Data Sovereignty: Firms, Accountability, and the Challenge of International Governance*

Carlos Felipe Balcázar[†], Martina F. Ferracane[‡] and Stephen Weymouth[§]

June 9, 2026

Abstract

International organizations have struggled to establish binding rules on cross-border data flows despite the growing importance of trade in digitally delivered services (DDS). This paper argues that cooperation is difficult because governments pursue *data sovereignty* through different instruments for asserting authority over data. We distinguish strict localization measures, which keep data, processing, or infrastructure under domestic jurisdiction, from conditional flow regimes, which permit transfers when firms or destination jurisdictions satisfy domestic legal requirements. We develop a theory in which DDS competitiveness raises the costs of territorial control, while political accountability shapes whether firm opposition and consumer privacy concerns influence policy. Using new data on cross-border data restrictions for more than 150 countries from 2005 to 2024, we show that data sovereignty measures differ from conventional goods trade protection. Leveraging exogenous variation in topographic constraints on DDS competitiveness, we find that revealed comparative advantage in DDS is associated with fewer strict localization measures, especially in higher-accountability systems. IOs struggle to govern data flows because domestic political environments lead governments to assert authority through different forms of data sovereignty.

*Tomás Rogaler and Simón González Ugarte provided excellent research assistance. The comments of Robert Basedow and participants at DIESL workshop at Oxford University and KCL's junior IPE scholar workshop, are especially acknowledged.

†Department of Political Science and School of Public Policy, University College London. 29-31 Tavistock Square London, WC1H 9QU, UK. E-mail:c.balcazar@ucl.ac.uk.

‡Robert Schuman Centre for Advanced Studies, European University Institute; Via Giovanni Boccaccio 121, 50133, Florence, Italy. E-mail: martina.ferracane@eui.eu.

§McDonough School of Business, Hariri Building 553, Washington, DC 20057. E-mail: stephen.weymouth@georgetown.edu.

1 Introduction

International organizations have struggled to produce binding rules on cross-border data flows. WTO members launched exploratory work on e-commerce in 2017 and negotiations in 2019, but the text announced in 2024 excluded core disciplines on cross-border data flows and data localization (Whittle, 2024). This failure is puzzling because cross-border data flows support one of the fastest growing areas of international trade: digitally delivered services (DDS). WTO estimates show that global DDS exports more than tripled after 2005, reached US$3.82 trillion in 2022, and accounted for 54 percent of global services exports.[1] DDS depend on the ability of firms to transmit information and deliver cloud-based services across borders (Azmeh, Foster and Echavarri, 2020; Weymouth, 2023), thus good governance of data flows matters for international cooperation in the digital space.

We argue that cooperation on data flows is difficult because governments are creating different forms of control over data, in response to domestic political institutions and incentives. Cross-border data mobility generates economic value, but it also moves sensitive information through platforms and infrastructures that states do not control (Chander and Sun, 2022; Farrell and Newman, 2019*b*; Weymouth, 2025; Hulvey and Simmons, 2025; Oppenheimer, 2025). Governments therefore face pressure from consumers to bring locally generated data under public authority even when data mobility supports DDS trade (Chander and Sun, 2023; Beaumier and Gjesvik, 2025). Meanwhile, governments in less accountable political systems face fewer political costs from strict data localization and can use these measures to bring politically sensitive data under domestic control (Chander and Lê, 2015; Roberts, 2018; Liu, 2021). We use the term *data sovereignty* to refer to policies that expand state authority over locally generated data by restricting its outbound movement, or by conditioning transfer on certain legal requirements and protections. This paper seeks to explain the form that this authority takes.

Governments asserting data sovereignty choose between two main instruments: i) strict localization measures require data, processing activity, or infrastructure to remain within local jurisdiction, and ii) conditional flow regimes permit cross-border transfers when firms satisfy certain legal requirements or when destination countries provide adequate protections (Ferracane, 2017; Ferracane et al., 2026; Mattoo and Meltzer, 2018). Both instruments expand public authority over data, but they serve different domestic purposes. Strict localization measures increase state leverage over firms and data infrastructures (Chander and Lê, 2015; Farrell and Newman, 2019*b*; Polyakova and Meserole, 2019). Conditional flow regimes preserve data mobility while extending legal accountability across borders (Bradford, 2023; Schwartz, 2019). Other governments assert little data

[1] https://www.wto.org/english/res_e/booksp_e/trade_in_serv_devpt_chp1_e.pdf

sovereignty and instead seek open data flows at home and abroad. International institutions must therefore reconcile governments that regulate data flows for different purposes.

The growth of data sovereignty measures makes this problem increasingly consequential. Restrictions on cross-border data flows have increased sharply since 2010, reaching more than 330 policies by December 2024 (Ferracane et al., 2026). Recent work demonstrates that these measures can slow commerce (Jia, Jin and Wagman, 2021; Goldfarb and Tucker, 2019; Zhang, Kim and Moon, 2025). Yet the policy response has remained fragmented. International commitments on data flows have developed primarily through preferential trade agreements rather than binding multilateral rules (World Trade Organization, 2025*b*; Burri and Polanco, 2020). The central puzzle is therefore why governments impose restrictions to data flows through different instruments despite the economic gains from data mobility.

Our explanation centers on the domestic politics of data sovereignty. Firms that export DDS face high costs from data flows restrictions because these policies disrupt the cross-border data flows on which digital service delivery depends. These firms, therefore, have incentives to oppose territorial control and support rules that keep data mobile. Firms pursue these objectives through lobbying and public position-taking over trade policy (Kim, 2017; Osgood, 2017, 2021; Kim et al., 2019*b*). Where DDS competitiveness is stronger, firms press governments to avoid strict localization measures that disrupt data mobility.

Political institutions influence their preferences into policy. The relevant institutional dimension is political accountability: the extent to which citizens and firms can participate in politics in ways that make public demands visible, and costly for governments to ignore. In lower-accountability regimes, incumbents face less public contestation over data policy and have greater room to pursue territorial control, especially when control over information and data infrastructure strengthens state control over firms and politically sensitive data (King, Pan and Roberts, 2013; Lorentzen, 2014; Roberts, 2018; Chen and Xu, 2017; Chander and Lê, 2015; Liu, 2021). In higher-accountability regimes, firms are better positioned to challenge costly restrictions, while consumer privacy concerns are more likely to enter policy through public pressure and organized advocacy. These pressures make territorial control more politically costly and increase demand for enforceable rules governing data transfers.

DDS competitiveness and accountability jointly shape the form of data sovereignty that governments pursue. Competitive DDS sectors raise the political cost of territorial control; accountable institutions make both firm costs and consumer privacy concerns harder for governments to ignore. Conditional flow regimes tend to emerge from this political setting because they preserve cross-border data flows while extending domestic privacy protections to data transferred abroad.

Our framework yields three expectations. First, countries with stronger revealed comparative advantage in DDS should adopt fewer data flows restrictions. Second, political institutions should shape the form of data sovereignty: lower-accountability regimes should rely more on strict localization, while higher-accountability regimes should rely more on conditional flow regimes. Third, accountability should condition the effect of DDS competitiveness. In higher-accountability regimes, stronger DDS competitiveness should be associated with fewer strict localization measures and with more conditional flow regimes—but only if data privacy concerns dominate firms' opposition to the regulation.

We test these expectations using new data on regulatory barriers to cross-border data flows for more than 150 countries from 2005 to 2024. We distinguish strict localization measures from conditional flow regimes rather than treating all data-flow policies as equivalent restrictions. We first show that data sovereignty measures do not follow the same patterns as goods-trade-protection, which motivates our assertion that a different political explanation is needed.

We estimate the relationship between revealed comparative advantage in DDS and the choice of data sovereignty instruments, using land ruggedness as a quasi-exogenous moderator of DDS competitiveness. Difficult geography reduces internet adoption and constrains the development of DDS, allowing us to identify how variation in this exogenous component of DDS competitiveness conditions data policy. To do this we develop an innovative method to estimate differential treatment effects using an exogenous moderator, which we call *instrumental moderation*. We provide formal proof of the sufficient conditions under which this differential treatment effect is statistically identified. We also carry out specification tests to evaluate the identification assumptions of the design.

The results show that DDS competitiveness and political accountability shape the form of data sovereignty. Countries with stronger DDS competitiveness adopt fewer strict localization measures vis-á-vis lower-accountability regimes. In contrast, we do not observe stronger DDS competitiveness being associated to a higher likelihood of conditional flow regimes. This occurs because the mechanisms of accountability and comparative advantage offset each other to some extent in regards to these policies: while data privacy concerns should increase when countries are better at producing data flows, especially in high accountability systems, firms' incentives to oppose them also increase opposing this effect. DDS competitiveness pushes policy away from territorial control in accountable systems, where firms can more effectively shape policy.

