THE VALUE OF CROSS-BORDER DATA FLOWS TO EUROPE: RISKS AND OPPORTUNITIES

Report prepared for DIGITALEUROPE

June 2021

CONTENTS

FOREWORD

International data flows are vital to Europe's economic success. Setting the right framework for data flows now can have a huge positive impact on our economy by 2030.

The European Commission recently launched its targets for the Digital Decade. Among them were that by 2030:

- 75 per cent of European enterprises should have taken up cloud computing services, big data and artificial intelligence.
- Europe should double its number of unicorns.
- More than 90 per cent of European SMEs should reach at least a basic level of digital intensity.

Underpinning each of these goals is the need to be able to transfer data smoothly and securely across borders. This study demonstrates that even in a limited assessment – by looking only at international trade – cross-border data transfers can have a huge positive impact on the economy by 2030 if we make the right decisions now.

The growth of the digital economy and the success of European companies is dependent on the ability to transfer data. This is especially so when we note that already in 2024, 85 per cent of the world's GDP growth is expected to come from outside the EU.[1] Our study shows that we could be missing out on around €2 trillion worth of growth by the end of the Digital Decade. This is roughly the same size as the Italian economy any given year.

Restrictions on cross-border data flows affect companies of all sizes and sectors. The EU manufacturing sector stands to lose the most in absolute value – indeed, more than half of our total losses from data restrictions. As SMEs account for almost a quarter of all goods exported from the EU, they will be heavily impacted. Sectors such as media and culture are some of the most impacted in relative terms, losing about 10 per cent of their exports. These are some of the industries that have made Europe what it is today.

A great example is that of our member Airbus. The latest A350 has 50,000 sensors on board collecting 2.5 terabytes of data every day. It shows just how much data has penetrated into traditional industries.

Data transfers are not only a key aspect of international trade, the value of which we have attempted to capture in this research, but are crucial for economic activity more at large. For example, moving HR information from a subsidiary to a parent company, transferring health data for ground-breaking research, or simply being able to use the perfect application for the tasks you need to do. Hampering the data flows behind these business decisions has a negative impact on all companies' economic prospects.

Europe stands at a crossroads. It can either set the right framework for data transfers, and win the Digital Decade, or it can follow its current trend and move

[1] https://data.consilium.europa.eu/doc/document/ST-6308-2021-INIT/en/pdf

towards data protectionism, and lose. Our analysis shows that the consequences of these decisions will have a huge impact on exports, jobs and growth, and will ultimately define whether Europe can reach its ambitious industrial and digital goals.

Hilary Mine
President
DIGITALEUROPE

Cecilia Bonefeld-Dahl
Director General
DIGITALEUROPE

EXECUTIVE SUMMARY

Frontier Economics was commissioned by DIGITALEUROPE to estimate the economic impact of cross-border data flows on the economy of the European Union (EU), and the potential impact of changes to existing restrictions on these flows. This report provides new, robust estimates of the economic cost to the EU of additional restrictions, and the potential economic gain that could result from liberalisation. Our estimates are relatively conservative. We focus on modelling the role of cross-border data flows in facilitating international trade (imports and exports), as this allows us to use an established robust methodology to quantify the economic impact of these flows. However, we do not quantify other benefits of cross-border data flows. In particular, our results do not take into account the role of cross-border data flows in supporting the delivery of public services, and they largely exclude the economic impact of cross-border activities within companies, which typically are not recorded in international trade data. The key messages from our analysis are summarised below.

KEY FINDINGS

- Cross-border data flows are central to the EU economy, and especially its trade and investment. Many sectors, including **high-value manufacturing**, **IT and information services, media, and cultural sectors, are highly reliant on data**. This suggests that data localisation internationally and in the EU carries serious risks for sectors that are part of both the EU's industrial strategy and its wider social and economic agenda.
- The future direction of global policy on data flows therefore matters. **The difference between a path that is moderately liberalising and one that is moderately restrictive is economically significant: worth a little over 1.5% in EU GDP per year**. This is equivalent to approximately one year of GDP growth for the EU according to the IMF's long-run forecasts. Over a ten-year period to 2030, the difference between a moderately liberalising path and a moderately restrictive path would amount to €2 trillion, in today's money.
- A **moderately restrictive scenario** is one in which the EU is unable to rely on GDPR transfer mechanisms, and in which trade partners increase their overall levels of restrictions on cross-border data flows. Our modelling suggests that such a **scenario leads to a reduction in EU exports of around 4%, and in 1% of GDP per year**. Cumulatively to 2030, losses amount to €1.3 trillion. The loss in output corresponds to around 1.3 million jobs in the impacted sectors.
- The majority of the pain is self-inflicted: **a majority of the EU's export losses in restrictive scenarios (around 58%) come from an increase in its own restrictions than from partner actions.** Domestic measures that increase data localisation act as a tax on a country's exports.
- **By contrast, if the EU and major trade partners adopted measures to facilitate cross-border data transfers**, **EU exports as a whole would**

grow by a little over 2% per year, adding 0.6% to GDP per year, which is around 0.7 million jobs in the impacted sectors. Cumulative effects to 2030 are worth around €720 billion.

- The downside risks from the moderate restrictiveness scenario outweigh the upside risks from the more liberalising scenario. The priority for the EU and its partners should in the first instance be **to avoid increased restrictiveness** by locking in existing levels of liberalisation.
- **The estimated GDP effects are conservative**. This is because the primary channel of impact that is modelled is via effects on international trade. Other channels of impact, such as costs to innovation are also important. The analysis is also based on trade and national account data that records transactions between firms, rather than internal flows of goods and services. The significant impacts of data restrictions on internal flows are therefore not captured. Finally, we do not model other policy measures (such as the restrictions on operations of non-national businesses) that might accompany data localisation.
- Data localisation requirements could also hurt sectors that do not participate heavily in international trade, such as healthcare. Indeed, research activities increasingly feature collaboration across borders: for example, it is estimated that in 2019 more than 5,000 collaborative projects between EEA countries and the US National Institutes of Health alone (NIH, 2019). Moreover, up to one-fourth of inputs into the provision of healthcare consists of data-intensive products and services.
- The impacts are likely to affect both large and small businesses. **In EU manufacturing, small and medium enterprises account for 23% of exports**. Therefore, a relatively crude apportioning of the results shown in section 4.3 suggests that exports by data-reliant manufacturing SMEs in the EU are worth around €280bn. In the challenge scenario, exports from EU SMEs would fall by €14bn, and in the growth scenario they would increase by €8bn.
- Assessing the impact of data flows on SMEs in service sectors is challenging due to data limitations. **However, we know that SMEs account for 61% of the turnover of EU data-reliant services sectors**. Therefore, it is likely that exports of data-reliant services from EU SMEs account for a significant proportion of the value of cross-border data flows to the EU economy.
- Our modelling also provides conservative estimates of the economic contribution of cross-border data flows to the EU. A **loss of cross-border data flows on exports from data-reliant sectors** would lead to an **annual reduction in EU GDP worth at least €330bn, or around 2.5% of total EU GDP**.

1 INTRODUCTION

1.1 Background

The cross-border flow of data plays a central role in an increasingly digitised economy. Cross-border data flows underpin modern international trade and investment, reflecting the wider role data plays in enhancing the operations of firms, and indeed in creating new business models based on the use, processing and storage of data. Cross-border transfers of data have facilitated the rise of global value chains, which according to the OECD account for close to three-quarters of the value of international trade.[2] These transfers have also facilitated the integration of services and manufacturing. In OECD countries, including the EU, between 25 and 40% of the value of manufacturing exports is accounted for by services inputs.[3] During the COVID-19 pandemic, cross border flows of data have helped to coordinate economic activity internationally, mitigate the adverse effects on trade, and support critical value chains such as medicines.[4]

The centrality of data to modern societies has also given rise to a number of public policy objectives. These include the protection of personal data, national security, the market power of digital networks, sectoral regulation matters, and industrial policy objectives. Some of these have led countries to develop policy frameworks that either restrict cross-border data flows, or have the potential to restrict by making these flows conditional on demonstrating compliance with policy objectives.

"Data localisation" refers to a range of policy interventions that restrict the extent to which data that is generated in a jurisdiction (e.g. by businesses, organisations or individuals) may be accessed, used or stored outside that jurisdiction. Data localisation involves to some degree restrictions on cross-border trade.

In view of the array of data policy issues discussed above, countries have also sought ways to maintain sufficient policy flexibility to regulate cross-border data flows, including via localisation requirements, while at the same time also seeking to secure the benefits of cross-border data flows. This gives rise to trade-offs, and data governance frameworks often reflect ways of trying to manage these trade-offs. The General Data Protection Regulation (GDPR) adopted by the EU reflects one particular type of approach: it makes liberalisation of cross-border data flows conditional on the trade partner demonstrating that its data governance framework provides equivalent levels of data protection to that accorded within the EU ("adequacy"), or on the existence of equivalent safeguards under other transfer mechanisms (e.g. standard contractual clauses and binding corporate rules). More generally many countries have sought to agree international rules, particularly through Free Trade Agreements, that eliminate restrictions on cross-border trade, and that attempt to limit the effects of measures, taken for other policy purposes, on cross-border data flows.

---

2 https://www.oecd.org/trade/topics/global-value-chains-and-trade/

3 https://www.oecd.org/industry/ind/tiva-2018-flyer.pdf

4 The terms 'cross-border data flows' and 'cross-border data transfers' are used interchangeably in this report.

The OECD has documented that the international context for cross-border data flows has become more restrictive over the last decade. Given the importance of data flows to trade, this raises the question of what costs such developments might impose on the EU should they persist, and, conversely, the benefits of securing more liberal arrangements.

1.2 Terms of reference

DIGITALEUROPE asked Frontier Economics (Frontier) to assess the extent to which the EU economy and the economies of EU Member States are reliant on cross-border data flows. The assessment is required to estimate how changes in policy restrictions on cross-border data flows (i.e. data localisation) would impact key economic indicators including GDP, employment and trade.

This report sets out our findings and the underpinning evidence. **Chapter 2** outlines the role of cross-border data flows in key sectors in the EU economy. **Chapter 3** sets out our modelling approach to estimate the impact of changes in policy restrictions on cross-border data flows on key economic indicators. **Chapter 4** reports the modelling results.

2 THE ROLE OF CROSS-BORDER DATA FLOWS IN THE EU ECONOMY

2.1 Uses of data in the modern economy

The use of data is an increasingly important component of all advanced economies. Businesses can generate value from data in several ways, summarised into four main economic purposes in recent work by the OECD, and described in the figure below.[5]

Figure 1 Ways in which businesses rely on data

Source: Adapted from Nguyen and Paczos (2020)

Cross-border data flows can occur as part of all four ways of using data:

- First, using data requires preliminary steps (a "data value chain"), and data may flow between different geographies between each step in the data value chain;
- Second, once the final step in that value chain is reached, using and monetising data may involve a cross-border exchange.

This applies to all forms of data, including personal data. In both cases, data flows may underpin monetary transactions between different organisations, or they may underpin operations that take place within the same company. This distinction matters for the measurement of the impact of data flows: in the former case, the exchange is recorded by official statistics on international trade as an import/export; in the latter, the exchange is often not recorded and measured outside of the organisation using the data. Therefore, as discussed in Section 3, we model the impact of data flows on international trade. This means that our estimates are conservative as they do not include the economic value of intra-company cross-border data flows.

Cross-border flows within the data value chain

Individual data points do not automatically generate value for firms, governments and consumers. Instead, the value of data is highly context specific. For example, the data of an individual person may be valuable to that person, but only hold broader value when aggregated with data from many other individuals and other sources of data.

