. The data intensity level enters significantly positive for both importers, with a particularly large effect for the US. This means that the data intensity of products is positively associated with imports. However, we see from the interaction data intensity and EU, that this effect is weaker for the EU than for the US.

Table 7 in Appendix A performs several robustness tests on our triple difference results. Model 1 and 2 use only indirect data intensity and therefore includes the products from NACE industries 1-9. Model 3 uses both types of data intensity and looks at a smaller time frame. Model 3 and 4 use perform the triple difference analysis on shorter time periods: the 24 months around the announcement and the 24 months around the coming into force of the GDPR. Model 4 performs a fake treatment test, to see whether the common trends assumption held in 2014-2016. That these coefficients are insignificant, indicates that the common trends assumption was not violated. Table 11 in the appendix looks at the strict exogeneity assumption. Model 1 and 2 regress imports on the announcement and coming into of the GDPR. We see that imports do not seem to predict the GDPR. Model 3 and 4 perform a Granger causality test. Only the fourth lead is significant.[9] This also points towards that future GDPR applicability is not associated with current imports. Finally, we perform our analysis with clustered standard errors on the BEA commodity level. This allows the residuals of imports to correlate within BEA product groups. This is intuitive because specific shocks could affect a certain commodity. Table 9 shows that this does not affect the results.

Overall, our aggregate analysis gives robust insignificant results. Therefore, we conclude that on the aggregate level, the GDPR did not seem to have an effect on EU imports of data-

[9]The earlier leads that we tested but not displayed were insignificant.

intensive goods. We discuss possible explanations for this in Paragraph 5.2.2.

Table 3: Aggregate: triple difference

	Dependent variable: log Imports (in US $)			
	Model 1	Model 2	Model 3	Model 4
Intercept	17.650***	17.607***	19.927***	19.927***
	(0.178)	(0.132)	(0.031)	(0.022)
Post	0.001	0.154	0.050	0.108**
	(0.226)	(0.238)	(0.038)	(0.040)
Data Int.	0.096***	0.098***	0.607***	0.631***
	(0.007)	(0.005)	(0.049)	(0.035)
EU	0.269	0.173	-0.204***	-0.234***
	(0.242)	(0.182)	(0.050)	(0.035)
Post × Data Int.	0.003	-0.000	0.064	0.054
	(0.009)	(0.009)	(0.061)	(0.065)
Post × EU	-0.272	-0.242	-0.097	-0.101
	(0.315)	(0.348)	(0.062)	(0.066)
Data Int. × EU	-0.020*	-0.017*	-0.117	-0.094
	(0.009)	(0.007)	(0.069)	(0.050)
Post × Data Int. × EU	0.007	0.005	0.036	-0.008
	(0.012)	(0.013)	(0.088)	(0.096)
Num. obs.	30816	30816	30672	30672
Data intensity	Continuous	Continuous	Median	Median
Treatment month	May 2016	May 2018	May 2016	May 2018
Start sample	Jan 2014	Jan 2014	Jan 2014	June 2017
End sample	Dec 2019	Dec 2019	Dec 2019	May 2019

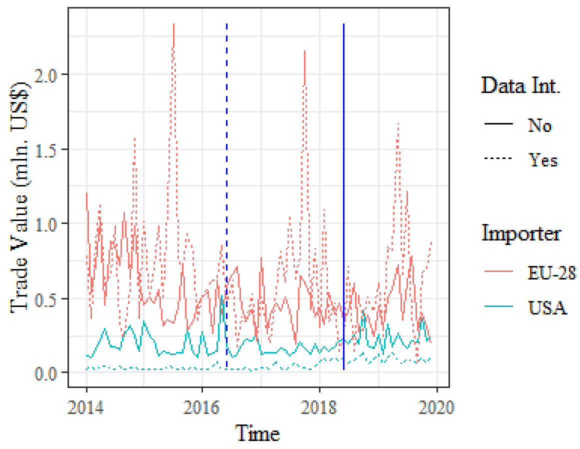
Note : $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$ Robust standard errors.

5.2 Exporter-specific

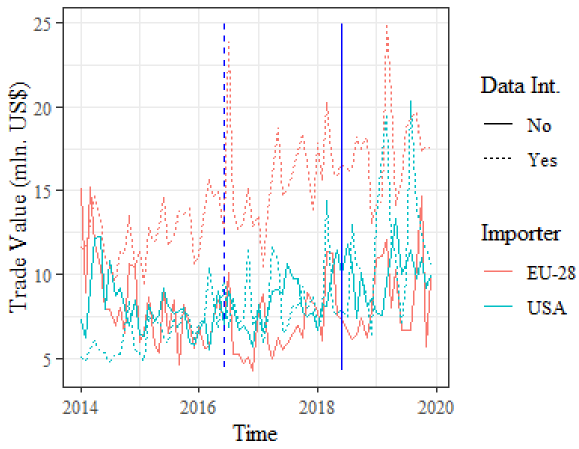
5.2.1 Difference-in-difference

We continue with an exporter-specific analysis, in which we group countries based on income or on the European Commission's adequacy decision. Figure 2 gives a graphical representation of the difference-in-difference estimations, splitted per exporter income group. Panel (a) shows that the US imported almost no data-intensive goods from low income countries. The EU imported relatively more data-intensive goods. Important to note, is that we do not control for world demand in this graph, which we do in the regression analysis. Panel (b) shows that for both the EU and the US, relative data-intensive imports seemed to increase over 2014-2019 from middle low income countries. Panel (c) displays that both importers imported large values of data-intensive products from middle high income countries. Non-data-intensive products were decreasing for the US. Panel (c) shows clearly that relative data-intensive imports from rich countries decreased over the time frame. The US experienced a data-intensive import boom in