Our estimates indicate that in high accountability regimes, an increase in one standard deviation in the exogenous component of the comparative advantage in DDS, translates into a drop in five percentage points in data flows restrictiveness. This estimated differential treatment effect of the comparative advantage in DDS is robust to sensitivity analysis for unobservable confounding and is not driven by any single country in jackknife tests.

The argument and findings contribute to debates on the global governance of new technologies by explaining how digital interdependence reshapes the domestic politics of economic openness. Existing work shows that open data flows support trade and innovation (Meltzer, 2015; Ferracane and van der Marel, 2021; Ferracane, González Ugarte and van der Marel, 2025). We explain why governments with strong economic stakes in data mobility may still adopt data sovereignty policies. Digitally competitive economies have incentives to avoid territorial constraints, but they may still adopt conditional flow regimes that preserve mobility while subjecting cross-border transfers to enforceable legal conditions when data privacy concerns are strong.

The paper also connects trade politics to debates on digital governance. Existing work demonstrates how internationally oriented firms shape economic policy (Baccini, Pinto and Weymouth, 2017; Kim, 2017; Kim et al., 2019*a*; Osgood, 2018, 2021; Jensen, Quinn and Weymouth, 2015). Research on digital interdependence shows how cross-border networks create new problems of state authority (Farrell and Newman, 2019*b*; Hulvey and Simmons, 2025; Oppenheimer, 2025; Weymouth, 2025). We bring these literatures together by showing how DDS competitiveness and political accountability shape data sovereignty. Data carries a politics of its own because the same flows that firms monetize also carry the information through which states govern and citizens hold rights. International cooperation is difficult because domestic political pressures vary across countries, leaving governments to pursue divergent instruments of data governance.

Empirically, the approach we develop allows us to estimate differential effects in a context where it is difficult to find instrumental variables and event studies to leverage exogenous variation using the standard tools of causal inference. It does so by providing an alternative strategy that can leverage the exogenous variation in the treatment as a function of an exogenous moderator. The advantage of this approach is that it relaxes the assumptions in instrumental variables design, using an instrumental moderator instead of an instrumental variable for statistical identification of the effect of the treatment. Since this approach identifies the effect on the full sample, it also circumvents concerns associated with the lack of power that can arise when estimating effects on compliers, using instrumental variables, when using country-level data (Arel-Bundock et al., 2026).

The paper proceeds as follows. The next two sections provide the conceptual background for the theory. We first introduce the dataset used for the analysis, which covers cross-border data flows restrictions. We document how data flow restrictions differ from traditional goods restrictions. We also distinguish strict localization measures from conditional flow regimes, since the argument depends on treating these instruments as politically distinct. We then develop the theory to explain variation in the form of data sovereignty. The empirical sections estimate the effect of DDS competitiveness on strict localization and conditional measures and examine how political accountability conditions this relationship. The discussion returns to the implications for international cooperation and IOs governance of emerging technologies.

2 Data flows restrictions reduce digitally delivered services

Data flows and DDS are intrinsically linked. Data serves as the means through which digital trade occurs, a means of production, and an asset that can be traded itself. It is also the core of new "information industries" such as cloud computing, big data analytics, and artificial intelligence, while contributing to nearly all sectors as an input into R&D, product design, production processes, logistics, marketing, sales, and customer engagement (Ker and Mazzini, 2020).

Several databases compile regulatory information that may be relevant to the analysis of data flows regulation in the context of DDS, even though they do not specifically target this area.[2] However, these databases have a limited geographical and temporal coverage, and data is not readily structured to be used in an empirical setting, which makes it difficult to carry out meaningful analyses.

The Digital Trade Integration (DTI) database, however, covers regulatory restrictions and enabling policies related to digital trade across 162 countries, including regulations applied to data governance (Ferracane, González Ugarte and Rogaler, 2026). Pillar 6 of the database is dedicated to the regulation of cross-border data flows, and the DTI project team released an extension of this pillar that provides a timeseries from 2005 to 2024 for 155 countries, which we use as a basis for our analysis.

This pillar contains five indicators (see Table 1): four relate to restrictions to data flows (Indicators 6.1-6.4) and one to enabling policies for data flows (Indicator 6.5). Bans to data transfers and local processing requirements (indicator 6.1) and infrastructure requirements (indicator 6.3) are considered to the most restrictive type of restriction on data flows because companies must conduct the main data processing activities in the country or even use a specific facility for data processing. Therefore, these policies are treated as strict localization measures.

In contrast, conditional flow regimes (indicator 6.4) forbid the transfer of the data abroad unless certain conditions are fulfilled—conditions that may apply to the recipient country or to the

[2]Among the most relevant ones: the Digital Policy Alert by the St. Gallen Endowment for Prosperity through Trade, OneTrust Data Guidance database, and the OECD's Digital Services Trade Restrictiveness Index.

company (Ferracane, 2024). Many of these resemble the European Union's General Data Protection Regulation (GDPR), which governs cross-border data flows, and which has exerted significant influence across the globe—becoming a blueprint for regulation elsewhere (Bradford, 2020, 2023).

Local storage requirements (indicator 6.2) are the least restrictive only requiring the firm to keep a copy of certain data within the country; as long as the copy of the data remains within the national territory, the firm can operate as usual. These requirements are generally associated to tax and accounting records.[3]

Taken together, indicators 6.1-6.3 define strict localization measures because they require firms to keep data, processing activity, or infrastructure in the country. This distinction becomes relevant later on when we present our theory.

Indicator 6.5 does not contain domestic regulation that applies on an Most-Favored-Nation (MFN) basis but rather bilateral and regional commitments enshrined in trade agreements. As such, while these commitments can also inform our analysis, they are preferential instruments shaped by negotiation dynamics rather than unilateral policy choices. Therefore, we exclude this indicator from our analysis.

[3] A notable exception in Kenya's data protection law, which, among other conditions, requires to keep a copy of all personal data in the country. See Section 50 of the Data Protection Act and Regulation 26.2 of the Data Protection (General) Regulations.

Table 1: DTI pillar 6 indicators

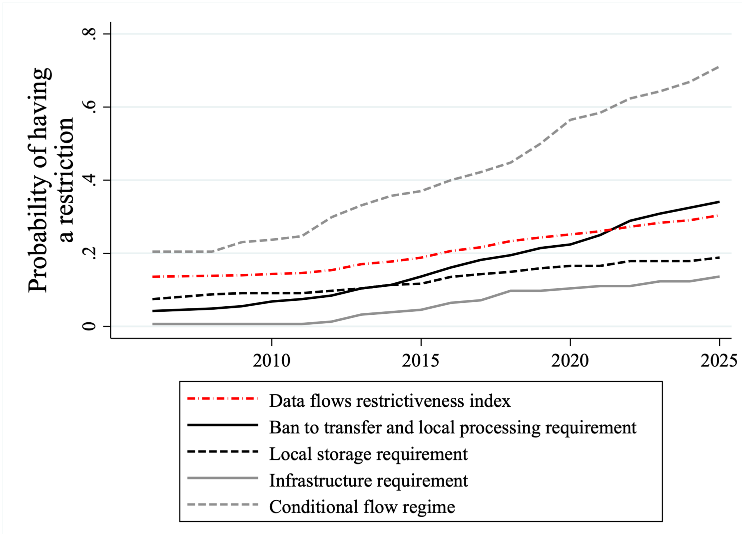
Indicator	Definition	Included in analysis
6.1. Ban on transfer and local processing requirement.	The indicator covers policies that impose a ban on data transfers abroad or require local processing. In the case of a ban on transfers, the company is not allowed to even send a copy of its data abroad.	Yes.
6.2. Local storage requirement.	The indicator covers measures requiring companies to retain copies of certain data within the country. In such cases, as long as a copy of the data is saved domestically, data storage and processing activities can also take place outside the country.	Yes.
6.3. Infrastructure requirement.	The indicator lists cases in which the company is obliged to set up a local data processing facility or use a specific data processing facility in order to provide certain services in the country.	Yes.
6.4. Conditional flow regime.	The indicator covers measures requiring that data can be transferred abroad only if certain conditions are fulfilled. When these conditions are not met, the data cannot be transferred and must be processed locally. The conditions can apply either to the recipient country (e.g., some jurisdictions require that data be transferred only to countries with an "adequate" level of protection) or to the company (e.g., the need to obtain the consent of data subjects for the transfer abroad of their data).	Yes.
6.5. Participation in trade agreements committing to open cross-border data flow.	The indicator highlights whether the country has joined any agreement with binding commitments to open transfers of data across borders. The commitments can be part of free trade agreements or other treaties and regional regulations.	No.

Source: Ferracane, González Ugarte and Rogaler (2026).