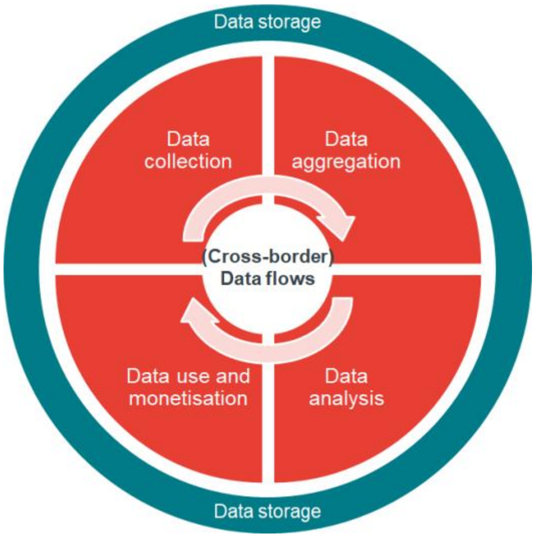
Value is extracted from data through a "data value chain": firms need to collect raw data (stage 1), aggregate the data (stage 2), and analyse the data (stage 3) before the data is finally used and monetised (stage 4). The use and monetisation stage may itself generate additional data, which is in turn aggregated, analysed, and

[5] https://www.oecd-ilibrary.org/docserver/6345995e-en.pdf?expires=1623000458&id=id&accname=guest&checksum=8319D67CA66D0C1DEBA1FFB53621747E

used. This feedback loop has been described as a **"**data value cycle**"** (see Figure 2).[6] Throughout this value chain, data may be stored in a different location compared to where the data is collected, aggregated, analysed and used.

Each step in the data value chain may take place in different geographies. Consider for example a hypothetical retailer headquartered in Germany. This retailer may commission a specialist market research company to collect data on its customers in Germany, France, and the UK (step 1); the data may then be aggregated and analysed by a data analytics provider based in Ireland (steps 2 and 3), and results from the analysis may inform strategic business decisions that are taken at the retailer's headquarters in Germany (step 4). In this example, each step in the data value chain involves international trade in services (imports/exports of services) and an associated flow of personal data between different companies. Even if the collection, aggregation and analysis, and use of data all took place within the same company (the hypothetical retailer), this data value chain may involve the same cross-border data flows. At the very least, data from customers in Germany, France and UK would be aggregated in one location, e.g. where the company's main analytics and insight function is based. Aggregation allows drawing insight from larger datasets and potentially learning from cross-country differences. The analytics and insight function may not be based in Germany, France or the UK: for example, this company may have decided that Dublin is the best and most cost-effective location to attract data analytics talent to the company.

Figure 2 The global data value cycle



Data use and monetisation involving cross-border data flows

Even when the collection, aggregation, analysis and storage of data all take place within one geography, using the data may involve cross-border data flows. A simple example is where the data is itself the product being sold. For example, consider a market research company based in the EU. This company may sell data

6 https://www.oecd-ilibrary.org/docserver/6345995e-en.pdf?expires=1623000458&id=id&accname=guest&checksum=8319D67CA66D0C1DEBA1FFB53621747E

on the habits of EU customers to retailers based elsewhere, for example a Japanese car manufacturer interested in designing new electric vehicles tailored to the EU market. Similarly, exporting and importing data-enabled products and services will typically involve cross-border data flows. For example, if a German manufacturer sells a “connected car” to a Spanish customer, this may involve data on the customer’s driving in Spain being transferred to Germany. Indeed, section 2.2.2 below provides examples of how digitalisation allows manufacturers to provide their goods to customers “as a service”: this involves an ongoing relationship with the customer with associated data flows.

On the opposite end of the spectrum described in Figure 1, cross-border data flows may also occur when data is used to improve an existing product or process. For example, an automobile parts manufacturer and an automobile manufacturer may exchange data across borders to ensure that levels of demand and supply are aligned, or in the research & development process, to ensure that the specifications of a new car part match the requirements of a new vehicle. This example is a very simple case of a global value chain. As already observed in the introduction, the growth of global value chains is a very important development in the evolution of the world economy. Developments in information technology have reduced costs of storing and transmitting data. Associated with falling transport costs and a lowering of trade barriers, these developments have stimulated a geographic unbundling of production. In other words, value chains are increasingly broken down into a large number of steps and components (“tasks”), and countries or regions increasingly specialise in such tasks.[7] In such circumstances, the cross-border flow of data is not only intrinsically tied to international trade, it is also a driving force behind the ways in which it develops.

Cross-border flows as a driver of international trade

Providing a comprehensive taxonomy of all the ways in which data flows across borders is beyond the scope of this report. However, the discussion presented so far shows how the popular perception that data flows only benefit search engines and social networks is misconceived. In fact, a range of sectors rely on cross-border data flows. And it is not just activities by firms that rely on cross-border data flows. Government agencies also need firms to be able to transfer data across borders as part of financial oversight, drug approval, law enforcement, counterterrorism, and other responsibilities.

In summary, cross-border data flows facilitate international trade in the following ways:

- Across all ways in which data is used, cross-border data flows allow **trade in the data value chain**: data collection, aggregation, analysis, and use can be imported or exported;
- **Data can be exported**: sold or licensed to customers abroad;
- Exports and imports of **data-enabled products and services** (e.g. connected cars, online platforms…) typically require cross-border data flows;

[7] See notably D.K Elms and P. Low (eds) 2013, Global Value Chains in a Changing World,

- When data is used **to improve existing products or processes**, cross-border data flows allow:
  - **Coordinating operations across global value chains:** many sectors feature operations that span international geographies, and a significant part of their trade is internal to these value chains;
  - **Facilitating the flow of inputs of goods and services** from different parties in different locations across the value chain, and notably facilitating the integration of services and goods (servicification), whether this takes the form of services inputs (e.g. software services in vehicles) or the bundling of services and goods (GPS data services and insurance services bundled into the sale of a vehicle). See also the discussion in section 2.2.2 below.
  - **Reducing transaction costs** with end-users and customers in export markets since the initiation and completion of cross-border transactions between buyers and sellers typically involves transferring some data, e.g. regarding contracts or specifications of orders, packaging and delivery. For services sectors in particular, cross-border data flows can increase the scope for supplying services remotely, limiting the need to avoid the fixed costs associated with establishing a physical presence.

2.2 Case studies of how different sectors rely on cross-border data flows

To illustrate the importance of cross-border data flows to the EU economy, consider the role these data flows play in two sectors: Pharmaceuticals and Healthcare, and Manufacturing.

2.2.1 Pharmaceuticals and Healthcare

Cross-border data sharing plays an important role across many activities within the Pharmaceuticals and Healthcare sectors. This includes the screening of chemical compounds, the optimisation of clinical trials, and in the post-market surveillance of drugs, leading to new treatments, improved patient outcomes, and lower costs.[8] While health information is a particularly sensitive category of data, its potential to improve and save lives also makes it highly valuable.

In the case of medical research the cross-border sharing and aggregation of data, including genetic and other health-related data, is an essential part of improving healthcare and disease prevention, in terms of ensuring sufficiently large sample sizes, identifying complex pathways, and comparing the determinants and outcomes of disease in different settings (ALLEA, EASAC and FEAM, 2021).

This is particularly the case for rare diseases. Without data flows and data aggregation it becomes significantly harder to identify and collect enough data in each and every country, given many countries only have small patient populations.

International co-operation in life sciences is widespread. It is estimated that in 2019 there were more than 5,000 collaborative projects between EEA countries and the US National Institutes of Health alone (NIH, 2019). In addition, the U.S. Food and

[8] https://itif.org/publications/2021/03/25/how-build-back-better-transatlantic-data-relationship

Drug Administration's (FDA) clinical trials database shows that more than a fifth of industry funded clinical trials were engaging patients on both sides of the Atlantic (including 6 COVID-related clinical trials).

As explained in Kepes et al. (2021), there are pharmaceutical companies which carry out their clinical trials in the EU and rely on cross-border data transfers of get authorisation for new treatments from regulatory authorities outside the EU. If this is not possible, these pharmaceutical companies may need to duplicate their research investment, carrying out clinical trials both within and outside the EU.

In addition, cross-border data transfers have been critical during the global coronavirus pandemic. Without such data transfers, the development of COVID-19 vaccines likely would have been delayed, with significant economic and health impacts (Kepes et al., 2021).

2.2.2 Manufacturing

The global nature of all facets of manufacturing – including research and development, supply chain management, design, production, sales, and customer and after-sales support – and the advent of smart manufacturing mean that modern manufacturing firms rely heavily on cross-border data flows.

Smart manufacturing is the application of information and communication technologies to manufacturing processes.[9] Through the aggregation and processing of data, smart manufacturing provides manufacturers with a comprehensive view of what is happening across the production system, along with the insights to make real-time adjustments in order to optimise production. This enables firms to effectively coordinate their research and development, supply chains, production, sales, and post-sales processes. A "plugged in" manufacturer can receive real-time information from suppliers to adapt to supply chain disruptions or use data analytics from across the supply chain to adjust to meet shifting demand.

The role of data flows within smart manufacturing is illustrated by the way manufacturers use key enabling technologies such as cloud computing and sensor technologies.

Cloud computing and sensor technologies

Cloud computing is the delivery of on-demand computing services such as applications, storage and processing power typically over the internet, and it allows firms to store and process the data necessary to manage their manufacturing operations, and use new production systems, from 3D printing to the Internet of Things (IoT) and industrial robots.

Sensors embedded within devices, machines, and products themselves measure everything from output, consumption, wear, and capacity to salient operating conditions such as temperature, humidity, and electrical flow, and play a key role in creating the information streams upon which smart manufacturing techniques rely. Over the past 15 years, the cost of sensor technologies has declined more

[9] Smart manufacturing is also known as "Industry 4.0". The term originated in 2011 from a project in the high-tech strategy of the German government, which promoted the digitalisation of manufacturing.

than a hundredfold.[10] This highlights that cross-border data flows are not just consumer-to-consumer or business-to-consumer. Cisco estimated that out of the approximately 18.4 billion networked devices in use in 2018, 24 per cent served business customers, and a growing number of business-to-business (or machine to machine) applications, such as smart meters, transportation, and package and asset tracking are now major drivers in the growth of Internet-connected devices.[11]

Cross-border machine-to-machine data flows allow manufacturers to move towards a predictive and preventative maintenance and repair model, rather than a “repair and replace” model, both on the factory floor and for products deployed in the field. The McKinsey Global Institute has estimated that the use of predictive maintenance techniques reduces factory equipment maintenance costs by up to 40 percent, while reducing equipment downtime by up to 50 percent, and capital-equipment investment costs (to replace defective equipment) by 5 percent.[12]

Sensor data can also have personal components. For workers this can include sensors that track their movements on the factory floor. For example, sensors in wristbands can be used to track where warehouse employees are placing their hands. This can feed into performance assessment, help manufacturers to optimally design the layout of factories, and even (via vibrations) nudge workers to take certain actions in real time. For example, wearable sensors can leverage personal data to predict and prevent injuries in the industrial workplace be notifying workers when they make dangerous movements.[13]

Sensors can also be used to transfer personal data from final customers. For example, automotive manufacturers rely on transfers of personal data on driving behaviour and geolocation. Several manufacturers, including Audi, Mercedes and Volvo, even offer drowsiness detection systems that monitor a vehicle's movements, such as steering wheel angle, lane deviation, time driven and road conditions.[14]

Data-reliant products and the servicification of manufacturing

One of the consequences of the trends described above is that services have become closely integrated into manufacturing, a phenomenon that goes by the (rather unwieldy) title of “servicification”. The phenomenon can be broken down into various aspects:

---

[10] http://www2.itif.org/2016-policymakers-guide-smart-manufacturing.pdf?_ga=2.229420636.1748280977.1622887938-1338740611.1622720210

[11] https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

[12] https://www.mckinsey.com/~/media/McKinsey/Industries/Technology%20Media%20and%20Telecommunications/High%20Tech/Our%20Insights/The%20Internet%20of%20Things%20The%20value%20of%20digitizing%20the%20physical%20world/Unlocking_the_potential_of_the_Internet_of_Things_Executive_summary.pdf

[13] https://unreasonablegroup.com/companies/strongarm-tech

[14] https://www.nytimes.com/2017/03/16/automobiles/wheels/drowsy-driving-technology.html#:~:text=Several%20manufacturers%2C%20including%20Audi%2C%20Mercedes,time%20driven%20and%20road%20conditions.