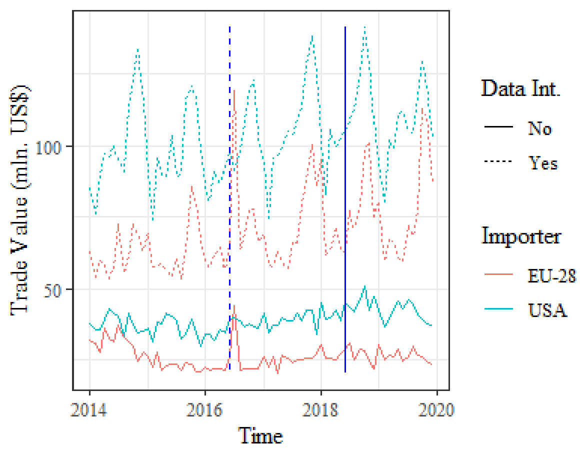
Figure 2: Exporter-specific: difference-in-difference



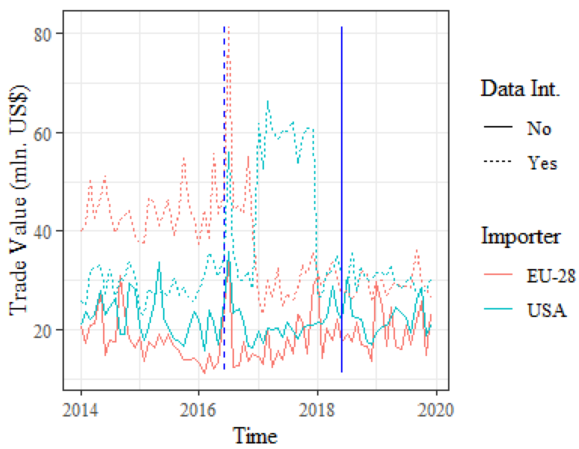
(a) Low income



(b) Middle low income



(c) Middle high income



(d) High income

Note : The development of above median data-intensive minus below median non-data-intensive products per exporter income group. The dotted vertical line is the announcement and the solid vertical line the coming into force of the GDPR.

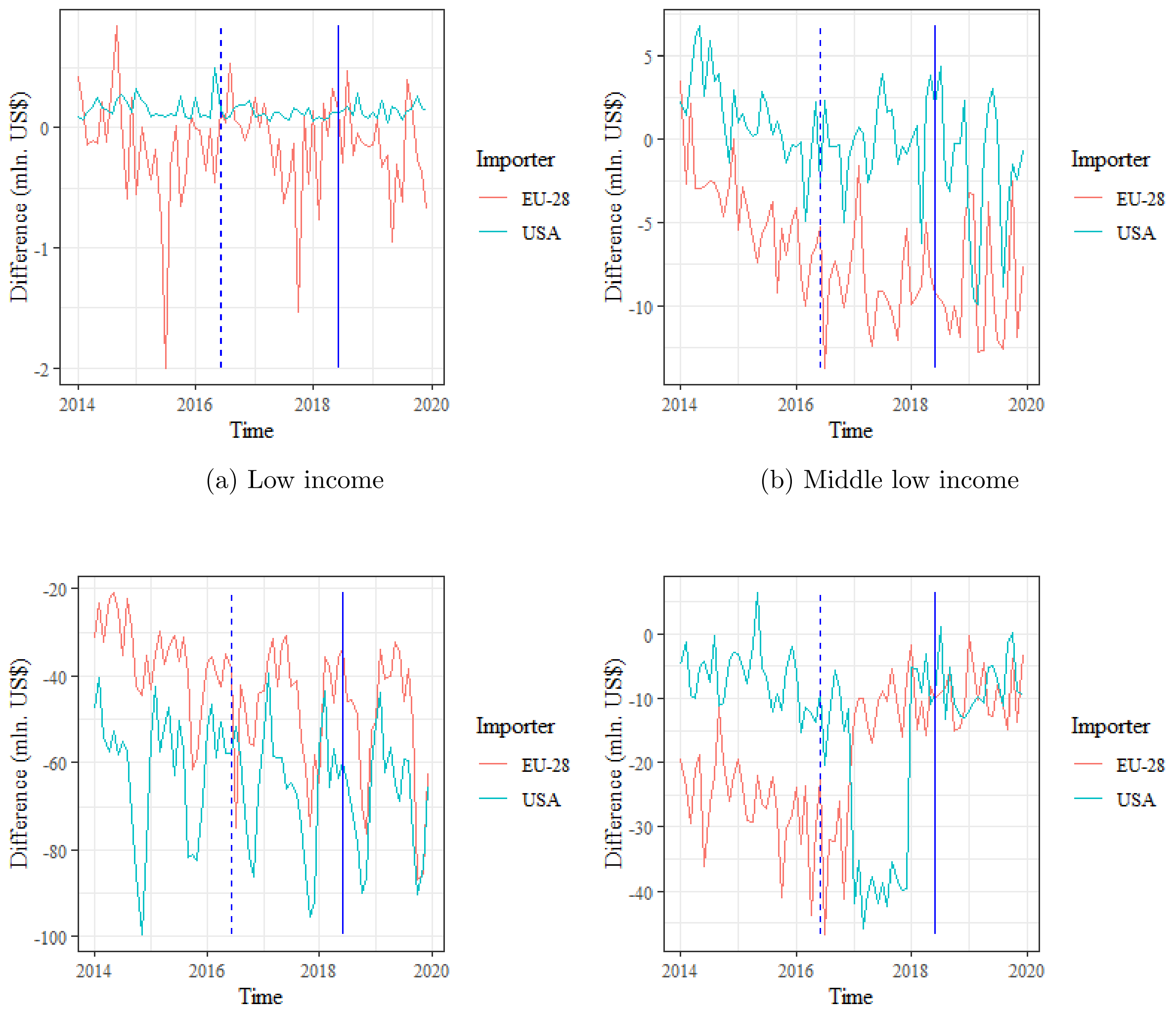
2017.[10] For the EU, we see a sharp fall in relative data-intensive imports in the beginning of 2017.