To start exploring these barriers to data flows, we construct an index of data flows restrictiveness using multiple correspondence analysis to produce data-driven weights (Table A1); more restrictive measures have higher weights. We re-normalize the index to lie between zero and one, such that zero is the lowest level of restrictiveness and one is the highest. Unfortunately, time series cross-section data on the number of policies by indicator is not available to investigate more deeply the intensive margin of these restrictions.

Figure 1, panel a, shows that barriers to data flows have increased significantly over time. Since 2010, local processing requirements have increased more than ten-fold, conditional flow regimes have increased more than nine-fold, and local storage requirements have doubled (Ferracane, González Ugarte and Rogaler, 2026). Moreover, while there was only one infrastructure requirement implemented in 2010, the DTI database now counts 29 of these obligations.

Figure 1: Likelihood of having a restriction to data flows (2005-2024)



Note: The data flows restrictiveness index is obtained from applying MCA and re-normalizing it so that it lies between zero (lowest level of restrictiveness) and one (highest level of restrictiveness).

Restrictions on data flows impose a direct cost on the import of DDS by companies and citizens residing in the country imposing the requirement. These higher costs derive from the need for foreign companies to comply with additional conditions to process certain data outside the country that imposes the restriction, or to switch to local providers of data processing services (see, e.g., Ferracane 2021; Rentzhog 2015; Rentzhog and Jonströmer 2014). Moreover, restrictions on data flows also raise the costs of using foreign data processing services as inputs for other digital services, thereby increasing the costs for local companies and potentially restricting the export of DDS. Local providers of data processing services, however, might benefit from the restrictions as they might increase demand for services that process data in the implementing jurisdiction.

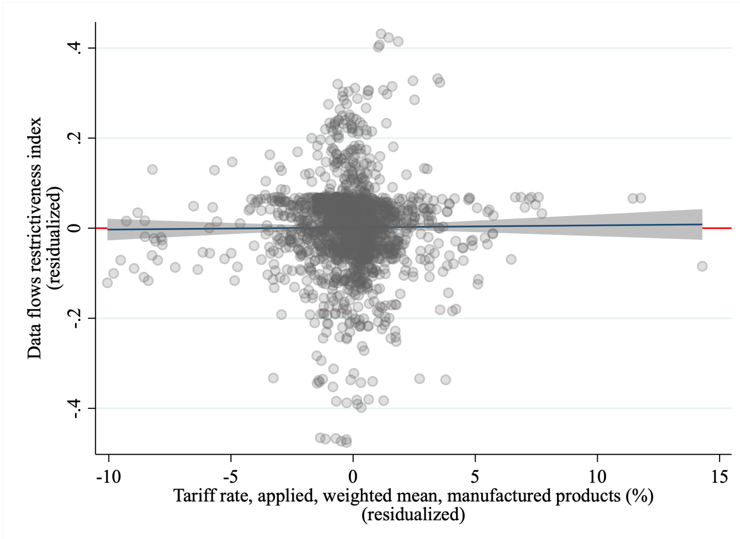
DDS encompass the World Trade Organization's supply mode 1, which are "all international trade transactions that are delivered remotely over computer networks" (OECD, 2023). They encompass all institutional sectors of the economy, and cover deliveries made over the internet (including via mobile devices) and via private networks (e.g., via an extranet). They reflect all electronic transmissions of data, including but not limited to cloud computing, data processing, software, apps, business services and financial services. These barriers emerge as governments seek to regulate data flows to protect citizens' and companies' information, which is both economically valuable and can create cybersecurity risks (Chander and Sun, 2022).

3 Stylized Facts and Motivating Trends

Because our focus are data flows, we do not expect barriers to these to be related to other, more traditional forms of protectionism, such as in manufactured goods, which reflect industrial protectionism. We find that higher levels of protectionism in goods do not correlate with barriers to data flows: Figure 2 shows that the correlation between these two dimensions is statistically zero. This indicates that traditional protectionism cannot account for protectionism in data flows.

Policies on data flows are often linked to the need to protect individuals' personal data and national security (e.g., Chander and Sun 2022; Gilbert et al. 2026), despite their expected negative effect on trade. We posit that countries that are more efficient vis-á-vis other countries in producing data flows should have lower barriers, as they exhibit a comparative advantage in their production. As a percentage of total trade, DDS have increased over time, from 5% to 13%, or about 4.6 trillion USD (from 3.3 to 7.9 trillion) in about a decade—and they are expected to continue growing. In fact, DDS are the fastest growing component of international trade, growing faster that goods trade, and trade in other services, as the Internet enables the provision of services across borders (OECD, 2025; World Trade Organization, 2025*a*). These flows are the core of cloud computing, big data analytics, Artificial Intelligence (AI), and one of the main inputs in the digital economy (Ker and Mazzini, 2020). Barriers to them could reduce the growth of the digital economy and increase the associated transaction costs (Balcázar and Pissarides, 2025). Hence we expect that higher productivity to be associated with lower barriers to data flows as firms oppose these barriers because they increase transaction costs.

Figure 2: Correlation between tariffs in manufacturing and restrictions to data flows (2005-2024)



Note: Values are computed using data gleaned from the WTO. We residualize both the x- and y-axis by controlling for country and year fixed effects. Variables in the x-axis are lagged one year.

3.1 Comparative Advantage and Restrictiveness

The theory of comparative advantage indicates that a country's ability to produce a good or service at a lower relative opportunity cost vis-á-vis other countries, means that the country will export the good. As such, countries with a comparative advantage in the production of services associated with data flows should therefore exhibit lower barriers to them. To start evaluating this economic logic, we start by measuring the revealed comparative advantage (comparative advantage from hereon) of DDS, or what is the same the relative advantage or disadvantage of a certain country in DDS services as evidenced by their trade flows.[4]

The first thing we note is that there is a low correlation between the comparative advantage in manufacturing and the comparative advantage in DDS (Figure 3, panel a). This indicates that having a revealed comparative advantage in manufacturing does not correlate with the revealed comparative advantage in DDS. This indicates that countries that produce more efficiently manufactures than other countries, are no necessarily producing DDS more efficiently; similarly for

[4]The revealed comparative advantage of country in DDS is given by:

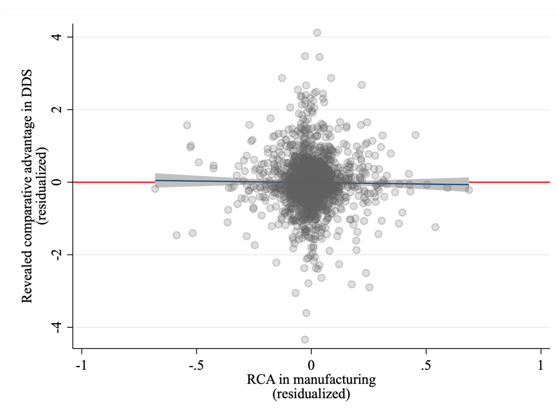
$$RCA_p = \frac{E_{cp}/\sum_{p \in P} E_{cp'}}{\sum_{c' \in C} E_{c'p}/\sum_{c',p' \in P} E_{c'p'}},$$

where E denotes DDS exports, c is the country index, p is the service-commodity index; C and P are the set of countries and commodities respectively. An RCA_p greater than one indicates that the country has a comparative advantage.

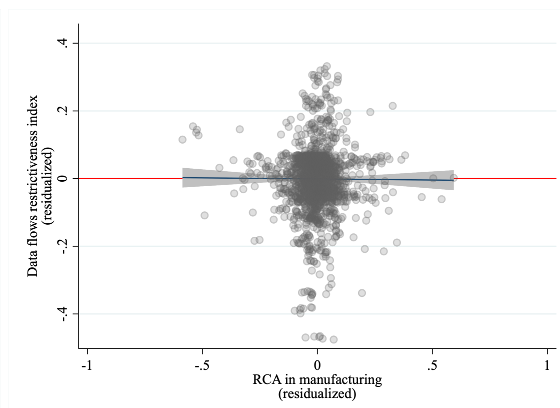
other goods (Figure A1). Second, there is no correlation between the comparative advantage in manufacturing and restrictions to cross border data flows (Figure 3, panel b). In contrast, the comparative advantage in DDS is related to less restrictiveness to data flows (Figure 3, panel c). Trade-related industrial characteristics therefore do not map one-to-one into trade-related characteristics in the digital sphere.