- The role of services inputs into the value added of the final manufactured products. These inputs can include R&D, IT and software, engineering services and so forth.
- The bundling of services and goods. Examples include after sales services plans and financing plans. Data services can also be embedded into goods. For example, a car may contain GPS systems and software, and this in turn enables the purchaser to conclude data tracking services with a third-party provider.

One specific example of how servicification can change business models can be found in Kaeser Kompressoren, a German-based manufacturer of compressed air systems and services. This business launched an “air-as-a-service” business model in which customers no longer purchase Kaeser compressors but rather lease the compressors and pay only for the compressed air itself. It means customers can scale consumption up or down as the needs of their manufacturing operations change, without needing to purchase new equipment.

The aviation and automotive sectors are often presented as paradigmatic examples of servicification via data, and we consider the cases of Scania and Airbus in the boxes below.

SCANIA

Scania is a global company offering sustainable transport solutions with a focus on trucks, buses, and engines. Its headquarters are in Södertälje (Sweden), as are its R&D, production, purchasing, sales, and IT systems. Scania also has production sites in Brazil, Netherlands, Argentina, France, and Poland as well as regional production centres in six other countries (all outside the EU).

When a Scania vehicle is driven a small box sends diagnostic data – speed, fuel use, engine performance, even driving technique – to the company's headquarters in Sweden.

The vehicles would not function effectively without transferring data, and neither would the repair of these vehicles. If a vehicle breaks down, data can be transferred to a regional or global help desk for help in tracking and solving the problem. The 'old repair manual is replaced by a global data base' and hence effective repairs hinge on data transfers.

Indicative of the growing role of software and data, Scania is developing an open, brand-neutral platform (the RIO platform) and operating system for software that will host its telematics services as well as software from third parties, thus making it easier for its partners and other transport manufacturers to adopt and use.

Scania has likewise increasingly transitioned to a services-based business model focused on fleet management services including logistics, repair, and others. For example, facilitated by the collection and sharing personal and nonpersonal data, Scania offers an educational service that coaches drivers to help them improve, such as with braking points and coasting. The purpose is to continuously coach the driver on how to operate the vehicle in a more efficient and environmentally friendly way. Overall Scania now generates one-sixth of its revenues through new services enabled by the data-connected devices built into its vehicles.

AIRBUS

Airbus is the world's largest airliner manufacturer. Headquartered in the Netherlands it has around 180 locations globally and 12,000 suppliers.[15] Each year Airbus' aircraft enable millions of passengers to travel across cities, countries, and continents.

The collection, aggregation, analysis and use of data has become fundamental to the airline industry – the latest Airbus A350 has 50,000 sensors on board collecting 2.5 terabytes of data every day.[16]

Airbus now believes that the *"aviation industry is now entering into its next revolution: connected, fully-digital aviation."*[17] For example, in June 2017, Airbus launched Skywise, an open data platform that aggregates and processes large volumes of data from aircraft sensors, as well as data operational and maintenance data recorded by pilots, flight attendants, engineers, maintenance technicians and airport representatives. Today, Skywise hosts data from Airbus, suppliers and more than 100 airlines including EasyJet, Emirates and AirAsia.

The data – and the cross-border data flows that unpin the collection, aggregation and processing of the data – enable services that provide real-time and longer-term benefits for Airbus, airlines, suppliers and consumers. Real-time remote access to in-flight data helps in identifying problems early on and enable maintenance crews on the ground to be prepared for when the plane arrives, reducing turn-around times, delays and cancellations.[18] Longer-term, Skywise data helps Airbus understand how their aircrafts behave in reality (across different geographies and business models) and to improve the design of new aircrafts models.[19] Better aircraft will decrease operational cost, and increase reliability, fuel efficiency and safety.

[15] https://www.airbus.com/company/we-are-airbus.html

[16] https://www.airbus.com/public-affairs/brussels/our-topics/innovation/data-revolution-in-aviation.html

[17] https://www.airbus.com/public-affairs/brussels/our-topics/innovation/data-revolution-in-aviation.html

[18] https://www.airbus.com/newsroom/press-releases/en/2019/04/airbus-launches-skywise-health-monitoring-with-us-airline-allegiant-air-as-early-adopter.html

[19] https://digital.hbs.edu/platform-digit/submission/skywise-airbus-bet-on-big-data/

3 APPROACH TO MODELLING THE IMPACT OF DATA LOCALISATION

3.1 Overview of approach

Our approach includes the four steps set out in Figure 3 below.

Figure 3 Overview of approach

Draw on OECD work on digital intensiveness
- Data either the core element of the business model
- Or key enabler in operations and value chains

- A numerical measure [0 to 1] capturing a range of policy interventions and regulations
- Policies affecting data are one subset.
- Measure effects of restrictions on data flows, once other factors are controlled for
- Changes in policy lead to a change in STRI value by sector by country
  - Increase= more restrictive

- Specify a baseline scenario that reflects existing levels of liberalisation/ restrictiveness in EU and partners
- Liberalising scenario reflects decrease in STRI values relative to baseline
- Restrictive scenario reflects increase in STRI values relative to baseline
- Changes apply to both EU and partners

- Use gravity model of trade to estimate effects of STRIs changes on trade between EU and partners
- Growth effects calculated based on relationship between trade and GDP

Source: Frontier Economics

The first step involves identifying sectors that are particularly data reliant, and specifically reliant on cross-border flows of data, by virtue of their participation in international trade and value chains (see also the analysis in section 2). This is consistent with our observations in section 2. This approach recognises that while data use is pervasive across all sectors, sectors vary in their degree of reliance on data and cross-border data flows specifically. Data localisation implicitly acts as a tax on cross-border data flows (see also the discussion in sections 3.3 and 3.4), and this tax will be disproportionately borne by some sectors: those that are more heavily reliant on cross-border data flows, with flow through consequences for the economy as a whole.

The next step consists of identifying the extent of restrictiveness of data policies. There are various ways of doing this. One is to collect data on actual costs faced by businesses as a result of complying with data localisation requirements. Another, which we adopt here in view of time constraints and tested robustness, is the OECD's services trade restrictiveness index. This allows us to develop a numerical representation of a qualitative question – the degree of policy restrictiveness.

The third step is to specify the scenarios for analysis. Quantitative analysis of policy changes are counterfactual exercises: in this case we wish to compare outcomes between projected future scenarios relative to a status quo baseline.

The last step is to model the outcome of these future scenarios using a gravity model of trade. This allows us to capture the effects of changes to bilateral trade between the EU and each of its main partners in the sectors of interest, as a result of changes to restrictions in cross-border data flows, once other factors are controlled for. These other factors include country fixed effects (e.g. size) and policy variables. The modelling of trade effects provides a channel through which to estimate growth effects, given the dependency of growth on trade.

We explain each of these steps in more detail, and the insights they yield, in the sections below.

3.2 We identify sectors that are particularly data reliant

We focus specifically on sectors that are data reliant. To identify these sectors we draw on the OECD's grading of sectors based on digital intensity, which itself is based on seven criteria: ICT investment, software investment, ICT intermediate goods, ICT intermediate services, robot use, online sales and ICT specialists)[20] and apply a filter to remove sectors that are rarely traded internationally.[21] Ideally, this assessment would rely on information on the amount and characteristics of data use in each sector. However, as mentioned earlier in this report, there is limited available information on this. Therefore, the measures above are the best available proxies on which sectors can be systematically compared across sectors. The sectors selected through this process are involved in international trade (to varying degrees), the only exception being public administration, which is thus excluded from the analysis.

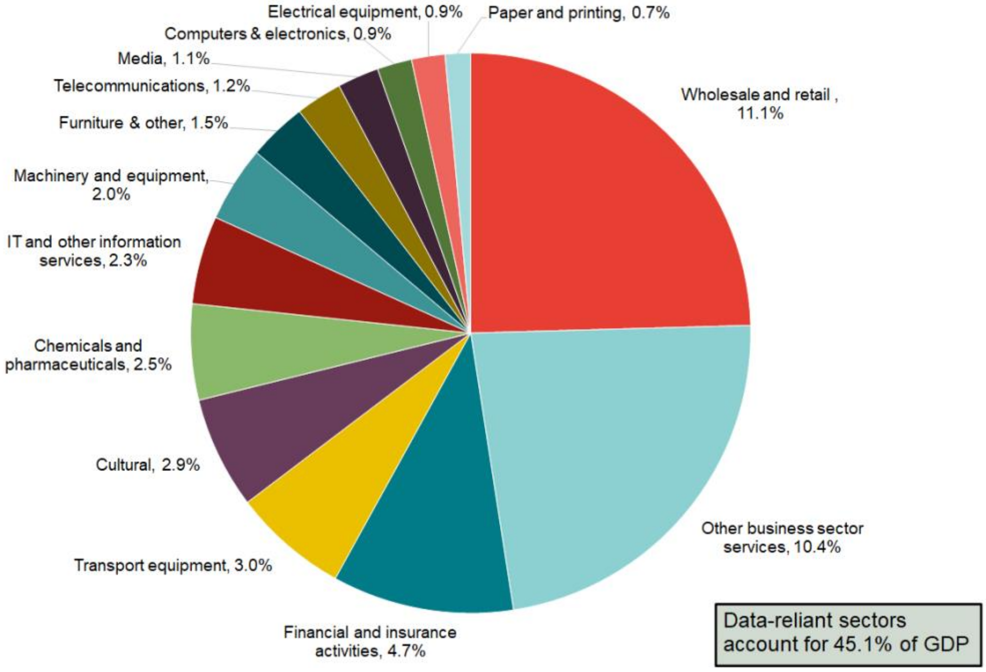
Based on this approach, the list of data-reliant sectors is set out in Figure 4, along with their shares of EU GDP and employment. The sectors as a whole account for 45% of EU GDP. Of this, 11.4% is attributable to data-reliant manufacturing sectors (in particular, the manufacturing of transport equipment, chemicals and pharmaceuticals, machinery and equipment, furniture, computers and electronics, electrical equipment, and paper and printing).[22]

[20] Calvino, F., et al. (2018), "A taxonomy of digital intensive sectors", OECD Science, Technology and Industry Working Papers, No. 2018/14, OECD Publishing, Paris, https://doi.org/10.1787/f404736a-en.

[21] Sectors are considered data-reliant if they are in the top quartile for at least two of the seven criteria or in the top two quartiles for at least four criteria.

[22] Note that the trade modelling includes impacts across *all* services sectors, including transport and construction. Although they fall outside the data-reliant category, they still exhibit some sensitivity to services trade restrictions (albeit less than other services sectors), so are included in the trade modelling.

Figure 4 Data-reliant sectors and their shares of EU GDP



Source: Frontier Economics analysis of OECD Trade in Value Added Database

Source: Other business sector services includes Legal & accounting, Scientific R&D, Marketing and Administrative services. The sector shares are as reported in TiVA for the year 2015. While subsequent modelling applies growth in line with overall IMF statistics and forecasts, the sector shares are not assumed to change.