Table 4 gives the PPML difference-in-difference results, with robust standard errors and controlling for world demand. We see heterogeneous developments in the difference between data-intensive and non-data-intensive imports. Model 1 and 3 regard the 12 months before and the 12 months after the coming into force of the GDPR in May 2018. For the EU, Model 1, the difference in data-intensive and non-data-intensive imports did not change significantly in the two years around the coming into force of the GDPR. In the longer time span around the announcement month, we see that the EU experienced a relative decrease in data-intensive imports from the two richer groups. For middle high income countries, the negative difference was 0.041

[10]The boom seems to be caused by vehicle imports from Japan and China, possibly related to an expected import increase of this good (Politico, 2018).

percentage point per extra standard deviation data intensity. For high income exporters this was 0.043 percentage point. In the US, Model 3, relative data-intensive imports decreased by 0.075 percentage point for high income exporters. The other income groups did not experience a significant relative change. In the years around the GDPR announcement, Model 4, the US's data-intensive imports from both low and high income countries increased significantly. For rich exporters this development was the opposite from the EU development. Overall, we conclude that relative data-intensive imports in the EU and the US moved in opposite directions. The EU experienced a decrease in data-intensive imports from the two high income country groups. The US experienced an increase for the richest and the poorest exporter group. In Paragraph 5.2.2 we test to what extent these EU and US developments are statistically different from each other.

Figure 3: Exporter-specific: triple difference



(a) Low income

(b) Middle low income

(c) Middle high income

(d) High income

Note : The development of above median data-intensive minus below median non-data-intensive products per exporter income group. The dotted vertical line is the announcement and the solid vertical line the coming into force of the GDPR.

Table 4: Exporter-specific: difference-in-difference

	Dependent variable: log Imports (in US $)			
	Model 1	Model 2	Model 3	Model 4
Low income				
Post	1.808	-0.130	0.015	-1.014
	(1.460)	(0.827)	(0.994)	(0.613)
Data Int.	0.024	-0.021	-0.215***	-0.314***
	(0.055)	(0.023)	(0.029)	(0.027)
Post × Data Int.	-0.080	-0.008	0.015	0.050*
	(0.059)	(0.033)	(0.039)	(0.024)
Num. obs.	4922	15408	4922	15408
Middle low income				
Post	0.521	0.818	-0.703	-0.654
	(0.994)	(0.726)	(1.057)	(0.559)
Data Int.	0.065**	0.116***	-0.023	-0.037*
	(0.023)	(0.015)	(0.022)	(0.017)
Post × Data Int.	-0.020	-0.034	0.027	0.033
	(0.038)	(0.025)	(0.041)	(0.022)
Num. obs.	4922	15408	4922	15408
Middle high income				
Post	0.173	1.059*	0.431	0.575
	(0.723)	(0.499)	(0.739)	(0.477)
Data Int.	0.123***	0.169***	0.160***	0.178***
	(0.019)	(0.011)	(0.018)	(0.014)
Post × Data Int.	-0.008	-0.041*	-0.018	-0.016
	(0.027)	(0.018)	(0.027)	(0.017)
Num. obs.	4922	15408	4922	15408
High income				
Post	0.114	0.912**	1.705*	-0.925*
	(0.595)	(0.284)	(0.702)	(0.380)
Data Int.	0.032	0.087***	0.054**	-0.010
	(0.017)	(0.007)	(0.017)	(0.011)
Post × Data Int.	-0.007	-0.043***	-0.075**	0.039**
	(0.022)	(0.010)	(0.028)	(0.015)
Num. obs.	4922	15408	4922	15408
Importer	EU	EU	US	US
Treatment month	May 2018	May 2016	May 2018	May 2016
Start sample	June 2017	Jan 2014	June 2017	Jan 2014
End sample	May 2019	Dec 2019	May 2019	Dec 2019

Note : $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Robust standard errors. World demand control not displayed.

Table 5: Exporter-specific: triple difference

	Dependent variable: log Imports (in US $)			
	Model 1	Model 2	Model 3	Model 4
Low income				
Post $\times$ Data Int.	-0.007	0.046	0.030	-0.012
	(0.038)	(0.024)	(0.024)	(0.039)
Post $\times$ EU	1.354	0.752	0.324	-2.217
	(1.783)	(1.022)	(1.000)	(1.554)
Data Int. $\times$ EU	0.292***	0.315***	0.293***	0.266***
	(0.054)	(0.031)	(0.025)	(0.038)
Post $\times$ Data Int. $\times$ EU	-0.074	-0.053	-0.038	0.092
	(0.072)	(0.041)	(0.040)	(0.062)
Num. obs.	9844	30816	30816	9844
Middle low income				
Post $\times$ Data Int.	0.002	0.027	0.022	-0.013
	(0.050)	(0.026)	(0.032)	(0.044)
Post $\times$ EU	0.662	0.518	1.209	-0.623
	(1.453)	(0.833)	(0.948)	(1.410)
Data Int. $\times$ EU	0.238***	0.254***	0.249***	0.255***
	(0.035)	(0.026)	(0.019)	(0.041)
Post $\times$ Data Int. $\times$ EU	-0.019	-0.029	-0.047	0.027
	(0.058)	(0.032)	(0.037)	(0.055)
Num. obs.	9844	30816	30816	9844
Middle high income				
Post $\times$ Data Int.	-0.022	-0.018	-0.017	-0.019
	(0.029)	(0.018)	(0.020)	(0.033)
Post $\times$ EU	-0.452	-0.494	-0.317	-1.164
	(1.069)	(0.701)	(0.723)	(1.297)
Data Int. $\times$ EU	0.060*	0.046*	0.052***	0.027
	(0.026)	(0.021)	(0.014)	(0.033)
Post $\times$ Data Int. $\times$ EU	0.020	0.015	0.013	0.041
	(0.039)	(0.026)	(0.027)	(0.047)
Num. obs.	9844	30816	30816	9844
High income				
Post $\times$ Data Int.	-0.078**	0.038*	-0.037*	-0.004
	(0.030)	(0.016)	(0.016)	(0.024)
Post $\times$ EU	-1.664	1.514**	-0.238	-0.280
	(0.894)	(0.479)	(0.515)	(0.756)
Data Int. $\times$ EU	0.015	0.123***	0.079***	0.125***
	(0.022)	(0.012)	(0.010)	(0.018)
Post $\times$ Data Int. $\times$ EU	0.072*	-0.069***	0.008	0.013
	(0.034)	(0.018)	(0.019)	(0.028)
Num. obs.	9844	30816	30816	9844
Treatment month	May 2018	May 2016	May 2018	May 2015
Start sample	June 2017	Jan 2014	Jan 2014	June 2014
End sample	May 2019	Dec 2019	Dec 2019	May 2016

Note : $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Robust standard errors. Levels and world demand not displayed.