Figure 3: Restrictions to data flows, comparative advantage in manufacturing and in DDS (2005-2024)

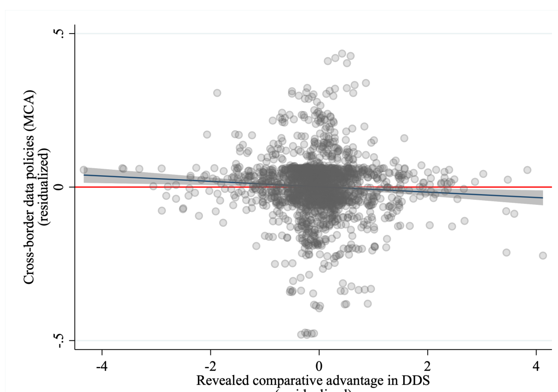
(a) Correlation between comparative advantage in manufacturing and in DDS (2005-2024)



(b) Correlation between comparative advantage in manufacturing and restrictions to data flows (2005-2024)



(c) Correlation between comparative advantage in DDS and restrictions to data flows (2005-2024)



Note: Both variables in the y- and x-axis are residualized by controlling for both country and time fixed effects. Variables in the x-axis are lagged one year.

3.2 Firms and Restrictiveness

To further substantiate the claim that restrictions to cross-border data flows restrictions increase transaction costs, we use data from the World Bank Enterprise Surveys (2022-2025). These are nationally representative firm-level surveys with top managers and owners of businesses in over 150 economies reliance on the internet for sales and purchases, and the costs of these transactions. They provide information at the firm-level on topics such as e-payments and their transaction costs.[5] For firms whose products and business models depend on cross-border data transfers, these barriers should raise their transaction costs.

To identify those firms that rely most on data flows, as this information is not surveyed, we use data by Ferracane and van Der Marel (2025) to identify data-intensive firms, using input-use tables on both capitalised and non-capitalised expenditures in ICTs and computer software using ICT census data for 2010. Cross-border data flow restrictions should affect firms in data-intensive sectors because their operations and service delivery depend more heavily on data flows and processing. These firms should then experience high transaction costs as a result of data barriers; relatedly, firms in countries with higher comparative advantage in DDS should experience lower transaction costs as well.

Intensity is given by:

$$I_j = \ln \frac{\sum_d c_{jd}}{LAB_j}$$

where c_{jd} denotes the software usage for from a set of data producing sectors d, for each downstream services sector j defined using 6-digit NAICS; LAB_j is the amount of labor employed in each downstream sector j. This expression of intensities is standard, and it has been used to define services linkages; in our case, the measure defines a data linkage index (Arnold et al., 2016). We aggregate this measure to 2-digit NACE codes to map the measure to World Bank Survey data.[6] While the measure is imperfect, there is no data on the extent to which data is used by sectors.

Our measure of data intensity is commensurate with our purpose as the transmission of data within and across borders over the internet is performed using software technologies, moreover technology advanced transmissions of data over the internet are done with the help of cloud computing technologies, which in themselves are software-intensive. Table A2 in the appendix shows that the telecom and computer service sectors are the most data-intensive, followed by information

[5] While these surveys also collect information on internet provision, and opinions on obstacles to international trade, such as customs and regulations, and the institutional environment, these don't measure barriers to data flows, whereas e-payments and their transaction costs are directly impacted by these.

[6] Higher data intensity companies show lower transaction costs for digital transactions, are more likely to agree that customs are obstacles, and to state that the workforce is unqualified (Table D2).

services such as data processing services and web search; both financial and insurance services are also data-intensive sectors. The least data-intensive sectors are construction and travel services. In the middle, there is a mix of modern and traditional sectors such as R&D and transport services.

We observe that firms in data-intensive sectors, in countries with high comparative advantage in DDS, report using more e-payments to some extent to both sell and purchase goods and services (Figure 4 panel a). They also report lower transaction costs for e-payments made (Figure 4 panel b). Congruently, firms facing lower cross-border barriers to data flows report using more e-payments to both sell and purchase goods and services (Figure 4 panel c), they also report higher transaction costs for e-payments received (Figure 4 panel c); firms that are more data intensive report the highest transaction costs. Taken together, this indicates that establishing restrictions to data flows in places with low comparative advantage in DDS is associated to the highest transaction costs in payments made (outflows); the opposite occurs for payments received (inflows). This stands to reason as restrictions to cross-border data flows should make outflows more difficult.

This evidence indicates that indeed firms report higher transaction costs as a result of data flows restrictions.

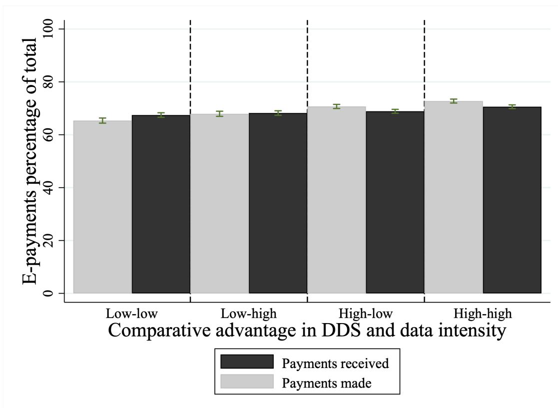
4 Theory: Firms, Institutions, and the Instruments of Data Sovereignty

Cross-border data flows create political economy conflicts distinct from goods trade. They are a core input into DDS, through which firms transmit personal information and deliver cloud-based services across borders (Weymouth, 2023; Ferracane, Kren and Van Der Marel, 2020). They also move sensitive information through platforms and infrastructures that governments do not fully control (Chander and Sun, 2022; Farrell and Newman, 2019*b*; Weymouth, 2025). We use *data sovereignty* to refer to policies that expand state authority over locally generated data by restricting or limiting its outbound movement, or requiring data processing to occur under domestic jurisdiction.

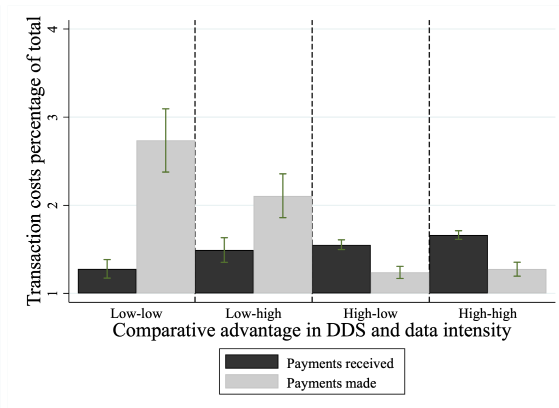
The argument explains why governments assert data sovereignty through different types of data flows restrictions. Strict localization measures require firms to keep data, processing activity, or infrastructure under domestic jurisdiction. Conditional flow regimes permit cross-border transfers when firms satisfy domestic legal requirements or when foreign jurisdictions provide adequate data protections (Farrell and Newman, 2019*a*; Ferracane, 2024; Mattoo and Meltzer, 2018). Both instruments expand public authority over data, but they impose different economic costs and reflect

Figure 4: Transaction costs for e-payments (2022-2025)

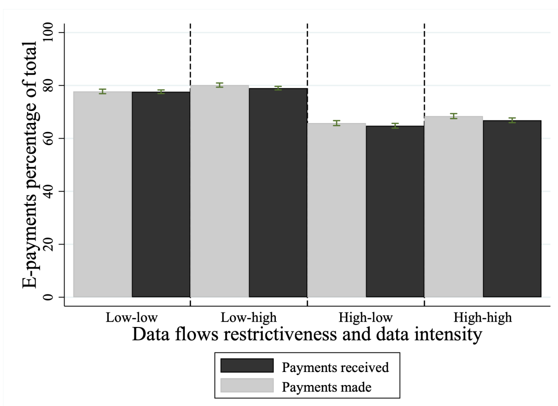
(a) Dependence on e-payments, by comparative advantage in DDS and data intensity



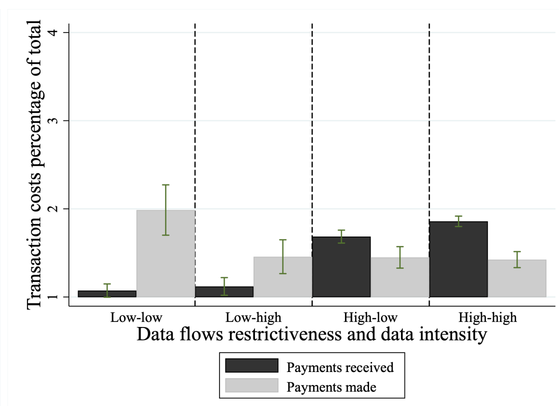
(b) Transaction costs, by comparative advantage in DDS and data intensity



(c) Dependence on e-payments, by Cross-border barriers to data flows and data intensity



(d) Transaction costs, by Cross-border barriers to data flows and data intensity



Note: Each bar displays a 95% confidence interval. Table D1 shows the questions we used to define e-payments and transaction costs. Descriptive statistics are presented in Table D2.

different political interests.

Our theory links DDS competitiveness to the domestic institutions that determine how governments assert authority over data. DDS competitiveness creates a pro-openness constituency because firms that export DDS or rely on cross-border data flows bear high costs from territorial restrictions. Political accountability shapes how that pressure enters policy because it gives firms and consumers stronger channels to shape the form of data sovereignty. Where firm pressure and consumer privacy concerns are both politically consequential, governments tend to avoid strict localization and establish conditional legal governance.