But that does not capture the value of cross-border data flows to the EU. To develop an understanding of that value, we look at how these sectors are impacted by changes to the levels of restrictiveness applied to cross-border data flows via the effects on trade between the EU and major partners involving these sectors.

We focus on trade for the reasons documented in section 2: cross-border data flows play an important role in facilitating international trades, and restrictions on cross-border data flows increase the costs of doing international trade.

3.3 We measure restrictions on cross-border data flows affecting these sectors

We capture the extent of restrictions on data through the OECD's services trade restrictiveness index (STRI). The STRI scores countries by assessing various aspects of regulation that affect cross-border trade. While the focus is on services trade, the STRI is also relevant to goods given the interdependence between goods and services, a feature that has been enhanced considerably through digitisation and data.

The STRI includes a specific component on cross-border restrictions on data flows, which captures different types of restrictions that if implemented increase the overall level of restrictions. Each of these restrictions is weighted equally for any given sector, though weights differ by sector. Adding (removing) any one restriction increases (decreases) overall restrictiveness by the same proportion as any other data restriction in that sector.

These specific restrictions are:

- Cross-border transfer of personal data is possible when certain private sector safeguards are in place;
- Cross-border transfer of personal data is possible only to countries with substantially similar privacy protection;
- Cross-border transfer is subject to approval on a case-by-case basis;
- Certain data must be stored locally; or
- Cross-border transfer of data is prohibited.

Each sector within a country is assessed according to these aspects, which can be cumulative. The higher the STRI score, the more restrictive a country is.

In terms of interpretation, the STRI should be understood as a tool for comparing the levels of restrictiveness across different countries and sectors, and to understand the specific types of restrictions that contribute to it. The components in it, and the weights attached to them, are arrived at in a detailed deliberative procedure, leveraging a consensus view across sector experts. The five restrictions specified in relation to data in the STRI will not capture all the nuance around restrictions on data, and there is scope for some overlap between them. They are intended to give a summary view of how different countries and sectors compare. From this we can derive indicative estimates of how trade would change if countries increased or decreased their restrictiveness.

3.4 We model two scenarios with different restrictions on cross-border data flows

3.4.1 Scenarios

We use a gravity model of trade which includes the STRI as one variable explaining bilateral trade flows in order to capture the effects of changes in data restrictiveness on trade between the EU and partners in the sectors of interest.

The modelling requires specifying a baseline against which changes to the degree of restrictions on cross-border data flows under hypothetical scenarios can be measured. The change in the number of STRI restrictions in place is used to compute changes in STRI score for both EU and trading partners, from which percentage trade impacts are calculated. The STRI changes under the different scenarios can be summarised as follows:

Figure 5 STRI changes by scenario

Country group	Challenge scenario:	Growth scenario
China, India and Russia	All 5 data restrictions apply’	Removing data restrictions so only one remains active
EU and other countries	Two new data restrictions are added	One data restriction is removed

Source: Scenario assumptions

The baseline

This represents current policy settings in the EU and major trade partners. We assume that existing GDPR data transfer mechanisms apply, although under increased scrutiny, and that the UK is granted adequacy. We assume that other trade partners maintain levels of restrictiveness reported as at end 2020.

The challenge scenario

In this scenario, global levels of restrictiveness increase. Liberal data regimes (the EU, the US, Canada, Japan, and Switzerland, for example) become moderately restrictive, and more restrictive regimes (India, Russia and China, for example) become highly restrictive. In the EU, GDPR transfer mechanisms generally cannot be relied on, and additional restrictions are introduced for international data transfers under the EU's data strategy (e.g. Data Governance Act and Data Act). This includes the possibility of requiring businesses that are data intermediaries to establish a legal representation in the EU (a form of localisation of business presence in addition to any data localisation requirements per se). In practical terms this means the restrictive countries move to the highest restrictiveness level in our model (five elements of STRI restrictions), while other countries add two STRI elements.

Growth scenario

Major trading partners undertake commitments to eliminate restrictions on cross-border data flows. Flexibility to pursue public policy objectives is retained, but countries accept disciplines that seek to ensure that public policy measures are no more trade restrictive than necessary to achieve the policy purpose that is sought. We assume that for restrictive jurisdictions, there is greater level of flexibility (reflecting their status as developing countries) so their restrictiveness does not converge to those of developed countries. In addition, any remaining uncertainty around GDPR transfer mechanisms is eliminated, and an enhanced Privacy Shield is negotiated with the US and upheld by the Court of Justice of the EU.

In terms of representing this scenario in the modelling framework, we assume that most countries shed one element of restrictiveness. For example, if jurisdictions, under the baseline, require that partners must have substantially similar privacy protections as a pre-condition for liberalising data transfers, we assume that this requirement is lifted in the growth scenario.

By contrast, we assume that the currently restrictive countries, notably China, India and Russia, substantially liberalise. Specifically, they converge to one line of STRI restriction. This is the same level that most OECD and EU countries are at now. But as China, India and Russia together account for a fairly small share of EU trade, the growth scenario is therefore relatively incremental. In practical terms the growth scenario involves the restrictive countries moving to one line of restriction, while other countries remove one line.

3.4.2 Use of gravity models

The technical annex explains these models in greater detail. Gravity models are the workhorse of international trade analysis. They model trade between pairs of countries as a positive function of their respective size and inversely proportional to the square of the distances between them. Distance acts as a basic variable of trade cost. Other variables that affect trade costs can be included, such as linguistic differences, institutional quality and trade variables. Trade variables include the STRI. We can thus capture the effects of changes to data restrictiveness once other factors are accounted for.

The key parameter of the gravity model, for our purposes, is the elasticity that captures the responsiveness of bilateral trade flows in the sectors of interest to changes in the STRI. In practical terms, we run the model for each scenario: in each scenario, the STRI assumes a different value in line with the level of restrictiveness assumed for the country/ jurisdiction in question.

The impact on a country's exports is the product of the increase in partner restrictions and the increase in *own* restrictions. The latter effect reflects the fact that when a country imposes localisation measures, it affects its own firms as well as foreign ones. Moreover, because localisation increases trade costs, it biases production towards home markets, i.e. it functions as an export tax.

3.5 Limitations

The overall approach is conservative, in the sense that it is likely to understate the economic effects of changes to restrictions. This is for several reasons:

- The STRIs may not capture all aspects of policy that affect data flows. That is, there may be other aspects of data policies that are not reflected in the specific elements of the STRI. For instance, limitation on the legal structure of data providers or ownership restrictions are not captured by the data segment of the STRI. Similarly, indices such as the STRI are not equipped to capture the effects on businesses of uncertainty concerning future policy changes.
- Building on the above, some aspects of data regimes (e.g. the draft Data Governance Act) also specify localisation requirements in terms of business operations, e.g. establishing a commercial presence and registering as a local entity for data intermediaries. These restrictions *are* captured by the STRI, but we do not include them in the model as we seek to focus primarily on restrictions imposed on data flows.
- Our approach uses trade and national account data, which is based on transactions whose value can be observed because they involve a flow of money between parties. However, data flows are also important to activity within firms (e.g. between the software and research division of a firm, and its production division). These activities are significant but are not captured in trade data. Moreover, national account data will not capture non-market benefits (such as health and environmental benefits) that are enabled by the activity of data-reliant sectors (e.g. the contribution of engineering services and wind turbines to reducing greenhouse gas emissions).

- The relationship between GDP and trade reported in this analysis reflects some of the longer-run effects of data restrictions, via trade, on innovation and productivity. However, our approach does not explicitly address the underlying mechanics of these relationships or some of the broader effects of data flows on innovation For example, research and development activities and product testing often occur through cross-border value chains that involve large scale flows of different types of data (industrial, IP, and personal data). While this is captured in the observed relationship between productivity and trade, it is not possible to see the mechanism of action. In addition, as these are average effects looking at trade overall (rather than the effects of data), the impact in the context of European countries and sectors might be more specific.
- Our modelling approach implicitly assumes that the proportion of activities in the EU economy that are “data reliant” remains constant over the next decade. However, it is likely that further digitalisation of the EU economy will take place, meaning both that more sectors will be data reliant, and that the importance of data may increase further in sectors that already are data reliant. As a result, the importance of cross-border data flows may be greater in the future.

4 THE IMPACT OF DATA LOCALISATION ON TRADE AND ECONOMIC GROWTH

In this section, we draw on the key role that cross-border data flows play in modern international trade to measure their overall economic contribution. We measure the impacts of policy scenarios that lead to changes in the extent of data localisation compared to current settings. Section 4.1 models the impacts of the policy scenarios described in the section 3.4 on trade providing detail in terms of countries and sectors. Section 4.2 reports estimates for GDP and employment by policy scenario, based on the estimated trade effects. Section 4.3 provides further evidence in terms of the exporter firms affected and the trading partners. Section 4.4 considers the hypothetical case in which cross-border data flows are prohibited, as a way of measuring the overall economic contribution they currently make to the EU via trade. Section 4.5 presents some concluding observations.

4.1 Trade impacts of policy scenarios

4.1.1 Challenge scenario

Definition

Under the challenge scenario, previously liberalised jurisdictions become more restrictive. In the EU, GDPR transfer mechanisms cannot be relied on to a large extent, and additional restrictions are introduced for international data transfers under the EU's data strategy (e.g. Data Governance Act and Data Act). This includes the possibility of requiring businesses that are data intermediaries to establish a legal representation in the EU (a form of localisation of business presence in addition to any data localisation requirements per se). Hitherto restrictive jurisdictions move to the highest level of restrictiveness possible, by mandating data localisation. In line with the relatively conservative approach taken, we do not assume any changes to other (non-data) aspects of policy.

Main results

Table 1 reports the main results for this scenario. The table below shows the impact on EU exports for the various sectors, first in absolute terms, then as a percentage of sector exports (to see which are relatively more affected), and finally as a percentage of total exports (across all sectors) to gain a sense of their overall impact. **Overall, we see annual impacts of €116bn,** which represent **4% of total (ex EU) exports**. To put this into perspective, **€116bn is equivalent approximately to 20% of the fall in exports experienced by EU economies in 2008, linked to the global financial crisis**.[23]

Manufactured goods account for **just over half of this impact.** The sectors most heavily impacted in terms of the effect on their exports are **IT, media, telecoms and cultural services (which all see a 10% or greater decrease in their**

[23] According to World Bank data, EU exports fell by 12% (€601bn in today's money) in 2008 compared to 2007.

exports). These sectors are particularly sensitive to restrictions in cross-border data flows.[24]

Table 1 Challenge scenario: sectoral impacts

Name	Annual impact (€m)	% of sector ex EU exports	% of total EU (ex EU) exports
Construction	-125	-2.4%	0.00%
Wholesale and retail	-6657	-2.3%	-0.23%
Transportation and storage	-6573	-2.5%	-0.23%
Media	-1758	-11.5%	-0.06%
Telecommunications	-1692	-8.2%	-0.06%
IT and other information services	-9926	-12.4%	-0.35%
Financial and insurance activities	-8725	-8.3%	-0.30%
Other business sector services	-17421	-8.7%	-0.61%
Cultural	-1693	-10.3%	-0.06%
Manufactured goods	-61137	-5.1%	-2.13%
Total	**-115707**	**-5.3%**	**-4.0%**

Source: Frontier Economics analysis of OECD Trade in Value Added (TiVA) Database

Note: Figures expressed in 2020 EUR, and include rescaling from 2015 TiVA year to 2019 baseline year

For the EU, a greater proportion of its own export losses (around 58%) comes from an increase in its own restrictions than from partner actions.[25] This reflects the fact that a loss of GDPR transfer mechanisms affects domestic and foreign firms. Moreover, domestic suppliers typically rely on competitively priced inputs. Imposing costs on the ability of business to access these impedes their competitiveness on global markets.