5.2.2 Triple difference

Figure 3 shows the development of the difference between data and non-data-intensive imports by the EU and the US. Panel (a) displays that from low income exporters the US always imported relatively more non-data-intensive goods. For the EU, this fluctuated around zero, with a shock mid 2015 and end 2017. Panel (b) shows a different picture for middle low income countries. For both the EU and the US, the difference became more negative over the time frame, meaning that data-intensive imports increased compared to non-data-intensive imports. In Panel (c), we see that for both countries the difference was decreasing further below zero. This exporter group experiences shocks at the end of the year in data-intensive products. This suggests that a certain data-intensive product from middle income exporters has strong seasonality in the US. Panel (d) shows that the EU exported relatively more data-intensive goods compared to the US from rich countries. After the shocks in 2017, the difference is just below zero. This suggests a strong negative relative supply shift in data-intensive products from rich countries, mainly from 2018.

Table 5 shows the PPML triple difference estimations in data-intensive imports by the EU compared to the US. All models have robust standard errors and control for world demand. Model 1 looks at the 12 months before and after the coming into force of the GDPR in May 2018. Only for high income countries, the EU's increase of data-intensive imports was stronger than the development in the US. This difference was 0.072 percentage point per extra standard deviation data intensity. Model 2 regards the full time frame and uses the announcement date of the GDPR as treatment month. Here we see no significant differences, except for in the high income group. However, the sign is not the same as the sign in Model 1. This is not what we would expect. We would expect that the announcement would have a similar effect as the coming into force, because firms are anticipating on future compliance issues. Another possibility is that more firms want to sell to EU consumers between May 2016 and May 2018, in order to gain on in this market before this becomes more costly. However, these results show the opposite. As we saw in 2 Panel (d), this result is more likely to be driven by the data-intensive import boom in the US. Model 3 regards the full time frame and uses the coming into force of the GDPR as treatment month. None of the coefficients are significant in this model.

Model 4 in 5 performs a fake treatment test around May 2015. This test is insignificant for all groups, meaning that the common trends assumption was not violated in those years. Table 11 in the appendix shows that for both the Granger causality test and the test with GDPR applicability as dependent variable the coefficients are insignificant. This means that the strict exogeneity assumption is not violated. Table 12 shows that our results hold if we cluster the standard errors on the BEA product level. The results are also robust to using different types of data intensity dummies.[11]

Table 10 in Appendix A shows the PPML triple difference results for the countries that received an adequacy decision of the European Commission. Model 1 covers the 24 months before and after the coming into force of the GDPR. This model shows a positive difference of 0.126 percentage point. After the announcement date, Model 2, we estimate a significant negative difference of 0.124 percentage point. Model 1 could be weak evidence of a positive

[11] Percentile 25, 50 and 75.

effect of having an adequacy decision on data-intensive exports to the EU. However, the results of those two models are similar to those for the rich country group. This could be related to the fact that Japan and Canada drive the boom. Those two countries also have an adequacy decision. In Figure 4 in the appendix, we see the same boom as for the rich country group. This means that this result is likely to be more related to the US boom in data-intensive goods than the GDPR. Using the full time frame and the coming into force as the treatment, Model 3, we find no significant difference. Our fake treatment test in 2015, Model 4, shows that the common trends assumption was probably not violated.

We conclude that we have robust insignificant results on the impact of the GDPR on EU imports of data-intensive goods. The significant differences we measure for the rich and adequacy country groups are most likely unrelated to this privacy regulation. An explanation for this could be that possible trade deterring effects are mitigated by the increase in trust or the unification of the regulations in European countries. Another possibility is that firms are simply willing to invest in the privacy measures in order to be able to be active in the EU now and in the future. Further research could study those incentives by means of a micro-analysis.

6 Conclusion

In one of the first ex-post empirical studies on the effect of the EU privacy law on trade flows, we found no effect of the GDPR on EU imports of data-intensive goods. Both aggregate imports and imports from different types of exporters did not respond to the announcement or the coming into force of the privacy regulation. To account for the various other factors that could have affected imports, we performed a difference-in-difference and a triple difference analysis. In the difference-in-difference approach, we compared the development of data-intensive imports to the development of non-data-intensive imports. In the triple difference approach, we compared the difference between the two product groups in terms of EU imports with this difference in terms of US imports. Our results could only be influenced by the limited set of factors that is time-varying, affects the difference between data-intensive and non-data-intensive imports and is EU or US specific. By not finding any effect, we contradict the many academics that argue that the GDPR serves as a non-tariff import barrier (Mishra, 2015; Ferracane and Van der Marel, 2018; Pasadilla et al., 2020; Bauer, Ferracane, Lee-Makiyama, et al., 2016).

Although our approach eliminated many sources of endogeneity, some possible pitfalls remain. First, there could still be factors that are not absorbed by the triple difference approach. The most important example is the data-intensive import boom in the US in 2017. In our data we observe that this boom is mainly driven by Canadian and Japanese vehicles. We suspect that this could be related to an expected tenfold tariff increase for US car imports (Politico, 2018). This import boom causes us to observe a positive effect after May 2016 and negative effect after May 2018 for rich exporters and exporters with an adequacy decision. This finding is probably unrelated to the GDPR. More generally, tariffs and other importer specific policies that affected data-intensive products differently could have biased our results. In future research, those political factors could be added to the analysis. Another important possible omitted variable is the trade war between the US and China. This impact could also be systematically different for the data-intensive product group (Financial Times, 2019). If US data-intensive imports became less due to the trade war and we suppose that the GDPR affects data-intensive imports negatively, our triple difference estimator would suffer a bias towards zero. A third example is the fluctuation in oil prices. The oil price was increasing in 2017 and dropped end 2018 (Oilprice, 2020). Oil is non-data-intensive, so a positive oil shock implies a fall in relative data-intensive imports and vice versa. Supposing that the GDPR had a negative effect on data-intensive imports, the oil price fluctuations caused an underestimation after May 2016 and an overestimation after May 2018. This would be mainly the case in the middle high income group, as many large oil exporters are part of this group. In future research, one could make a selection of products to include in the analysis. This should be the result of a careful assessment of the political factors that have been present in the years on the sample. Another pitfall is that supply effects are not fully accounted for by comparing the EU with the US. It is commonly known in international trade that factors such as distance are highly explanatory for trade values. This would mean that the EU is subject to different supply effects than the US. Future studies could control for supply effects by collecting data on world supply per exporter group.