4.1 Digital Competitiveness and the Demand for Governed Data Flows

Strict localization measures impose uneven domestic costs. These costs fall most heavily on firms that depend on cross-border data flows. These measures raise compliance costs and disrupt cross-border service delivery by limiting where firms can process and transfer data. Firms exposed to these costs have strong incentives to support alternative rules that keep cross-border data transfers legally viable.

The constituency exposed to data flow restrictions extends beyond DDS exporters. Multinational firms organize production through cross-border data, and firms across sectors rely on foreign cloud services in their production process (Kim et al., 2019*a*; Weymouth, 2023; Ferracane, Kren and Van Der Marel, 2020; Girard, 2024). Strict localization measures disrupt these data architectures by forcing firms to change where data are stored and processed (Girard, 2024).

Strict localization can also benefit certain domestic actors. Local providers of data-processing or related digital services may benefit when strict data flows restrictions shift demand toward domestic alternatives (Chander and Sun, 2022; Gilbert et al., 2026; Girard and Wilhelm, 2025; Nebbiai, 2026). This coalition is strongest where domestic infrastructure can substitute for foreign providers. Where such substitutes are weak, strict localization measures impose broader adjustment costs on firms and generate fewer domestic beneficiaries, even when state officials value control.

Firms shape policy through organized political channels. Industry associations and digitally competitive multinational firms have pushed for open data flow commitments in trade negotiations and predictable cross-border transfer mechanisms in domestic regulation. The strength of this opposition should depend on the domestic importance of firms exposed to cross-border data transfer costs.

Comparative advantage in DDS captures a central dimension of those economic interests and the strength of this coalition. It measures the importance of DDS exports and proxies for the economic competitiveness of local firms in digital trade. The broader exposed constituency also includes firms that import digital services and firms integrated into global value chains, whose data dependence is only partly visible in services export data (Ferracane, Kren and Van Der Marel, 2020). Where comparative advantage in DDS is higher, the domestic costs of strict localization restrictions should be larger, and opposition to instruments that disrupt cross-border data flows should be stronger.

DDS competitiveness should therefore increase the political cost of territorial control. Strict localization directly interferes with the cross-border flows on which DDS firms depend. Where

DDS competitiveness is stronger, opposition to strict localization should be stronger as well.

Hypothesis 1. *Countries with a stronger revealed comparative advantage in digitally delivered services will adopt fewer strict localization measures.*

4.2 Political Accountability and the Form of Data Sovereignty

Open data flows create political vulnerabilities. Citizens expect governments to protect personal data and manage security risks when firms move data abroad, but data generated inside the jurisdiction often move through platforms and infrastructures outside domestic control (Chander and Sun, 2022; Weymouth, 2023, 2025; Farrell and Newman, 2019*b*).

Digital trade gives consumers a different role than standard accounts of goods trade emphasize. Consumers use foreign digital services, but they also become data subjects exposed to privacy risks (Weymouth, 2023; Zuboff, 2019). Those risks arise from how firms collect and transfer personal data. Individual consumers have limited capacity to monitor foreign data practices or bargain over transfer conditions.

Data sovereignty policies expand public authority over locally generated data. We argue that political institutions shape the form this authority takes. The key dimension is democratic accountability, which captures the degree to which citizens participate in politics and can hold governments accountable through that participation.

In higher-accountability regimes, citizens can sanction governments electorally and remain politically engaged between elections. These institutions make consumer privacy concerns harder for governments to ignore and give firms greater capacity to contest costly data restrictions. In accountable systems, these pressures generate demand for public rules that keep outbound data flows subject to domestic privacy protections while preserving access to digital services (Newman, 2008; Farrell and Newman, 2019*b*; Liu, 2021; Bradford, 2023). Conditional legal governance implements this conception of personal data protection. These regimes permit cross-border transfers when firms satisfy domestic legal requirements or when foreign jurisdictions provide adequate protection. Conditional transfer rules can diffuse because they give firms and governments a way to integrate digital markets while addressing consumer demand for enforceable privacy protections.(Greenleaf, 2019; Bradford, 2023).

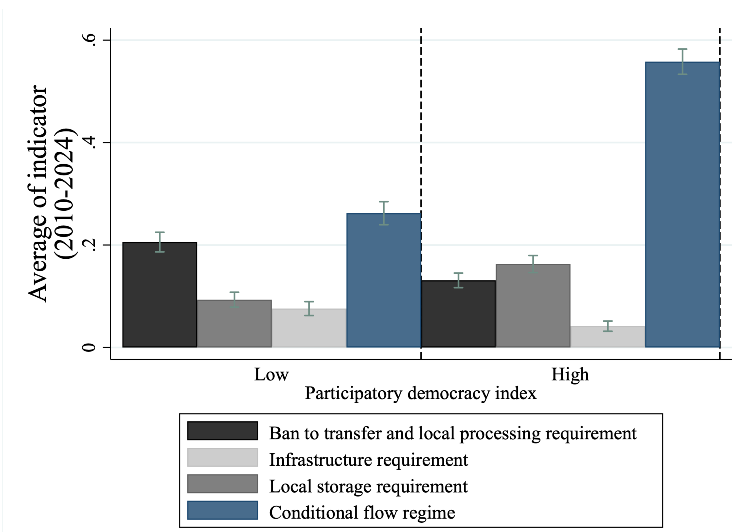
In lower-accountability regimes, incumbents face less public contestation over data policy and have greater ability to pursue territorial control, especially where control over data infrastructure strengthens the state's capacity for surveillance and censorship (King, Pan and Roberts, 2013; Lorentzen, 2014; Roberts, 2018; Gohdes, 2024). Strict localization measures fit lower-accountability settings because they bring data and processing activity under domestic jurisdiction and increase state influence over firms holding politically or economically valuable data.

Political accountability should therefore shape the form of data sovereignty. Lower-accountability regimes should rely more on territorial control because incumbents value state leverage and face less organized contestation over data policy. Higher-accountability regimes should rely more on conditional legal governance because consumer demand for privacy protections creates support for data governance that preserves mobility while extending domestic privacy protections to data transferred abroad.

Hypothesis 2. *Political institutions shape the form of data sovereignty. Lower-accountability regimes will rely more on strict localization measures, while higher-accountability regimes will rely more on conditional flow regimes.*

Figure 5 provides descriptive evidence consistent with this expectation. High-accountability regimes are twice more likely to adopt conditional flow regimes, while we observe that low-accountability regimes rely more often on territorial measures. Because political accountability is not plausibly exogenous, we interpret this pattern as evidence of association rather than a causal effect. We discuss below how accountability conditions the form of data sovereignty.

Figure 5: Data flow restrictions by political accountability (2005-2024)



Note: Low and high in the x-axis correspond to being below and above the mean of the participatory democracy index by Varieties of Democracy.

4.3 Digital Competitiveness, Accountability, and Instrument Choice

DDS competitiveness shapes data sovereignty most clearly in higher-accountability systems. Competitive DDS firms raise the political cost of territorial control because localization disrupts cross-border data flows. In these systems, firms have greater capacity to contest strict data policies, and consumer privacy concerns are harder for governments to ignore. Governments therefore face pressure to preserve cross-border data mobility while extending domestic privacy protections to data transferred abroad.

These political pressures act most clearly against strict localization. Firms oppose localization because it raises the costs of DDS delivery. Consumers are unlikely to favor territorial restrictions that limit access to DDS without clearly improving privacy protection. In higher-accountability systems with competitive DDS sectors, firm and consumer pressures should therefore most consistently reduce strict localization measures.

Conditional flow regimes can reconcile firm and consumer interests because they permit transfers to locations that satisfy domestic legal requirements. Consumers have strong reasons to support these rules because they protect personal data abroad. Firms, however, are unlikely to support conditional regimes unless privacy protections have diffused widely enough to keep transfer channels open. Without that diffusion, conditional regimes can operate like transfer bans, raising the costs of DDS delivery. The effect on conditional flow regimes should therefore depend on the balance between firm demand for open transfer and consumer demand for protected transfer.

Hypothesis 3. *Political accountability conditions the effect of DDS competitiveness on the form of data sovereignty. In higher-accountability regimes, stronger revealed comparative advantage in digitally delivered services will be associated with fewer strict localization measures.*

5 Empirical Strategy

Our main purpose is to estimate the effect of having a comparative advantage in DDS (the treatment) on data flows restrictiveness (the outcome), conditional on having low or high political accountability (the moderator). A relevant problem we face is that other variables that are hard to measure may affect both the treatment and the outcome. For instance, the existence of unmeasurable factors that generate both a comparative advantage and more liberalization in data flows regulation—like the existence of economies of scale in the digital economy—mean that a negative effect of comparative advantage on data flows restrictiveness could be overestimated due to this omitted variable bias.