Figure 6 Impact of EU and trading partner restrictions

Source: Frontier Economics analysis of OECD Trade in Value Added (TiVA) Database

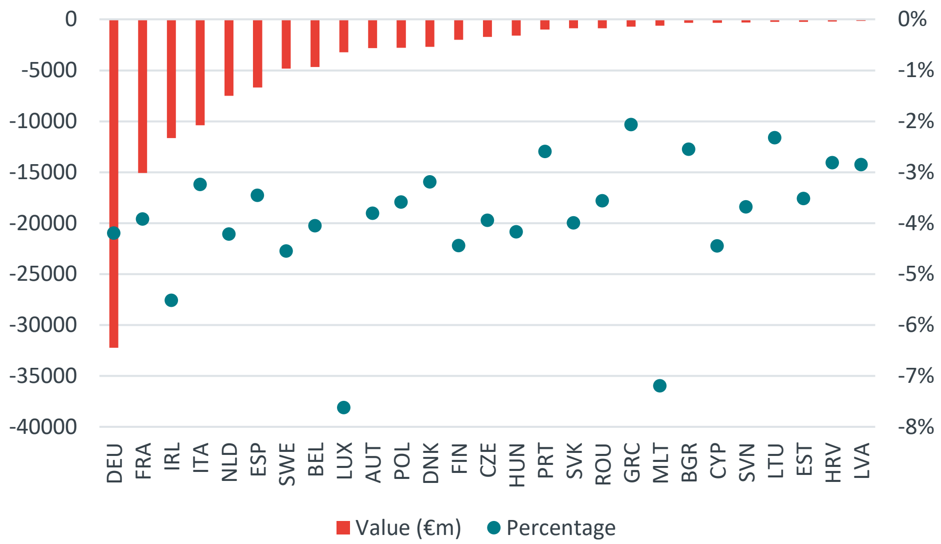
[24] The data component of the STRI carry a greater weight in these sectors than in others.

[25] This is estimated by comparing the magnitude of impact when 1) only the EU STRI is increased 2) both EU and trading partners increasing their STRI.

We can link this impact analysis to the various aspects of the EU's policy agenda. The impact on manufacturing highlights the effects of data localisation, and especially the EU's own policy stance, on the EU's strategy for an industrial renaissance and a competitive positioning based on high value-added activities. The effects on media and cultural sectors highlight the effects on sectors to which the EU attaches considerable importance from a social as well as economic perspective.

We can also break down impacts by country. These effects depend mainly on whether the country's exports are more reliant on sensitive sectors, as well as the mix of trading partners. Germany and France undergo export reductions of 4.2% and 3.9% respectively. The biggest impacts seen are for Luxembourg (-7.6%) and Malta (-7.2%), reflecting their reliance on services exports.[26] This is shown in Figure 7.

Figure 7 Trade impacts by EU country in the challenge scenario



Source: Frontier Economics analysis of OECD Trade in Value Added (TiVA) Database

Note: Figures expressed in 2020 EUR, and include rescaling from 2015 TiVA year to 2019 baseline year

4.1.2 Growth scenario

Summary of approach

Under this scenario, major trading nations enter into regional or plurilateral arrangements. These arrangements provide a framework for:

- Progressively reducing impediments to cross-border data flows; and
- Managing potential trade-offs between wider policy interests and the trade-enhancing benefits of liberalisation.

For the EU and major developed trading partners, who currently have low levels of restrictions on cross-border trade, the impact of these scenarios is likely to be relatively limited in terms of further liberalisation against the existing baseline.

[26] Unless otherwise mentioned, exports impacts are expressed as a proportion of exports outside of the EU, i.e. exclusive of intra-EU trade.

However, the arrangements contemplated in these scenarios will limit the scope for discretionary policy changes that lead to increased levels of restrictiveness. In particular, where countries seek to curtail cross-border data flows, the arrangements would require them to be able, if challenged, to justify the measures taken in view of the effects the measures might have on trade.

We assumed that developing countries, in line with general practice in trade agreements, are granted a higher degree of flexibility in implementing restrictions. This means that while their overall restrictiveness drops sharply under this scenario (and more so than for developed countries), it remains higher than developed countries.

Results

EU exports increase by around 2%. As in the challenge scenario, Manufactured goods account for the biggest contributor in absolute terms, while IT, media and culture gain especially significantly in proportionate terms. The results could help to contribute to the EU's Digital Decade objectives, particularly the development of international partnerships.

The positive impacts in the growth scenario are around half the size of the losses in the challenge scenario. This indicates there is more risk from slipping from current levels to less liberalised than there is scope to further liberalise. That in turn points to the value of trade agreements in locking in existing liberalisation.

Table 2 Growth scenario: sectoral impacts

Name	Annual impact (€m)	% of sector exports	% of total EU (ex EU) exports
Construction	64	1.2%	0.00%
Wholesale and retail	3428	1.2%	0.12%
Transportation and storage	3288	1.2%	0.11%
Media	1004	6.6%	0.04%
Telecommunications	1119	5.4%	0.04%
IT and other information services	5476	6.9%	0.19%
Financial and insurance activities	4530	4.3%	0.16%
Other business sector services	8814	4.4%	0.31%
Cultural	919	5.6%	0.03%
Manufactured goods	32975	2.7%	1.15%
Total	**61617**	**2.8%**	**2.15%**

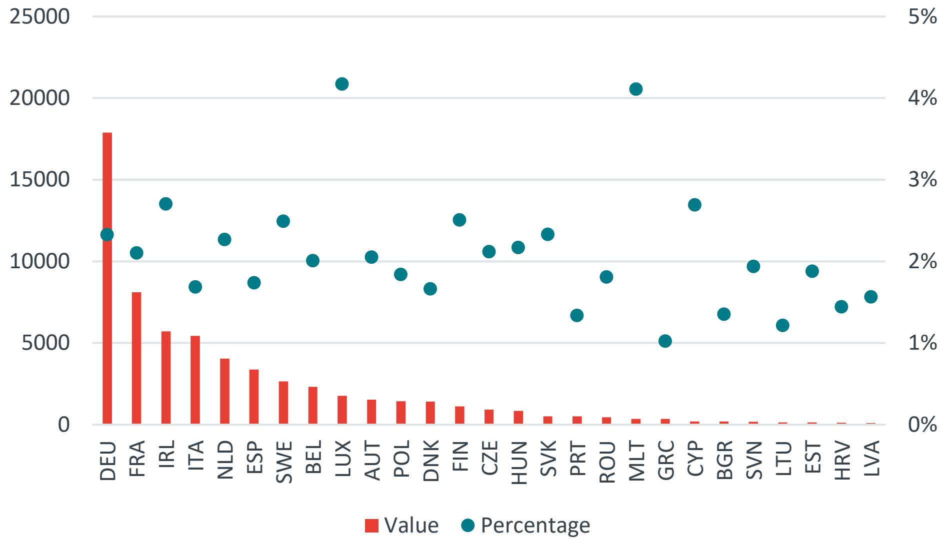
Source: Frontier Economics analysis of OECD Trade in Value Added (TiVA) Database

Note: Figures expressed in 2020 EUR, and include rescaling from 2015 TiVA year to 2019 baseline year

On balance the results suggest that the downside risk from more data restrictions (under the challenge scenario) is greater than the upside opportunities from liberalisation (under the growth scenario). This in part reflects the fact that many jurisdictions of importance to the EU are already relatively liberal and that bilateral data flows with these are relatively liberalised. This limits the incremental gains from further liberalisation, but also means that the losses from the introduction of new restrictions, particularly the EU's own, are greater.

The relative impacts by country are similar to the challenge scenario, as they reflect the size of the exporting economies and their sector mix. Germany and France see export increases of 2.3% and 2.1% respectively. Luxembourg (+4.2%) and Malta (+4.1%) are most affected, due to their reliance on services exports. This is shown in Figure 8.

Figure 8 Trade impacts by EU country in the growth scenario



Source: Frontier Economics analysis of OECD Trade in Value Added (TiVA) Database

Note: Figures expressed in 2020 EUR, and include rescaling from 2015 TiVA year to 2019 baseline year

4.1.3 Impacts by partner country

We can also explore which trading partners drive the impacts in the different scenarios. This reflects several factors:

- The partner's share of EU trade, which will be a function of size, distance and complementarity of their economies;
- Sector mix, i.e. whether the EU exports are in sectors that are more sensitive to data;
- How much room there is for further liberalisation / restriction.

In Figure 9 we show the export changes for the top 10 trading partners in the challenge and growth scenarios. Together these countries account for around two-thirds of the impacts.

Figure 9 Changes to EU exports by trading partner and scenario (€bn)

Source: *Frontier Economics analysis of OECD Trade in Value Added (TiVA) Database*

Note: *Figures expressed in 2020 EUR, and include rescaling from 2015 TiVA year to 2019 baseline year*

4.2 Impact on GDP and employment

4.2.1 Impact on GDP

Changes in exports would lead in turn to changes in the Gross Domestic Product[27] of the EU, as trade leads to higher productivity, innovation, and diffusion of knowledge. This is confirmed by a range of empirical estimates that we use to quantify the impact of cross-border data restrictions on GDP. The relationship between trade and GDP is captured by the elasticity of GDP to changes in trade: the percentage in GDP resulting from a percentage change in trade. A range of empirical estimates have been presented for this elasticity. We use a mid-point assumption that a 1 percentage point change in trade is associated with a 0.5 percentage point in GDP (between HMT range and Feyrer).[28]

Looking at the two scenarios, we see annual GDP impacts of -€139bn in the challenge scenario (-1%) and +€74bn in the growth scenario (+0.6%).

The difference between a path that is moderately liberalising (the growth scenario) and one that is moderately restrictive (the challenge scenario) is worth 1.6% of EU GDP per year. This is equivalent to approximately one year of GDP growth for the EU according to the IMF's long-run forecasts.[29]

We can also express these impacts in cumulative terms for the 10 years to 2030. This takes into account IMF projections of economic growth, together with a social

[27] We use GDP measured at factor costs, that is the value of goods and services, plus subsidies minus taxes (a measure sometimes known as GVA).

[29] https://www.imf.org/en/Publications/WEO

discount rate.[30] **This points to losses of €1.3tn in the challenge scenario, and gains of €720bn in the growth scenario.** The difference between the two can be interpreted as the gains accruing to the EU if it and partners were to embark on a modestly liberalising pathway, and avoid slippage into a more restrictive setting globally for cross-border trade.