Another possible source of endogeneity is measurement error. Our direct data intensity measure is not available for hunting, mining and fishing industries. This means that we drop some of the likely to be non-data-intensive products. This would only be a problem if the shocks and trade values belonging to these products were substantial. Another possible measurement error, is that our direct data intensity classification is not very precise. The classification covers relatively large industry or product groups. This means that there is no variation within these groups identified. Ideally, one would have a more detailed measure of data intensity. We expect this measurement error in the independent variable to be random, causing a bias towards zero. A measurement error that is more of a concern, is that our data intensity is time-fixed. It could be that there was an increase in data intensity of specific products. In that case, real data-intensive imports after the treatment would be higher than measured. If we suppose the GDPR had a negative effect, this measurement error would cause an overestimation of our triple difference coefficient. This could be solved by including time-varying data intensity, if one is able to deal with reverse causality and data availability problems. In that sense an instrumental variable approach might be a more appropriate solution.

A remaining issue is reverse causality. For the assignment of data intensity of products, this is mitigated by using an input-output table from 2012 instead of a more recent year. Although there could still be some reverse causality bias via autocorrelation in the input-output shares, we do not expect a large confounding effect. There could be a more direct causality from GDPR applicability to import values. The urge of the European Commission to impose data localization measures could be triggered by large import values of data-intensive goods. However, our tests for the strict exogeneity assumption do not find predictive power in trade flows towards the applicability of the GDPR.

Many more interesting empirical questions remain in the field of privacy regulation and the GDPR in specific. One of the policy objectives of the GDPR was to enhance the EU internal market (ITGP, 2017). It would be interesting to look at intra-EU trade in data-intensive goods. Another relevant topic is to study services trade. Although we did account for services in the indirect data intensity measure, we did not test direct effects on services trade. As direct data intensity is highest in services industries, this would be an interesting topic for future studies. Finally, others could use the approach we used to assess privacy regulations from other countries or to study effects in the more long run. For now, we suggest that the GDPR did not negatively affect EU imports of data-intensive goods, which is positive news for all policy makers that want to protect their citizens' privacy.

References

Aaronson, S., & Leblond, P. (2018). Another digital divide: The rise of data realms and its implications for the wto. *Journal of International Economic Law*, *21*(2), 245–272.

Acquisti, A., Taylor, C., & Wagman, L. (2016). The economics of privacy. *Journal of economic Literature*, *54*(2), 442–92.

Arnold, J. M., Javorcik, B., Lipscomb, M., & Mattoo, A. (2012). Services reform and manufacturing performance: Evidence from india. *World Bank Policy Research Papers Series no. 5948*.

Bauer, M., Erixon, F., Krol, M., & Lee-Makiyama, H. (2013). The economic importance of getting data protection right: Protecting privacy, transmitting data, moving commerce.

Bauer, M., Ferracane, M., Lee-Makiyama, H., & Van der Marel, E. (2016). Unleashing internal data flows in the EU: An economic assessment of data localisation measures in the eu member states.

Bauer, M., Ferracane, M., & Van der Marel, E. (2016). Tracing the economic impact of regulations on the free flow of data and data localization.

Bauer, M., Lee-Makiyama, H., Van der Marel, E., & Verschelde, B. (2014). The costs of data localisation: Friendly fire on economic recovery.

BEA. (2012). Input-output accounts data: Commodity-by-industry. U.S. Bureau of Economic Analysis (BEA). https://www.bea.gov/industry/input-output-accounts-data

Bratt, M. (2017). Estimating the bilateral impact of nontariff measures on trade. *Review of International Economics*, *25*(5), 1105–1129.

Christensen, L., Colciago, A., Etro, F., & Rafert, G. (2013). The impact of the data protection regulation in the eu. *Intertic Policy Paper, Intertic.*

Ciriani, S. (2015). The economic impact of the european reform of data protection. *Communications & Strategies*, (97), 41–58.

Crivelli, P., & Groeschl, J. (2015). The impact of sanitary and phytosanitary measures on market entry and trade flows. *The World Economy*, *39*(3), 444–473.

DPD. (1995). European Commission. Retrieved May 3, 2020, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%5C%3A31995L0046

European Commission. (2020). EU adequacy decisions. https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en

Eurostat. (2019a). E-commerce sales. European Commission. https://ec.europa.eu/eurostat/web/products-datasets/product?code=isoc_cisce_ra

Eurostat. (2019b). Security policy: Measures, risks and staff awareness. European Commission. https://ec.europa.eu/eurostat/web/products-datasets/product?code=isoc_cisce_ra

Ferracane, M., & Van der Marel, E. (2018). Do data flows restrictions inhibit trade in services?

Financial Times. (2019). How the trade war is damaging the us tech industry. https://www.ft.com/content/16fa93ba-bf69-11e9-b350-db00d509634e

Fugazza, M. (2013). The economics behind non-tariff measures: Theoretical insights and empirical evidence. *UN*, *57*.

GDPR. (2018). European Commission. Retrieved May 3, 2020, from https://eur-lex.europa.eu/legal-content/NL/TXT/?uri=CELEX%5C%3A32016R0679

Goldberg, S., Johnson, G., & Shriver, S. (2019). Regulating privacy online: The early impact of the GDPR on European web traffic & e-commerce outcomes. *SSRN Working Paper*.

Goldfarb, A., & Trefler, D. (2018). AI and international trade.

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

Hirshleifer, J. (1971). *Foundations of the theory of speculation: Information, risk, and markets*. Western Management Science Institute, University of California.

ITGP, P. T. (2017). *EU general data protection regulation (GDPR): An implementation and compliance guide*. IT Governance Ltd.