Furthermore, to the best of our knowledge there is no evidence of a natural experiment or unforeseen shock taking place globally that would help us addressing the endogeneity issue through the use of instrumental variables or an event study design. While one could consider the advent of ChatGPT as a global event affecting the comparative advantage in DDS, our data is annual and there is no well-defined measure of cross-country differential intake of this treatment to guarantee exogeneity, leaving us with little room to no room to properly estimate statistically-consistent post-treatment effects using these designs. To address this issue, we develop an approach to obtain statistically-consistent estimates using exogenous interaction effects.

5.1 Instrumental Moderator Design

First, consider the simplified data generating process

$$Y = D\beta_1 + I\beta_2 + D \times I\beta_3 + \varepsilon,$$

where Y denotes the outcome, D is the treatment, I is a pre-treatment moderator, and ε is the idiosyncratic error term. If the moderator I is as good as random: $E(I'\varepsilon) = 0$, then the coefficient of the interaction effect (β_3) is statistically consistent. That is, the estimated value of the interaction effect ($\widehat{\beta_3}$) approaches the true coefficient (β_3) as the sample size grows: $plim\widehat{\beta_3} = \beta_3$ (where $plim$ denotes the probability limit), as we demonstrate in Lemma 1 below. This is a well-known result.

Note that if for instance I and D are dummy variables, where $D = 1$ indicates the observation is treated, and is zero otherwise, then

$$\text{plim}\widehat{\beta_3} = \beta_3 = \underbrace{\overbrace{E(Y|D=1,I=1) - E(Y|D=0,I=1)}^{\text{difference when } I=1} - \overbrace{[E(Y|D=1,I=0) - E(Y|D=0,I=0)]}^{\text{difference when } I=0}}_{\text{differential treatment effect}}.$$

Hence the differential treatment effect, or the difference in the treatment effect between those with $I = 1$ and $I = 0$, is statistically-identified in large-enough samples.

Lemma 1. *Consider the simplified data generating process*

$$Y = D\beta_1 + I\beta_2 + D \times I\beta_3 + \varepsilon.$$

If $E(I'\varepsilon) = 0$ *then* $plim\widehat{\beta_3} = \beta_3$.

Proof. *In the Appendix.*

Lemma 1 indicates that the exogeneity of I is a sufficient condition to estimate β_3. Note, however, that unlike an instrumental variable, I need not to satisfy the exclusion restriction—i.e., it can act through another channel on the outcome Y besides D. While I does not need to satisfy relevance either, we do want I to be related to D's data generating process for the sake of interpreting the β_3 as the differential effect of the quantity we measure with the treatment—in the spirit of instrumental variables design. In this case where $D = f(I, \cdot)$, however, Lemma 1 is insufficient to identify β_3.

Define $D = f(I, \chi)$ where an endogenous treatment (D) is a function of moderator (I) which is as good as random: $E(I'\varepsilon) = 0$. χ are other variables regarding the relationship between the treatment and the outcome (Y), including confounders. If f is a continuous function of its parameters and differentiable, then Proposition 1 follows directly from applying Lemma 1.[7] The implication is that if $D = \omega_1\chi + \omega_2 I + \nu$, with ν some idiosyncratic random error, we can estimate the coefficient associated to $\chi \times I$—even though we use the interaction term $D \times I$ in the regression—if we introduce the square of the moderator in the regression. This adjustment implies that the coefficient of interest is statistically-consistent as we show in Corollary 1.

Proposition 1 and Corollary 1 indicates that we can identify the differential effect of the endogenous component of the treatment as its exogenous component varies: the *differential treatment effect*. We refer to I in this case as a instrumental moderator (IM), and to this approach as *Instrumental Moderator Design* or IMD. The advantage of it is that the by using an instrumental moderator we can relax the assumptions needed in a standard instrumental variable design. In fact, we do not claim to estimate the treatment effect on the compliers—the local average treatment effect—but rather the differential treatment effect on the sample. Since the compliers are a sub-sample, this approach also addresses recent concerns regarding low statistical power (Arel-Bundock et al., 2026).

Proposition 1. *Consider*

$$Y = D\gamma_1 + I\gamma_2 + D \times I\gamma_3 + I^2\gamma_4 + \vartheta,$$

if $E(I'\vartheta) = 0$ *and* $D = f(I, \chi)$ *with* $f(\cdot)$ *linear in its parameters, then* $plim\widehat{\gamma}_3 = plim\widehat{\beta}_3 = \beta_3$.
Proof. *In the Appendix.*

Corollary 1. *Consider*

$$Y = \chi\alpha_1 + I\alpha_2 + \chi \times I\alpha_3 + I^2\alpha_4 + \mu,$$

and that

$$D = \omega_1\chi + \omega_2 I + \nu$$

then $plim\widehat{\alpha}_3 = \alpha_3$ *where* $\widehat{\alpha}_3 = \widehat{\omega_1\beta_3}$.

[7] While non-nonlinearities can be important, they are outside of the scope of this paper. The assumption of f linear suffices for our purposes; f can also be log-linearized if multiplicative functions are assumed.

Proof. *This follows directly from the proof of Proposition 1 and by applying Lemma 1.*

5.2 Land ruggedness as an IM

We use a measure of land ruggedness to proxy for the quasi-exogenous interaction term, I. Land ruggedness can determine the comparative advantage in DDS and is quasi-exogenous: inland broadband cable installation in rugged terrain faces significant challenges, requiring robust planning and specialized equipment to handle complex topography, steep gradients, weather, environmentally sensitive areas, and security security threats from non-state armed actors. These factors—especially the topographic hurdles—have hindered the expansion of the internet backbone in developing countries, as well in rural areas across the world (Bates, Geidner and Zhu, 2015).[8] The topographic hurdles then can increase costs and reduce productivity in the digital sectors, which are responsible of producing and trading data flows, negatively affecting the comparative advantage. Further, these topographic hurdles are extremely difficult and prohibitively expensive to alter, thus we can safely assume that they are largely determined by nature, providing credence to their quasi-exogenous nature.

We proxy land ruggedness using the standard index of terrain ruggedness by Nunn and Puga (2012); we refer the reader to the manuscript for details. Recall that we do not need to impose assumptions over the instrumental moderator as in instrumental variables design, apart from linearity. So in our regressions below, we provide a statistical test for this linearity by comparing a linear fit to a non-linear one for $RCA = f(ruggedness, \cdot)$, under the null hypothesis that they are the same. Despite this, we also make sure that $RCA = f(ruggedness, \cdot)$ is non-trivial. In other words, we check that *ruggedness* is a statistically-significant determinant of *RCA* conditional on covariates.

To facilitate interpretation, we perform a linear transformation of ruggedness such that a lower value of the index indicates more ruggedness and a higher one more smoothness. We call this variable *smoothness* and use it instead of ruggedness.

[8]Satellite solutions like Starlink have low adoption and are expensive, especially for single users, hence they do not bypass this issue.

5.3 Estimating the effect of comparative advantage in DDS using IMD

We estimate

$$DR_{ct} = \beta_1 RCA_{ct} + \beta_2 RCA_{ct} \times Smoothness_c + X'_{ct-1}\omega_1 + \psi_c + \lambda_t + \omega_{rt} + \varepsilon_{ct}.$$

DR_{ct} are the data flow restrictions. Since our main hypotheses make distinctions between strict localization measures and conditional regimes, we focus on these outcomes. Given that the data flow restrictiveness index defined earlier in Section 2 is additive, we separate it into a sub-index of strict localization measures (indicators 6.1-6.3 in Table 1), which we also normalize between zero and one, and the conditional flow regime indicator (6.4 in Table 1).[9] RCA_{ct} is the treatment, which represents our measure of comparative advantage in DDS defined in Section 3. $Accountability_{ct}$ is the index of participatory democracy by Variaties of Democracy (V-Dem), which measures the extent to which citizens actively engage in deliberation and decision-making, rather than solely exercising their right to vote; this is commensurate with our definition of accountability in Section 4.

X_{ct-1} is a vector of lagged pre-treatment confounding covariates including but not limited to cybersecurity capacity, digital authoritarianism, political accountability, GDP per capita, population, urbanization, employment, access to government-related business opportunities, and other. The data for these confounders is gleaned from the World Development Indicators and the V-Dem data set, since they provide wide country and temporal coverage.[10] We present descriptive statistics for all these variables in Table C1. ψ_c and λ_t control for country and time fixed effects respectively. Since $Smoothness_c$ is time invariant, ψ_c subsumes its square: $Smoothness_c^2$.[11] We also control for region-year fixed effects, ω_{rt}, to capture regional common regional trends using the World Bank regions, to further isolate country-level policy from regional-level covariates.