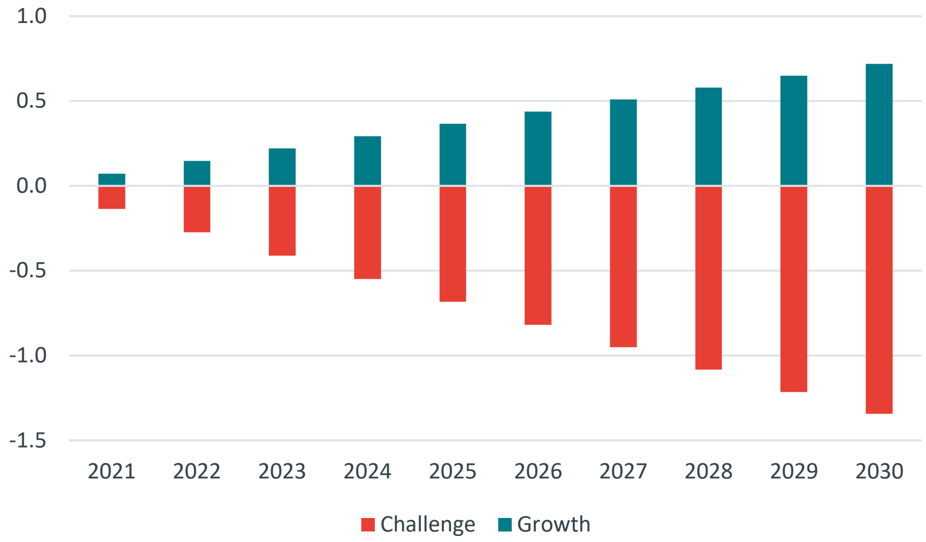
Table 3 GDP impacts

	Annual GDP impact €(m)	GDP impact as % of total GDP	Cumulative impact on GDP over 10 years (€m), NPV
Challenge	-138,806	-1.0%	-1,344,806
Growth	74,210	0.6%	718,969

Source: Frontier Economics analysis of OECD Trade in Value Added (TiVA) Database
Note: Figures expressed in 2020 EUR, and include rescaling from 2015 TiVA year to 2019 baseline year

We can also see how the impacts accumulate over time. This is shown in Figure 10 below. In the challenge scenario, each year the economy is below the baseline level it suffers losses. By 2030, the difference between challenge and growth scenarios is worth around €2tn.

Figure 10 Cumulative GDP impacts by year (€tn, Net Present Value)



Source: Frontier Economics analysis of OECD Trade in Value Added (TiVA) Database
Note: Figures expressed in 2020 EUR, and include rescaling from 2015 TiVA year to 2019 baseline year

4.2.2 Employment impacts

The annual GDP impacts can be converted into indicative employment figures by dividing through by GDP per worker. Over the whole EU, GDP is around €68,000 per worker. Across the data-reliant sectors, we find an average GDP per worker of €73,000. However, the modelling suggests that some sectors are more impacted

[30] In terms of calculation, the percentage trade impacts give a percentage GDP impact. This is then applied to a growth projection with discounting. So in other words GDP is (say) 1% lower per year, which is overlaid on the baseline growth path. The discounting uses a rate of 2.5%. There is a considerable literature on the question of what discount rate to use in appraisal, with a range of values suggested. For example, France uses a rate of 2.5%, whereas the UK uses a rate of 3.5%.

than others and pull down the average for the data-reliant group. For example, wholesale and retail has GDP of €52,000 per worker. Weighting by the sectors' share of trade impacts in the challenge scenario, we find an average GDP per worker of €105,000.

To illustrate, the €139 billion annual impact shown in the challenge scenario corresponds to the output of 1.9 million workers in the data-reliant sectors, i.e. by dividing through by €73,000 per worker.

Table 4 Illustrative employment impacts (million employees)

	€68k per worker	€73k per worker	€105k per worker
Challenge	-2.0	-1.9	-1.3
Growth	1.1	1.0	0.7

Source: Frontier Economics analysis of OECD Trade in Value Added (TiVA) Database and Eurostat NAME_A64_E

It is important to bear in mind that these results are indicative. The GDP impacts are based on high-level relationships observed between productivity and openness to trade. They are not specific about the mechanism of impact, and will include many different channels, such as knock-on effects on other sectors. For example, a reduction in activity by an exporting sector will feed through into impacts on the various sectors that supply it. Most importantly, these results should not be interpreted as net employment impacts, as even workers who directly lose their job as a result of export reduction will in many cases find employment in other sectors. But the exercise does give an indication of the number of highly productive jobs that could be compromised in each scenario.

4.3 Further detail on modelling results

4.3.1 Impacts on trade by trading partner

We can also explore how the EU's export impacts are broken down by partner. Here the countries are ranked by size of impact (as measured by millions of Euros) in the challenge scenario. The countries appearing on this list are either the EU's largest trading partners, or are the more restrictive jurisdictions, which are assumed to undergo sharper changes in the scenarios. **We see the role played by the US, the UK, China and Switzerland.**

Table 5 Impacts by trading partner and scenario

	Annual impact (€m)		Percent of EU exports to partner	
	Challenge	Growth	Challenge	Growth
USA	-20983	6212	-2.4%	0.7%
GBR	-15960	8408	-2.3%	1.2%
CHN	-8326	10562	-1.7%	2.1%
CHE	-8323	4420	-2.6%	1.4%
JPN	-4574	2422	-2.6%	1.4%
RUS	-4375	3328	-2.5%	1.9%
IND	-2970	1322	-2.9%	1.3%
TUR	-2966	1544	-2.0%	1.0%
KOR	-2650	1390	-2.3%	1.2%
BRA	-2557	1733	-2.1%	1.4%
Rest Of the World	-20749	9273	-2.3%	1.0%

Source: Frontier Economics analysis of OECD Trade in Value Added (TiVA) Database

Note: Figures expressed in 2020 EUR, and include rescaling from 2015 TiVA year to 2019 baseline year

We can also break down the partner-level impacts by sector to reveal particularly strong drivers, such as financial services with the UK and Switzerland.

Table 6 Annual export impacts by partner and sector in challenge scenario (€m)

	Construction	Wholesale retail	Transport	Media	Telecoms	Computer services	Financial	Other business	Cultural	Manufacturing
USA	-18	-948	-787	-398	-75	-1320	-1300	-4352	-208	-11576
GBR	-13	-933	-524	-478	-428	-1079	-1855	-2095	-747	-7808
CHN	-9	-589	-339	-194	-39	-300	-158	-590	-66	-6041
CHE	-8	-422	-229	-144	-70	-2082	-1085	-806	-79	-3398
JPN	-3	-235	-219	-56	-25	-973	-291	-757	-16	-1999
RUS	-6	-215	-139	-50	-143	-628	-366	-749	-86	-1993
IND	-3	-132	-174	-25	-19	-332	-437	-522	-13	-1312
TUR	-3	-203	-151	-11	-11	-18	-160	-154	-12	-2243
KOR	-2	-151	-152	-19	-17	-150	-73	-569	-12	-1506
BRA	-3	-176	-78	-37	-16	-386	-85	-855	-33	-887
Other	-26	-1423	-1992	-122	-415	-801	-1721	-2594	-231	-11423

Source: Frontier Economics analysis of OECD Trade in Value Added (TiVA) Database

Note: Figures expressed in 2020 EUR, and include rescaling from 2015 TiVA year to 2019 baseline year

4.3.2 Impacts by size of exporter

Gravity models of trade are not detailed enough to estimate the impact of restrictions specifically on small and medium enterprises (SMEs) or large enterprises. However, the effect of data policies on SMEs are likely to account for a significant proportion of the overall impact of these policies on the EU economy.

In the manufacturing sector, small firms account for 23% of goods exported from the EU.[31] Therefore, a relatively crude apportioning of the results shown in section 4.1 suggests that exports by data-reliant manufacturing SMEs in the EU are worth around €280bn. In the challenge scenario, exports from EU SMEs would fall by €14bn, and in the growth scenario they would increase by €8bn.

This apportioning implicitly assumes that the exports of SMEs and the exports of larger firms are equally responsive to changes in data policies. This assumption may underestimate the effect of data policies on SMEs. This is because data policies often impose fixed costs on businesses, and larger firms are generally better able to absorb changes to fixed costs by virtue of their size.

Assessing the importance of EU SMEs for trade in data-reliant services is more challenging because data on the value of their exports is not available. However, SMEs account for 61% of turnover in EU services. Therefore, even if exports were more skewed towards larger firms, a significant proportion of the export impact would be with smaller firms. Figure 11 below provides shows the split of sector turnover by firm size for all data-reliant services sectors. This should give some indication of the extent to which SMEs might bear the brunt of any trade impacts.

Figure 11 Breakdown of turnover by size band

Sector	SMEs (<250 employees)	Large firms (250+ employees)
Wholesale and retail	63%	37%
Media	55%	45%
Telecommunications	15%	85%
IT	46%	54%
Business services	65%	35%
Total data-reliant sectors modelled	61%	39%

Source: Frontier analysis of Eurostat Structural Business Statistics data

4.4 Assessing the overall role of data flows in facilitating trade

As an extension to our results, in this section we model the effects of a hypothetical prohibition on cross-border data flows. This is useful for two reasons:

- It provides a guide to the economic contribution of cross-border data flows via international trade: we measure a hypothetical counterfactual case in which

[31] This is in relation to manufacturing overall. Comparable statistics are not available at the level of data-reliant manufacturing sectors.

these data flows are not possible and observe how far trade and economic activity fall compared to current settings.

- This hypothetical provides a guide as to the "insurance" value of specific rules on trade and data flows: even if new rules do not lead to further liberalisation (as in the growth scenario) they can provide safeguards against bad outcomes. The hypothetical provides an estimate of the upper bound of losses that can be avoided through rules, even if these are limited to locking in existing levels of liberalisation.

In terms of our modelling framework, this assumes that STRIs for all countries are turned up to their maximum setting from their current levels. Though unlikely, this hypothetical is in principle possible. It provides an estimate of the total economic value at risk stemming from the effects of the prohibitions on trade.

Impacts on trade

In the most sensitive sectors (IT, media, cultural industries), maximum restrictions on data would lead to a fall in exports of around one fourth. **Over all sectors, the effect would be a 9.6% reduction in exports, worth around €270bn in total.** To put this into perspective, this is about the same as the total value of Spain's exports of goods in 2020.[32]

Table 7 Hypothetical case of full restrictiveness: sectoral impacts

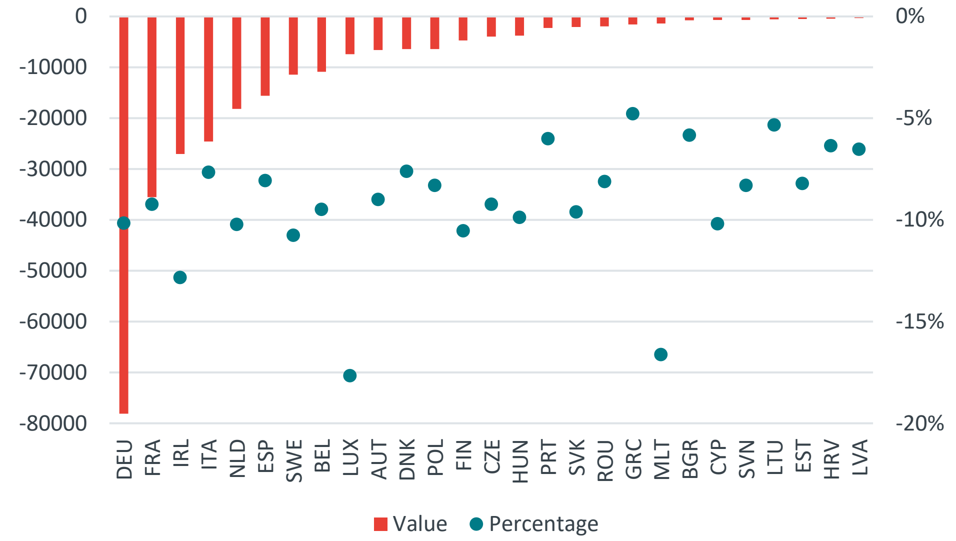
Name	Annual impact (€m)	% of sector exports	% of total EU (ex EU) exports
Construction	-303	-5.9%	0.0%
Wholesale and retail	-16190	-5.7%	-0.6%
Transportation and storage	-15704	-5.9%	-0.5%
Media	-4092	-26.9%	-0.1%
Telecommunications	-4068	-19.7%	-0.1%
IT and other information services	-22570	-28.3%	-0.8%
Financial and insurance activities	-20083	-19.1%	-0.7%
Other business sector services	-40559	-20.3%	-1.4%
Cultural	-3882	-23.7%	-0.1%
Manufactured goods	-147022	-12.2%	-5.1%
Total	**-274473**	**-12.5%**	**-9.6%**

Source: Frontier Economics analysis of OECD Trade in Value Added (TiVA) Database
Note: Figures expressed in 2020 EUR, and include rescaling from 2015 TiVA year to 2019 baseline year

The impacts by EU Member State are shown in Figure 12. These show a similar relativity of impacts in terms of values and percentages.