Jia, J., Jin, G., & Wagman, L. (2018). The short-run effects of GDPR on technology venture investment. *SSRN Working Paper*.

Jia, J., Jin, G., & Wagman, L. (2019). GDPR and the localness of venture investment.

Korinek, J., Melatos, M., & Rau, M. (2008). A review of methods for quantifying the trade effects of standards in the agri-food sector. *OECD Trade Policy Papers*.

Koski, H., & Valmari, N. (2020). Short-term impacts of the GDPR on firm performance.

Lee, K., Kang, J., & Kang, M. (2019). Welfare effects of the EU GDPR and data localization measures. *KIEP Research Paper, World Economy Brief*, 19–07.

Li, Y., & Beghin, J. C. (2017). A meta-analysis of estimates of the impact of technical barriers to trade. *Nontariff Measures and International Trade*, 63–77.

Meltzer, J. (2019). A WTO reform agenda: Data flows and international regulatory cooperation, *180*.

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

Mishra, N. (2015). Data localization laws in a digital world: Data protection or data protectionism? *The Public Sphere (2016)*.

Oilprice. (2020). Oil price charts. https://oilprice.com/oil-price-charts

Pasadilla, G. O., Duval, Y., & Anukoonwattaka, W. (2020). Next generation non-tariff measures: Emerging data policies and barriers to digital trade.

Politico. (2018). Trump considering 25 percent tax on car imports, citing national security. https://www.politico.eu/article/donald-trump-considering-25-percent-tax-on-car-imports-citing-national-security/

Posner, R. (1981). The economics of privacy. *The American economic review*, *71*(2), 405–409.

Privacy Affairs. (2020). GDPR fines list: Find all GDPR fines and detailed statistics. *PrivacyAffairs GDPR Fines Tracker Statistics Comments*. https://www.privacyaffairs.com/gdpr-fines/

Ronen, E. (2017). Quantifying the trade effects of ntms: A review of the empirical literature. *Journal of Economics and Political Economy*, *4*(3), 263–274.

Seo, J., Kim, K., Park, M., Park, M., & Lee, K. (2018). An analysis of economic impact on iot industry under GDPR. *Mobile Information Systems*, *2018*.

Silva, J. M. C. S., & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, *88*(4), 641–658.

Sobolewski, M., & Paliński, M. (2017). How much consumers value online privacy? welfare assessment of new data protection regulation (GDPR).

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics.*

Stigler, G. (1980). An introduction to privacy in economics and politics. *The Journal of Legal Studies*, *9*(4), 623–644.

Stone, S., M., J., & F., D. (2015). Emerging policy issues: Localisation barriers to trade, *180.*

Swire, P., & Litan, R. (1998). *None of your business: World data flows, electronic commerce, and the European privacy directive.* Jessica Kingsley Publishers.

TBT Agreement. (1995). WTO. Retrieved May 3, 2020, from https://www.wto.org/english/docs_e/legal_e/17-tbt_e.htm

UNCTAD. (2013). Non-tariff measures to trade: Economic and policy issues for developing countries. *Developing Countries in International Trade Studies.*

United Nations. (2020). Imports at the 6 digit hs product level. ComTrade.

Van der Marel, E., Bauer, M., Lee-Makiyama, H., & Verschelde, B. (2016). A methodology to estimate the costs of data regulations. *International Economics*, *146*, 12–39.

Wigand, C., & Voin, M. (2018). Questions and answers – general data protection regulation. *EC Europa.* https://ec.europa.eu/commission/presscorner/detail/en/MEMO_18_387

Wing, C., Simon, K., & Bello-Gomez, R. A. (2018). Designing difference in difference studies: Best practices for public health policy research. *Annual review of public health*, *39.*

Wooldridge. (2007). Difference-in-differences estimation. https://www.nber.org/WNE/lect_10_diffindiffs.pdf

WorldBank. (2020). World bank country and lending groups. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

WTO. (2020). Technical barriers to trade information management system.

Yuan, B., & Li, J. (2019). The policy effect of the general data protection regulation (GDPR) on the digital public health sector in the European Union: An empirical investigation. *International journal of environmental research and public health*, *16*(6), 1070.

A Appendix

Table 6: Direct data intensity indicators

NACE	Label	Sells online	Documents
10-12	Mf. of beverages, food and tobacco products	23	25
13-15	Mf. of textiles, wearing apparel, leather and related products	15	24
16-18	Mf. of wood & products of wood & cork, except furniture	17	34
19-23	Mf. of coke, refined petroleum, chemical & basic pharmaceutical products, rubber & plastics, other non-metallic mineral products	19	40
24-25	Mf. of basic metals & fabricated metal products excluding machines & equipments	12	30
26	Mf. of computer, electronic and optical products	19	52
27-28	Mf. of electrical equipment, machinery and equipment n.e.c.	12	46
29-30	Mf. of motor vehicles, trailers and semi-trailers, other transport equipment	28	44
31-33	Mf. of furniture and other manufacturing	14	31
35-39	Electricity, gas, steam, air conditioning and water supply	11	47
41-43	Construction	5	23
45	Trade of motor vehicles and motorcycles	22	37
46	Wholesale trade, except of motor vehicles and motorcycles	31	39
47	Retail trade, except of motor vehicles and motorcycles	27	30
49-53	Transportation and storage	15	29
55	Accommodation	63	37
58-60	Publishing activities	35	51
61	Telecommunications	25	66
62-63	Computer programming, consultancy and related activities, information service activities	19	72
68	Real estate activities	10	44
69-74	Professional, scientific and technical activities	8	53
77-82	Administrative and support service activities	13	39
95	Repair of computers and communication equipment	28	55

Note: Percentage of companies in upstream industries that sells online (more than 1 percent of turnover) or has official documents on ICT security. Mf. is manufacturing.