Our standard errors are clustered at the country level. We also use generalized doubly-robust estimators to address the problem of negative weights in two-way fixed effects models (Arkhangelsky et al., 2024).

Our coefficient of interest is β_2, or the differential treatment effect of the endogenous component of the treatment as its exogenous component varies. Since $Smoothness_c$ is quasi-exogenous, the coefficient is statistically identified as we demonstrate in Proposition 1. This identification

[9] We opt for using indexes as testing hypotheses on a family of outcomes can increase the probability of type II error (Anderson, 2008).

[10] To reduce the loss of information when using country-level data, we perform a selection process for our variables in Appendix C which identifies the variables that minimize the information loss. This also allows us to address statistical-power concerns recently highlighted by Arel-Bundock et al. (2026).

[11] Recall we need to control for the squared of the exogenous moderator by Proposition 1.

strategy allow us to approximate a quasi-experiment whereby we estimate the change in the effect of the comparative advantage in DDS if we could randomly assign its exogenous geographic component. We thus can interpret β_2 as the effect of comparative change on data flow restrictiveness, when geographies quasi-randomly improve.

Table 3 shows the results of estimating the aforementioned regression. Column (1) presents the results for the territorial control measures; column (2) does so for the conditional flow regime measure. To facilitate interpretation we standardize both RCA_{ct} and $Smoothness_c$ to one standard deviation.

Column (1) shows that an increase of one standard deviation (1SD) in the exogenous component of the revealed comparative advantage translates into a fall of two percentage points in the strict localization measures of data flows restrictiveness, when the revealed comparative advantage goes up by 1SD. This effect is substantive—it is equivalent to around a tenth of a standard deviation in the outcome of interest. This lends support to hypothesis 1. Column (2), in contrast, reveals that there is no statistically significant effect of the comparative advantage in DDS on the conditional flow regime. This indicates, as we theorize above, that the push for lower regulations by firms may offset concerns regarding private data security.

To interpret the effect on strict localization measures, we can consider the following case: Austria and Germany—which we pick due to their broad macro-level similarities—have both similar average values for the comparative advantage in DDS (1.049SDs v. 1.056SDs), but Austria has lower smoothness than Germany (2.57SDs v. 4.89SDs), thus the interaction diverges by two (2.69 v. 5.17). Having this in mind, if we could carry out an experiment where we could improve Austria's difficult geographies, we could increase Austria's exogenous component of the comparative advantage, and reduce strict localization measures by about four percentage points. This differential effect is not unrealistic since the observed difference in the value of the strict localization measure index between these countries is about ten percentage points.

Note our regressions pass the specification tests discussed—see also Table 3. First, we check the linearity assumption for $RCA = f(Smoothness, \cdot)$. To do this, we apply Frisch-Waugh-Lovell on two regressions: i) Where we assume *RCA* is a linear function of *Smoothness* and all other co-variates, and ii) Where we assume *RCA* is a non-linear function of *Smoothness*—using dummies for the percentiles of *Smoothness*—and all other co-variates. Then we test whether predictions for the residualized *RCA* as a function of *Smoothness* are statistically similar between both specifications—which would indicate that there are no non-nonlinearities. We accept this null hypothesis with an average difference between both specifications that is statistically zero. Hence the linearity assumption in Proposition 1 holds in this case.

Second, we check the significance on *Smoothness* for $RCA = f(Smoothness, \cdot)$. To do this we run a regression of *RCA* on *Smoothness* controlling for all covariates except the time fixed effects, as *Smoothness* is time-invariant. We observe not only that *Smoothness* is positively associated to *RCA*, indicating that better geographies do increase the comparative advantage, but that this relationship is statistically significant at 95% confidence.

Table 2: Effect of revealed comparative advantage on cross-border barriers to data flows

	Strict localization (1)	**Conditional flow regime** (2)
Differential treatment effect	-0.017**	0.006
of comparative advantage in DDS	(0.008)	(0.011)
IMD specification tests		
Linearity test for IM	0.001	0.001
	(0.006)	(0.006)
Significance test for IM	0.047**	0.047**
	(0.021)	(0.021)
Observations	2354	2354
Country FE	Yes	Yes
Year FE	Yes	Yes
Confounders	Yes	Yes

Notes: Standard errors clustered by country in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Controls include year fixed effects, country fixed effects, and region-year fixed effects. They also include includes lagged measures of a host of socioeconomicc confounders including cybersecurity capacity, digital authoritarianism, political accountability, GDP per capita, population, urbanization, employment, scientific production, and other. Estimates are heterogeneity robust.

Importantly, one may suspect that the "Snowden effect" is driving these results: this effect refers to the global paradigm shift in public awareness, corporate practices, and legal frameworks triggered by Edward Snowden's 2013 leaks, leading to massive legislative overhauls globally. It prompted countries to enact strict data flows regulations, forcing companies to store local citizens' data within domestic borders. The confounding effect of the Snowden effect is, however, of no concern to us because we control for year fixed effects and region-specific time trends, which capture these common overhauls. Furthermore, if any residual Snowden effect remains, it should attenuate the estimated relationship of the revealed comparative advantage with data flows restrictiveness. The estimates should therefore be interpreted as conservative rather than overestimated.

6 Results

We now estimate

$$
\begin{aligned}
DR_{ct} =& \beta_1 RCA_{ct} + \beta_2 RCA_{ct} \times Smoothness_c + \beta_3 RCA_{ct} \times Smoothness_c \times Accountability_{ct-1} \\
& + X'_{ct-1}\omega_1 + \psi_c + \lambda_t + \omega_{rt} + \varepsilon_{ct}
\end{aligned}
$$

to evaluate hypothesis 3—our main hypothesis. We define $Accountability_{ct-1} = 1$ if the value of V-Dem's participatory democracy index is above the median, zero otherwise, to reduce instability due to potential overfitting when computing interaction effects (Blackwell and Olson, 2022). To save on notation, X'_{ct-1} also contains all remaining component terms not explicit in the equation.

β_2 is the differential effect of the endogenous component of the treatment as its exogenous component varies for low accountability regimes: $Accountability_{ct-1} = 0$. In contrast, $\beta_2 + \beta_3$ is the differential effect of the endogenous component of the treatment as its exogenous component varies for high accountability regimes: $Accountability_{ct-1} = 1$. Since our hypothesis states that in high accountability regimes an increase in the comparative advantage in DDS should be associated with reduced strict localization measures, it must follow that $\beta_2 \geq 0$. We expect that $\beta_2 + \beta_3 \leq 0$ since high accountability should not have strong incentives to reduce digital data flows, consistent with our theory. The expected effect on conditional flow regimes is ambiguous on the basis of our theory because it depends on the balance between firm demand for open transfer and consumer demand for protected transfer.

Although we introduce an additional moderator—to estimate triple interaction effects—Proposition 1 and its corollary can be applied straightforwardly to the differential treatment effect conditional on political accountability; we do not require any additional assumptions.[12] Hence the estimates of β_2, β_3 and by extension $\beta_2 + \beta_3$, are statistically identified.

Table 3 shows the results. Overall, we observe that the effect we find for the comparative advantage on strict localization measures is driven by high accountability regimes. Further, we observe that political accountability conditions the effect of DDS competitiveness on the form of data sovereignty as hypothesized: we observe that in high-accountability regimes, stronger revealed comparative advantage in DDS is associated with fewer strict localization measures; we find once more no statistically significant effect on conditional flow regimes. In line with our theory, the latter result may emerge because the mechanisms of the comparative advantage and accountability can offset each other to some extent. This result suggests that firms demand for

[12] This is easy to check because we can define $D = RCA \times Accountability$, such that the mathematical proofs provided above hold.

Table 3: Effect of revealed comparative advantage on cross-border barriers to data flows

	Strict localization	**Conditional flow regime**
	(1)	(2)
Differential effect for of comparative advantage in DDS low political accountability	-0.001 (0.011)	-0.016 (0.022)
Differential effect for of comparative advantage in DDS high political accountability	-0.046** (0.0125)	0.026 (0.046)
IMD specification tests		
Linearity test for IM	0.001 (0.006)	0.001 (0.006)
Significance test for IM	0.047** (0.021)	0.047** (0.021)
Observations	2354	2354
Country FE	Yes	Yes
Year FE	Yes	Yes
Confounders	Yes	Yes

Notes: Standard errors clustered by country in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Controls include year fixed effect, country fixed effects, and region-year fixed effects. They also include includes lagged measures of a host of socioeconomicc confounders including cybersecurity capacity, digital authoritarianism, political accountability, GDP per capita, population, urbanization, employment, scientific production, and other. Estimates are heterogeneity robust.

open transfer and consumers demand for protected transfer cancel each other out.

In summary, our estimates indicate that, in high accountability regimes, an increase of 1SD in the exogenous component of the revealed comparative advantage, translates into a fall of five percentage points in the strict localization measures, when the revealed comparative advantage goes up by 1SD. This confirms hypothesis 3.