[32] € 268bn according to Eurostat data on International Trade in Goods.

Figure 12 Trade impacts by EU country in the hypothetical case of full restrictiveness



Source: Frontier Economics analysis of OECD Trade in Value Added (TiVA) Database
Note: Figures expressed in 2020 EUR, and include rescaling from 2015 TiVA year to 2019 baseline year

GDP and employments impacts

Annual GDP losses are around 2.5% with cumulative losses of close to €3.2tn out to 2030. These values provide an estimate of the economic value to the EU of cross-border data flows via the contribution these flows make to international trade. The results for employment effects under different assumptions provide a guide to the social contribution of cross-border data flows.

Table 8 GDP impacts under the hypothetical case of full restrictiveness

	Annual GDP impact €(m)	GDP impact as % of total GDP	Cumulative impact on GDP over 10 years (€m), NPV
Full restrictiveness hypothetical	-329,732	-2.5%	-3,194,563

Source: Frontier Economics analysis of OECD Trade in Value Added (TiVA) Database
Note: Figures expressed in 2020 EUR, and include rescaling from 2015 TiVA year to 2019 baseline year

Table 9 Illustrative employment impacts (million employees) under the hypothetical case of full restrictiveness

	€68k per worker	€73k per worker	€105k per worker
Full restrictiveness hypothetical	-4.8	-4.5	-3.1

Source: Frontier Economics analysis of OECD Trade in Value Added (TiVA) Database and Eurostat NAME_A64_E

4.5 Conclusions

This report has set out to quantify the impacts that policy changes that restrict or liberalise cross-border data flows can have on the EU economy. We have done this by:

- Identifying sectors of the EU economy that are particularly reliant on data;
- Defining possible scenarios for future policy restrictions to cross-border data flows determined by the EU and its major trade partners; and
- Using an econometric model to estimate the impact of these restrictions on EU exports, and consequently on EU GDP.

We find that cross-border data flows make a significant contribution to the EU economy. Because of this, changes in policy at the EU level and internationally in trade partners can have a significant effect. Indeed, both the EU and trading partners are at an important juncture in determining policy conditions around data.

Our research underscores that the difference in economic impact between a modestly liberalising path and a more pessimistic restrictive path is significant: around the equivalent of one year of forecasted long-term EU-wide growth. And that the EU can secure the majority share of these benefits through its own efforts to maintain free flows of cross-border data, regardless of what partners do.

The research also underscores the value to the EU, and partners, of pursuing international arrangements that help to reduce fragmentation in data governance: specifically, the value of agreeing common principles for balancing the benefits of cross-border data flows with other policy objectives.

ANNEX A DATA-RELIANT SECTOR DEFINITIONS

Figure 13 sets out the sectors analysed by the OECD in its taxonomy grading of sectors based on digital intensity (including those not considered data reliant). For the purposes of the analysis in this paper, sectors are considered data-reliant if they are in the top quartile for at least two of the seven criteria or in the top two quartiles for at least four criteria, with the exception of public administration, which is removed because trade accounts for negligible proportion of the sectors contribution to GDP. In our results reported in Section 4, we aggregate some sectors together due to data availability.[33]

[33] For example, Arts and entertainment, and other services are combined into a single 'Cultural Sector' since this is now the data is recorded in the OECD's Trade in Value-Added database.

Figure 13 Sectoral taxonomy of digital intensity, by indicator

	Software investment	ICT intangible investment	Intermediate ICT goods	Intermediate ICT services	Robot use	Revenue online sales	ICT specialists	Top quartile in 2+ criteria	Top 2 quartiles in 4+ criteria
Agriculture								No	No
Mining								No	No
Food & beverages								No	No
Textiles and Apparel								No	No
Wood & paper production								No	Yes
Coke & petroleum								Yes	No
Chemicals								Yes	No
Pharma								Yes	No
Rubber & plastics								No	No
Metal products								No	No
Computers and electronics								Yes	Yes
Electrical equipment								Yes	Yes
Machinery & equipment								Yes	No
Transport equipment								Yes	Yes
Furniture								Yes	Yes
Electricity & gas								No	No
Water & sewerage								No	No
Construction								No	No
Wholesale & Retail								Yes	No
Transport & storage								No	No
Hotels & restaurants								No	No
Media								Yes	Yes
Telecommunications								Yes	Yes
IT								Yes	Yes
Finance								Yes	Yes
Real Estate								No	No
Legal & accounting								Yes	Yes
Scientific R&D								Yes	Yes
Marketing								Yes	Yes
Admin services								Yes	Yes
Public Admin								No	Yes
Education								No	No
Health								No	No
Social work								No	No
Arts & entertainment								No	No
Other services								Yes	Yes

Source: *Adapted from Calvino, F., et al. (2018)*

Note: *All underlying indicators are expressed as sectoral intensities. For each indicator, sectoral values represent averages across countries and years (for the period 2013-15 only). The colour of the cells in the table corresponds to the quartile of the sectoral distribution to which the sector belongs.*

ANNEX B TRADE MODELLING METHODOLOGY

The modelling approach is to represent changes to data localisation as changes to the OECD's services trade restrictiveness index (STRI), and quantify how these feed through into effects on trade. This uses estimates of the responsiveness of trade to the STRI, which are calculated in a gravity model. The trade modelling is done on a bilateral, sector-level basis. For example, we model the how changes to data would affect telecoms trade between Germany and the UK, France and the US, etc. separately, and add these impacts together to give overall trade impacts by scenario.

We explore both services trade and data-reliant goods sectors, which are analysed separately, as the necessary data is different in the two cases. We begin with a discussion of the services trade restrictiveness index and operationalisation of the scenarios, followed by discussion of the gravity modelling for services trade and goods respectively.

Services trade restrictiveness index

The STRI assesses how restrictive a jurisdiction is to foreign services providers, with a value of zero meaning completely open, and a value of one meaning completely closed.

Barriers to services trade are defined in terms of:

- Restrictions to foreign entry;
- Movement of people;
- Discriminatory measures;
- Barriers to competition; and
- Regulatory transparency.

The STRI is calculated using a scorecard approach, containing a long list of restrictions pertaining to the above categories. Each of the restrictions carries a weight, and if in place, the corresponding weight[34] is added to the score. If all of the restrictions were in place, the weights would sum to one and the jurisdiction would be seen to be completely closed.

The STRI incorporates five restrictions ('lines') that describe the stance in relation to cross-border data flows. This sits within the 'restrictions to foreign entry' category. The five restrictions are set out below, with numbering from the OECDs codification:

- *1.20.2 Cross-border transfer of personal data is possible when certain private sector safeguards are in place*[35]
- *1.20.3 Cross-border data flows: cross-border transfer of personal data is possible to countries with substantially similar privacy protection laws*

[34] The weights attached to restrictions reflect the consensus view of sector experts.

[35]A "NO" to this line is considered restrictive, whewreas a "YES" is considered restrictive in the other four lines.

- *1.20.4 Cross-border data flows: cross-border transfer is subject to approval on a case-by-case basis*
- *1.20.5 Cross-border data flows: certain data must be stored locally*
- *1.20.6 Cross-border data flows: transfer of data is prohibited*

Each of the five lines carries the same weight in the STRI. The weights vary by sector, reflecting the relative importance attributed to this category of restriction by the experts consulted in the OECD exercise.

The change in the STRI associated with a change in data localisation policy is given by multiplying the number of STRI lines changed by the weight for that sector. In computer services each line carries a weight of 0.016. This means that moving from having none of the data restrictions in place to having all five in place would increase the STRI by 0.07.

The overall modelling approach is to simulate the impact on trade of turning these restrictions “off” or “on”. For most EU countries, only line 1.20.3 is on, although for a handful 1.20.5 is also on. Most trade is done with countries with one line on, while China, India, and Russia are a lot more restrictive.

Scenarios

As we find, the restrictiveness of both the importer and the exporter in a pair of countries affects the trade between them. The scenarios model the effects of both EU countries and trading partners changing their restrictions, both of which feed through into effects on trade. These result in a percentage change in EU exports, from which changes in GDP are then computed.

Baseline scenario

The baseline scenario keeps EU and partner restrictions at their current level, so there is no change in trade from current levels. From this, further projections of GDP growth are made, in line with IMF forecasts.

The other scenarios are all deviations from the baseline. They follow the same baseline GDP growth trajectory, but are modelled as levels deviations from it, i.e. trade is X% lower and GDP Y% lower per year as a result of the data restriction, which is overlaid on the baseline GDP projection.

Challenge scenario

The “challenge” scenario assumes that the more liberalised countries, including EU, become less liberalised, and add two lines of restrictiveness regarding data flows. For example, France would add two lines, which would add 0.032 (= 0.016 per line x 2) to the STRI. In the case of the more restrictive jurisdictions (China, India, Russia), we assume all five line would be switched on.

Growth scenario

The “growth” scenario assumes that the more liberalised countries add one line of restrictiveness. For example, France would add two lines, which would add 0.032

(= 0.016 per line x 2) to the STRI. In the case of the more restrictive jurisdictions (China, India, Russia), we assume all five line would be on.

Full value-at-risk scenario

This scenario looks at the difference between full and minimal restrictiveness with regard to data flows. It is modelled as the difference between having all five lines with respect to data switched on and off. This gives a percentage reduction in trade, which is applied to current levels. This is intended to illustrate how different trade and GDP would be as a result of such a change.

Services trade gravity modelling

The effect of services trade restrictiveness on trade flows is estimated using a gravity model. This largely follows the approach pursued by the OECD STRI analysis.[36] The gravity model predicts services trade flows as a function of distance, GDP, common language, contiguity, colonial relationship, STRI scores of exporting and importing countries (relating to the sector in question), and whether the trading pair are both EU. The aim is to generate elasticities for the STRI that will estimate how changes to the value of STRI affect bilateral flows between a pair of countries.

Data

The main dataset, including distance, GDP and dyadic variables is from CEPII.[37]

Bilateral services trade data is from OECD EBOPS[38] and is reported for a number of different sectors. The Services Trade Restrictiveness Index is also from the OECD[39] and the dataset used here covers the years 2014-16.[40]

The following sectors have both STRI and trade flow data available, and are incorporated in the model:

- Sea transport;
- Air transport;
- Other modes of transport;
- Postal and courier services;

[36] Nordås and Rouzet The Impact of Services Trade Restrictiveness on Trade Flows, Hildegunn K., The World Economy (2016)

[37] http://www.cepii.fr/cepii/en/bdd_modele/presentation.asp?id=8, dataset originally developed for

HEAD, K., T. MAYER AND J. RIES, 2010, "The erosion of colonial trade linkages after independence" Journal of International Economics, 81(1):1-14.

[38] https://stats.oecd.org/Index.aspx?DataSetCode=TISP The TISP dataset covers modes of supply 1 (cross-border supply), 2 (consumption abroad), and 4 (movement of natural person). It does not capture mode 3 (commercial presence).

[39] http://stats.oecd.org/Index.aspx?DataSetCode=STRI

[40] It should be noted that although the dataset is a technically a panel, there is little variation in STRIs over time, which means there is very little scope to use the panel aspect of the data, such as with fixed effects. Therefore the cross-sectional aspect of the data drives the results, and one-year pure cross-section models generate very similar results.