Table 7: Aggregate: robustness tests

	Dependent variable: log Imports (in US $)			
	Model 1	Model 2	Model 3	Model 4
Intercept	20.429***	20.244***	17.564***	17.766***
	(0.278)	(0.198)	(0.278)	(0.278)
Post	-0.418	-0.229	0.223	-0.400
	(0.342)	(0.355)	(0.374)	(0.389)
Data Int.	-0.001	0.006	0.101***	0.092***
	(0.011)	(0.008)	(0.011)	(0.011)
EU	1.442***	1.495***	0.019	0.242
	(0.379)	(0.274)	(0.396)	(0.380)
Post × Data Int.	0.020	0.014	-0.004	0.013
	(0.014)	(0.014)	(0.014)	(0.015)
Post × EU	0.159	0.152	-0.068	0.027
	(0.480)	(0.528)	(0.550)	(0.528)
Data Int. × EU	-0.068***	-0.070***	-0.014	-0.019
	(0.015)	(0.011)	(0.015)	(0.014)
Post × Data Int. × EU	-0.008	-0.009	0.000	-0.002
	(0.019)	(0.021)	(0.021)	(0.020)
Num. obs.	33696	33696	10272	10272
Data intensity	Indirect	Indirect	Both	Both
Treatment month	May 2016	May 2018	May 2018	May 2015
Start sample	Jan 2014	Jan 2014	June 2017	June 2014
End sample	Dec 2019	Dec 2019	May 2019	May 2016

Note : $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Robust standard errors.

Table 8: Aggregate: strict exogeneity

	Dependent variable:			
	Post × Data Int. × EU		log Imports (in US $)	
	Model 1	Model 2	Model 3	Model 4
Intercept	1.450***	2.173***	20.218***	20.191***
	(0.019)	(0.012)	(0.018)	(0.016)
Imports lag 1	0.000	0.000		
	(0.000)	(0.000)		
Imports lag 2	-0.000	-0.000		
	(0.000)	(0.000)		
Imports lag 3	-0.000	-0.000		
	(0.000)	(0.000)		
Imports lag 4	0.000	0.000		
	(0.000)	(0.000)		
GDPR lead 1			-0.001	0.004
			(0.005)	(0.006)
GDPR lead 2			0.001	-0.001
			(0.007)	(0.007)
GDPR lead 3			0.004	0.004
			(0.007)	(0.007)
GDPR lead 4			-0.009*	-0.011*
			(0.005)	(0.005)
Treatment month	May 2018	May 2016	May 2018	May 2016
Num. obs.	20544	20544	20544	20544
Start sample	Jan 2014	Jan 2014	Jan 2014	Jan 2014
End sample	Dec 2019	Dec 2019	Dec 2019	Dec 2019

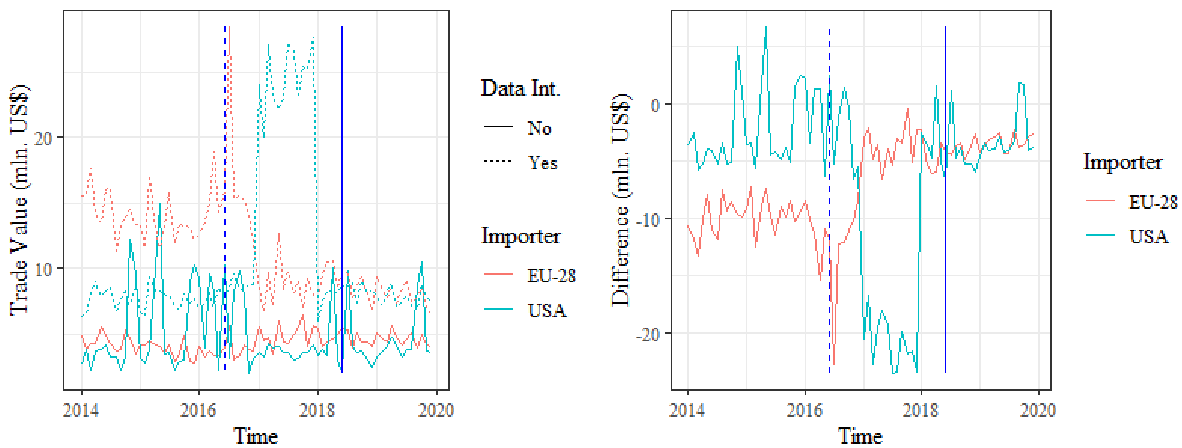
Note : $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Robust standard errors. GDPR is Post × Data Int. × EU.

Table 9: Aggregate: triple difference clustered

	Dependent variable: log Imports (in US $)			
	Model 1	Model 2	Model 3	Model 4
Intercept	17.650***	17.607***	19.927***	19.927***
	(0.935)	(0.941)	(0.160)	(0.150)
Post	0.001	0.154	0.050	0.108***
	(0.130)	(0.153)	(0.037)	(0.027)
Data Int.	0.096**	0.098**	0.607*	0.631*
	(0.036)	(0.036)	(0.256)	(0.249)
EU	0.269	0.173	-0.204**	-0.234***
	(0.472)	(0.461)	(0.071)	(0.070)
Post × Data Int.	0.003	-0.000	0.064	0.054
	(0.004)	(0.006)	(0.046)	(0.050)
Post × EU	-0.272	-0.242	-0.097***	-0.101**
	(0.188)	(0.227)	(0.025)	(0.032)
Data Int. × EU	-0.020	-0.017	-0.117	-0.094
	(0.019)	(0.018)	(0.146)	(0.145)
Post × Data Int. × EU	0.007	0.005	0.036	-0.008
	(0.007)	(0.009)	(0.058)	(0.072)
Num. obs.	30816	30816	30672	30672
Data intensity	Continuous	Continuous	Median	Median
Treatment month	May 2016	May 2018	May 2016	May 2018
Start sample	Jan 2014	Jan 2014	Jan 2014	June 2017
End sample	Dec 2019	Dec 2019	Dec 2019	May 2019

Note : $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$. BEA commodity clustered standard errors.

Figure 4: Adequacy exporters



(a) Difference-in-difference (b) Triple difference

Note : Panel (a) is the development of above and below median data-intensive and non-data intensive products. Panel (b) is the development of the difference between above and below median data-intensive and non-data intensive products. The dotted vertical line is the announcement and the solid vertical line the coming into force of the GDPR.