- Construction;
- Insurance and pension services;
- Financial services;
- Telecommunications services;
- Computer services;
- Legal services;
- Accounting, auditing, bookkeeping, and tax consulting services;
- Architectural services;
- Engineering services; and
- Audio-visual and related services.

Note that for audio-visual, although STRIs are provided separately for broadcasting, motion pictures and sound recording, the trade flow data is only broken down as far audio-visual, with no further disaggregation available. Therefore the gravity modelling includes the audio-visual sector as a whole, with the STRI values averaged across the three sub-sectors.

The range of countries covered is constrained by the availability of STRI data. As well as OECD, the STRIs are calculated for Brazil, China, Colombia, Costa Rica, India, Indonesia, Lithuania, Russia, and South Africa.

It is important to note limitations of the services trade data. In particular there are many gaps in the published TISP data, for example, due to data being redacted for confidentiality reasons, or not split out into detailed sector. In these situations, the gaps in the data can be addressed using "mirror flows". For example, if the UK does not report imports from Germany, we can instead use Germany-reported exports to the UK. However, for whatever reason, the trade flows reported by the exporter are generally larger than the same flows as reported by the importer. Therefore, to "infill" the data we first estimate importer-reported flows using exporter-reported flows, derive the predicted values, and use these where the importer-reported trade flows are missing.

Specification and results

The regression is estimated using a poisson pseudo-maximum likelihood (PPML) approach, following the Nordas-Rouzet paper. The coefficients in a PPML regression give the proportional change in the dependent variable in the same way as in an OLS regression with a logged dependent variable[41] The PPML approach is argued to be better for dealing with missing observations and is described in detail in Silva and Tenreyro.[42]

We predict trade from country I to country J as a function of size of the two countries (log GDP), the STRI scores of the two countries, a series of dyadic variables X (log distance, and dummies for common language, contiguity, colonial relationship, and whether EU pair), year dummies and sector dummies.

[41] In a PPML model the coefficients give a proportional change in the dependent variable. In both cases the percentage change in the dependent variable for a change in variable X is given by exp(βvar*Δvar)-1

[42] The Log of Gravity, Review of Economics and Statistics, 2006. The authors use Monte Carlo simulations to compare the performance of log-linear OLS and PPML estimators.

This can be written as follows:

$$\text{Trade flow}_{ijst} = b_0 + b_1 \log GDP_i + b_2 \log GPD_j + b_3 STRI_i + b_4 STRI_j + b_5 X_{ij} + b_6 yr_t + b_7 sector_s + u_{ijst}$$

The results of the pooled regression are shown in Figure 14 below. The first column includes all 14 sectors for which STRI data are available. The exporter STRI coefficient of -1.53 means that if the score is reduced by 5 percentage points, trade would be increased by 8%.[43] The other coefficients, e.g. on GDP and distance, are comparable to other services trade gravity model estimates. The second column shows results relating to 9 sectors most comparable to the sectors of interest, focusing on communications and professional services (transport, logistics and construction are excluded). The STRI coefficients become somewhat larger. The third column shows results for the physical sectors. Here the coefficients are smaller and the importer STRI is statistically insignificant.

Figure 14 Services trade regression results

	Pooled model	Digital sectors	Distributive sectors
Log distance	-0.591	-0.666	-0.541
	[22.29]**	[14.47]**	[-22.47]***
Log GDP exporter	0.548	0.604	0.494
	[39.78]**	[26.72]**	[32.69]***
Log GDP importer	0.612	0.61	0.618
	[33.61]**	[18.79]**	[37.14]***
Contiguity dummy	-0.022	-0.293	0.220
	[0.37]	[3.05]**	[3.34]***
Common language dummy	0.544	0.781	0.287
	[8.16]**	[7.53]**	[3.81]***
Colonial dummy	0.417	0.437	0.357
	[5.27]**	[3.90]**	[3.32]***
STRI exporter	-1.535	-2.441	-0.626
	[6.51]**	[5.49]**	[-2.59]***
STRI importer	-0.923	-1.718	-0.249
	[5.22]**	[6.12]**	[-1.1]0
EU pair	0.133	0.157	0.102
	[2.54]*	[1.71]	[1.94]*
Constant	-21.9	-27	-21.317
	[33.0]**	[23.9]**	[-29.1]***
R2	0.28	0.24	0.38
N	39232	24977	14255

Source: Frontier analysis of OECD and CEPII data

*Note: T-statistics in parentheses, significance levels:***p<0.01 ** p<0.05; * p<0.01*

Choice of control variables

Given that the sample is relatively small, care is needed in terms of the control variables that can be included together in the model. They may be highly correlated with each other, which may cause some of them to take a counterintuitive sign. For example. We therefore undertake a model selection approach to retain a limited number of variables, and which do not give counterintuitive results. For example, while we find that an EU dummy has a moderate positive impact, including a

[43] Using the marginal effects formula above, this is given by Exp(-1.53*-.05) -1 = 8%.

generic FTA dummy causes the EU dummy to get larger, but the FTA dummy has a negative sign, while the STRI variables stay the same. Meanwhile the FTA dummy can change sign depending on the inclusion of other variables. We therefore exclude the FTA dummy, as there is essentially not enough variation in this data to pick apart the relative effects of different types of FTA alongside the effects of the STRI.

We also explored supplementing the model with intra-EEA STRIs. In theory, these should have the benefit of offering more precise bilateral measures of restrictiveness, whereas the standard STRI is reported on an MFN basis. They could therefore be included to explore the effects bilateral deviation from the MFN level. However, we again find that these variables take a positive sign, i.e. so that having less restrictiveness vis-à-vis the partner trading relative to MFN levels results in lower trade than would otherwise be the case. This result is counterintuitive, and we suspect it arises because there is insufficient variation IN the data to be able to reliably pick apart these effects.

Pooled regressions

The limited variation in the data gives a strong justification for using a pooled regression, as cutting the data too is seen to give overfitted models with counterintuitive results. In a pooled regression, the intercept (sector) dummies control for the average sizes of sectors (e.g. telecoms is larger than audio-visual). Meanwhile, the coefficients estimated for the STRI, the dyadic variables, and GDP are constrained to be the same across sectors. This exploits the maximum amount of information in estimating an average STRI effect.

The alternative to a pooled approach is to run separate regressions for each sector in turn, which is explored in Nordas and Rouzet This gives a much larger range of STRI coefficients, with some becoming very large and others taking on a counterintuitive sign.[44] This amount of variation in the parameters is implausible and not consistent with prior empirical evidence on their effects. While there may be some genuine sectoral variation in responses to the variables (e.g. sea freight is less responsive to distance than terrestrial transport), it is not obvious, for example, why architecture should be much less responsive to distance than engineering or legal services are. This suggests that the sector-level regressions are overfitted and that the pooled results are to be preferred. On this basis, we consider that the pooled sub-sample results are reasonable, as they use the maximum amount of variation available in the data by drawing on trade relationships for similar sectors, and are less prone to influence from quirks in the data.[45] For similar reasons, we seek to run regressions on the maximum sample of countries rather than an subsets of them, which again causes the coefficeinst to become less stable.

[44] For example, banking has an elasticity of -13 while road has a positive sign.

[45] For a broader discussion of pooling, refer to the following papers. Baltagi, B.H. & J.M. Gri¢ n (1997) Pooled Estimators versus their Heterogeneous Counterparts in the context of dynamic demand for gasoline, Journal of Econometrics, 77, 303-327.

Baltagi, B.H,.J.M. Gri¢ n & W. Xiong (2000) To pool or not to pool: homogeneous versus heterogeneous estimators applied to cigarette demand, Review of Economics and Statistics, 82, 117-126

Goods trade gravity modelling

We hypothesise that a selection of goods manufacturing sectors are particularly sensitive to services inputs, of which data flows are a component. We therefore wish to estimate a model in which trade flows in these goods are the dependent variable.

We extend the analysis by estimating a gravity model of goods in which we include measures of services trade restrictiveness to allow us to derive the impacts of data localisation measures, proxied for by changes to the STRI. In this respect it follows the services trade analysis. There are two key differences: first, there may be additional drivers that affect goods trade and not services; second, sector-specific STRIs cannot be linked to trade flows at the sector level in the same way as can be done for services. A further complication is that some of the relevant variables for the dataset are captured at different points in time and it is difficult compiling a genuinely contemporaneous dataset.

Data

The goods trade data is extracted from UN Comtrade using product codes that correspond most closely to high-value manufacturing.[46] We extracted data for 2019. In addition to the variables used elsewhere, it is considered appropriate to control for conventional barriers to goods trade. The main indicator here is the Overall Trade Restrictiveness Index estimated by Kee et al,[47] which measures the effect of tariff and non-tariff barriers to goods trade. The problem is this data was last estimated in 2012, and more recent indices for goods trade are not available. As before, the dataset uses other controls from CEPII.

Note that as we use the STRI as control variable, this limits the number of countries that can be included in the regression, which is smaller than might often be used in goods analysis.

Analysis

Again, we undertake a parsimonious approach to variable selection, looking to retain variables that have a statistically significant and sensible interpretation. However, the overarching problem with the dataset is that it either omits, or poorly measures, the effects of tariff and non-tariff barriers to trade.

A further point to note is that China acts as a distortive outlier in the sample. It is a restrictive jurisdiction, but has much higher than expected levels of goods and manufacturing trade than is predicted by the model. When included, many of the restrictiveness variables take on the “wrong” sign. On this basis, China is excluded from the sample.

[46] The following HS codes are used: 28 inorganic chemicals; 29 organic chemicals; 30 pharmaceuticals; 37 photographic / cinematographic; 38 chemicals n.e.c.; 85 electrical machinery / equipment; 86 railway locomotives;87 vehicles; 88 aircraft;90 optical / medical instruments;91 clocks / watches;92 musical instruments; 93 arms and ammunition.

[47] https://www.researchgate.net/publication/23646050_Estimating_Trade_Restrictiveness_Indices

Results

The first specification shows that both importer and exporter STRIs have a negative effect on trade in high-value manufactured goods. However, the exporter term is statistically insignificant. The second specification adds in the OTRI term. This results in slightly larger STRI coefficients. However, the OTRI term itself has the wrong sign, which raises concerns as to how well it is measuring goods restrictions, which is unsurprising given it is potentially out-of-date.

Figure 15 Gravity model with high-value manufactured goods

Variable	Specification 1		Specification 2	
	Beta	*T-statistic*	*Beta*	*T-statistic*
STRI exporter	-1.074	[0.886]	-1.348	[0.767]*
STRI importer	-0.946	[0.476]**	-0.965	[0.501]*
FTA dummy	0.154	[0.149]	0.154	[0.151]
EU dummy	0.414	[0.153]***	0.417	[0.158]***
Log distance	-0.367	[0.097]***	-0.383	[0.088]
Log GDP exporter	0.804	[0.035]***	0.802	[0.036]
Log GDP importer	0.905	[0.066]***	0.902	[0.062]***
Common language dummy	0.425	[0.148]***	0.423	[0.147]***
Time difference	-0.080	[0.041]**	-0.076	[0.038]**
Colonial dummy	-0.475	[0.186]**	-0.465	[0.187]**
Importer OTRI			0.19	[1.028]
Exporter OTRI			1.128	[1.927]
Constant	-10.633	[2.385]***	-10.397	[2.225]***
R-squared	0.61		0.61	
N	1979		1979	

Source: Frontier analysis of OECD, Comtrade and CEPII data

*Note: T-statistics in parentheses, significance levels:***p<0.01 ** p<0.05; * p<0.01* www.frontier-economics.com