Table 10: Adequacy: triple difference

	Dependent variable: log Imports (in US $)			
	Model 1	Model 2	Model 3	Model 4
Intercept	12.672***	16.225***	14.477***	16.238***
	(0.526)	(0.536)	(0.413)	(0.961)
Post	2.846**	-2.343***	1.042	0.457
	(0.948)	(0.657)	(0.635)	(1.139)
Data Int.	0.111***	-0.031	0.042**	-0.028
	(0.020)	(0.019)	(0.015)	(0.034)
EU	1.985**	-2.438***	-0.503	-2.332*
	(0.631)	(0.546)	(0.434)	(0.971)
World Demand	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Post × Data Int.	-0.124***	0.094***	-0.051*	-0.019
	(0.038)	(0.024)	(0.024)	(0.039)
Post × EU	-2.935**	2.865***	-0.521	-0.825
	(1.034)	(0.696)	(0.676)	(1.185)
Data Int. × EU	-0.079**	0.108***	0.026	0.101**
	(0.025)	(0.019)	(0.016)	(0.034)
Post × Data Int. × EU	0.126**	-0.124***	0.020	0.031
	(0.041)	(0.026)	(0.025)	(0.041)
Num. obs.	9844	30816	30816	9844
Data intensity	Continuous	Continuous	Median	Median
Treatment month	May 2018	May 2016	May 2018	May 2015
Start sample	June 2017	Jan 2014	Jan 2014	June 2014
End sample	May 2019	Dec 2019	Dec 2019	May 2016

Note : $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Robust standard errors.

Table 11: Exporter-specific: strict exogeneity

	Dependent variable:			
	Post × Data Int. × EU		log Imports (in US $)	
	Model 1	Model 2	Model 3	Model 4
Low income				
Imports lag 2	0.000	0.000		
	(0.000)	(0.000)		
Imports lag 4	0.000	0.000		
	(0.000)	(0.000)		
GDPR lead 2			0.009	0.002
			(0.022)	(0.018)
GDPR lead 4			0.024	0.035
			(0.017)	(0.024)
Num. obs.	20544	20544	20544	20544
Middle low income				
Imports lag 2	0.000	0.000		
	(0.000)	(0.000)		
Imports lag 4	0.000	0.000		
	(0.000)	(0.000)		
GDPR lead 2			-0.000	-0.006
			(0.025)	(0.026)
GDPR lead 4			0.012	0.013
			(0.016)	(0.018)
Num. obs.	20544	20544	20544	20544
Middle high income				
Imports lag 2	0.000	-0.000		
	(0.000)	(0.000)		
Imports lag 4	-0.000	0.000		
	(0.000)	(0.000)		
GDPR lead 2			0.001	0.004
			(0.018)	(0.019)
GDPR lead 4			-0.009	-0.005
			(0.013)	(0.013)
Num. obs.	20544	20544	20544	20544
High income				
Imports lag 2	-0.000	0.000		
	(0.000)	(0.000)		
Imports lag 4	0.000	-0.000		
	(0.000)	(0.000)		
GDPR lead 2			0.010	0.001
			(0.016)	(0.009)
GDPR lead 4			0.003	-0.007
			(0.012)	(0.006)
Num. obs.	20544	20544	20544	20544
Start sample	Jan 2014	Jan 2014	Jan 2014	Jan 2014
End sample	Dec 201	Dec 2019	Dec 2019	Dec 201

Note : $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Robust standard errors. Other lags give similar results.
GDPR is Post × Data Int. × EU.

Table 12: Exporter-specific: triple difference clustered

	Dependent variable: log Imports (in US $)			
	Model 1	Model 2	Model 3	Model 4
Low income				
Post × Data Int.	-0.007	0.046	0.030	-0.012
	(0.046)	(0.038)	(0.037)	(0.028)
Post × EU	1.354	0.752	0.324	-2.217
	(1.977)	(1.519)	(1.123)	(1.661)
Data Int. × EU	0.292**	0.315***	0.293***	0.266***
	(0.110)	(0.084)	(0.084)	(0.075)
Post × Data Int. × EU	-0.074	-0.053	-0.038	0.092
	(0.081)	(0.056)	(0.046)	(0.066)
Num. obs.	9844	30816	30816	9844
Middle low income				
Post × Data Int.	0.002	0.027	0.022	-0.013
	(0.049)	(0.026)	(0.034)	(0.020)
Post × EU	0.662	0.518	1.209	-0.623
	(1.562)	(1.164)	(1.435)	(0.883)
Data Int. × EU	0.238***	0.254***	0.249***	0.255**
	(0.046)	(0.069)	(0.059)	(0.083)
Post × Data Int. × EU	-0.019	-0.029	-0.047	0.027
	(0.061)	(0.044)	(0.054)	(0.032)
Num. obs.	9844	30816	30816	9844
Middle high income				
Post × Data Int.	-0.022	-0.018	-0.017	-0.019
	(0.029)	(0.017)	(0.013)	(0.010)
Post × EU	-0.452	-0.494	-0.317	-1.164**
	(0.846)	(0.408)	(0.603)	(0.383)
Data Int. × EU	0.060*	0.046	0.052	0.027
	(0.028)	(0.030)	(0.028)	(0.031)
Post × Data Int. × EU	0.020	0.015	0.013	0.041**
	(0.032)	(0.013)	(0.021)	(0.014)
Num. obs.	9844	30816	30816	9844
High income				
Post × Data Int.	-0.078	0.038	-0.037	-0.004
	(0.067)	(0.045)	(0.031)	(0.012)
Post × EU	-1.664	1.514	-0.238	-0.280
	(1.613)	(0.890)	(0.864)	(0.461)
Data Int. × EU	0.015	0.123*	0.079*	0.125**
	(0.044)	(0.048)	(0.040)	(0.044)
Post × Data Int. × EU	0.072	-0.069	0.008	0.013
	(0.069)	(0.036)	(0.036)	(0.018)
Num. obs.	9844	30816	30816	9844
Treatment month	May 2018	May 2016	May 2018	May 2015
Start sample	June 2017	Jan 2014	Jan 2014	June 2014
End sample	May 2019	Dec 2019	Dec 2019	May 2016

Note : $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. BEA commodity clustered standard errors. Levels and world demand not displayed.