Robustness. In Appendix B, we perform several robustness checks: i) we show our results are robust to the presence of unobservable confounders via sensitivity analysis (Imbens, 2003), ii) using Jackknife, we also show that they are not driven specifically by any given county, and iii) we evaluate the robustness of our results using different measures of political accountability and an alternative expert-driven weighting scheme for the index. Our results hold for all of these alternative specifications. However, we would like to highlight that we also observe consistent differential effects for the comparative advantage on the data flows restrictiveness index—which considers both strict localization and conditional flow regimes measures—indicating that the comparative advantage in DDS reduces data flows restrictiveness altogether (Table B1).

7 Discussion

Our findings clarify why international cooperation on cross-border data flows has been so difficult to achieve. The conventional account points to preference heterogeneity: governments want different levels of restriction, and bargaining over a common standard is hard when underlying preferences diverge. The evidence presented here suggests a deeper obstacle. Governments that look similar on aggregate measures of restriction often use different instruments to assert authority over data, and those instruments embed different domestic political bargains.

A country that protects personal data through conditional transfer regimes is pursuing a different form of data sovereignty from one that imposes strict localization, even when both have a high level of restrictiveness in cross-national indices. Conditional regimes operationalize a rights-based conception of personal data through conditionality. Strict localization measures bring data infrastructure under domestic jurisdiction to expand state control. These instruments serve different coalitions, rest on different enforcement mechanisms, and respond to different political demands. Common rules are therefore harder to negotiate because the instruments are not politically interchangeable.

The WTO Joint Statement Initiative on E-Commerce illustrates the problem. WTO members launched exploratory work on e-commerce in 2017 and negotiations in 2019, but the stabilized text announced in 2024 excluded core disciplines on cross-border data flows. Countries have instead liberalized data flows selectively through preferential agreements and ad hoc mechanisms, including unilateral adequacy decisions, producing heterogeneous policies (Burri and Polanco, 2020; Burri, 2023; Ferracane et al., 2026). This pattern is consistent with instrument heterogeneity rather than simple differences in preferences across countries.

The political economy story we document explains why instrument heterogeneity persists. DDS competitiveness generates firm pressure for open data flows, but the form that pressure takes depends on the institutional setting. In accountable systems, firm pressure coexists with rights-based demands for legal protection, and conditional transfer regimes tend to emerge. In less accountable systems, control incentives dominate, and strict localization measures expand state leverage over firms and data infrastructures. Multilateral standard-setting requires harmonization across these different approaches.

This implies that building an IO regulatory authority in data governance will be difficult. Each instrument generates a distinct problem for IOs. Strict localization measures engage core sovereignty claims that IOs cannot easily address. Conditional regimes require regulatory cooperation and adequacy assessments that demand legal capacity. This helps explain why progress has occurred more often in preferential than multilateral venues.

Two implications follow for the study of IO governance over emerging technologies. First, aggregate measures of restriction obscure the politically meaningful variation. Future work on IO regulatory authority should disaggregate by instrument and ask which IOs are positioned to govern which instrument. Second, the durability of competing governance models depends on the durability of the domestic political bargains that anchor them.

Our findings also speak to the broader study of emerging technology governance. New technologies often outpace the rules governments use to regulate them, leaving states responsible for harms they cannot fully control. IOs may seek to govern these gaps, but their authority depends on whether domestic political bargains can be translated into agreements about the appropriate instruments. The central challenge is building international rules across states with strong divergences in digital competitiveness and governance incentives.

8 Conclusion

This paper began with a puzzle: why has international cooperation on cross-border data flows proven so difficult despite the rapid growth of DDS and the substantial economic gains associated with data mobility? We argue that variation in data-flow regimes is shaped by competitive DDS firms and the political institutions that determine whether firm and consumer demands influence policy.

Where firms possess a strong comparative advantage in DDS, the political costs of territorial control increase. Yet the consequences of this competitiveness depend on domestic institutions. In more accountable political systems, competitive digital sectors are associated with fewer restrictions on data transfers and, in particular, fewer strict localization measures. The relationship does not extend to conditional flow regimes, suggesting that firm pressure for data mobility does not simply produce deregulation. It can also coexist with consumer demand for rules that keep data transfers subject to domestic legal protections.

These findings carry broader implications for the study of international political economy. Data-flow regulation requires a political economy of data sovereignty rather than a simple extension of goods-trade protection. Governments rely on digital markets, but they make different choices about how much authority to exercise over the data that sustain those markets. The politics of goods trade do not extend cleanly to data flows because the rules governing data mobility also define citizens' rights and the state's control over information.

This perspective also sheds light on the persistent difficulties facing IOs. Efforts to construct multilateral rules for the digital economy confront competing models of governance rooted in different domestic political bargains. Governments that favor localization and governments that favor conditional transfer regimes may appear equally committed to data sovereignty, yet they are pursuing different visions of the relationship between markets, rights, and state authority. The challenge for international cooperation is therefore not merely to reduce barriers to data flows, but to reconcile these competing governance objectives.

References

Anderson, Michael L. 2008. "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American statistical Association* 103(484):1481–1495.

Arel-Bundock, Vincent, Ryan C Briggs, Hristos Doucouliagos, Marco M Aviña and Tom D Stanley. 2026. "Quantitative political science research is greatly underpowered." *The Journal of politics* 88(1):36–46.

Arkhangelsky, Dmitry, Guido W Imbens, Lihua Lei and Xiaoman Luo. 2024. "Design-robust two-way-fixed-effects regression for panel data." *Quantitative Economics* 15(4):999–1034.

Arnold, Jens Matthias, Beata Javorcik, Molly Lipscomb and Aaditya Mattoo. 2016. "Services reform and manufacturing performance: Evidence from India." *The Economic Journal* 126(590):1–39.

Azmeh, Shamel, Christopher Foster and Jaime Echavarri. 2020. "The International Trade Regime and the Quest for Free Digital Trade." *International Studies Review* 22(3):671–692.

Baccini, Leonardo, Pablo M. Pinto and Stephen Weymouth. 2017. "The Distributional Consequences of Preferential Trade Liberalization: Firm-Level Evidence." *International Organization* 71(2):373–395.

Balcázar, Carlos Felipe and Christopher Pissarides. 2025. *World Openness Report 2025*. Royal Collins chapter Cross-border data flows: the backbone of collaborative AI and digital trade in services.

Balcazar, Carlos Felipe and Matt Malis. 2022. Think locally, regress globally: making the most of conventional IR data. In *Handbook of Research Methods in International Relations*. Edward Elgar Publishing pp. 317–340.

Barnum, Miriam. 2022. Dealing with missing and incomplete data. In *Handbook of Research Methods in International Relations*. Edward Elgar Publishing pp. 425–445.

Bates, Benjamin J, Nick Geidner and Yi-Xi Zhu. 2015. The economics of broadband. In *Handbook on the Economics of the Media*. Edward Elgar Publishing pp. 188–217.

Beaumier, Guillaume and Lars Gjesvik. 2025. "Digital governance in a rubber band: Structural constraints in governing a global digital economy." *Global Studies Quarterly* 5(2):ksaf043.

Blackwell, Matthew and Michael P Olson. 2022. "Reducing model misspecification and bias in the estimation of interactions." *Political Analysis* 30(4):495–514.

Bradford, Anu. 2020. *The Brussels Effect: How the European Union Rules the World*. New York: Oxford University Press.

Bradford, Anu. 2023. *Digital empires: The global battle to regulate technology*. Oxford University Press.

Burri, Mira. 2023. "Trade Law 4.0: Are We There Yet?" *Journal of International Economic Law* 26(1):90–100.

Burri, Mira and Rodrigo Polanco. 2020. "Digital trade provisions in preferential trade agreements: introducing a new dataset." *Journal of international economic law* 23(1):187–220.

Chander, Anupam and Haochen Sun. 2022. "Sovereignty 2.0." *Vand. J. Transnat'l L.* 55:283.

Chander, Anupam and Haochen Sun, eds. 2023. *Data Sovereignty: From the Digital Silk Road to the Return of the State*. New York: Oxford University Press.

Chander, Anupam and Uyên P. Lê. 2015. "Data Nationalism." *Emory Law Journal* 64(3):677–739.

Chen, Jidong and Yiqing Xu. 2017. "Why Do Authoritarian Regimes Allow Citizens to Voice Opinions Publicly?" *Journal of Politics* 79(3):792–803.

Farrell, Henry and Abraham L. Newman. 2019*a*. *Of Privacy and Power: The Transatlantic Struggle over Freedom and Security*. Princeton, NJ: Princeton University Press.

Farrell, Henry and Abraham L. Newman. 2019*b*. "Weaponized Interdependence: How Global Economic Networks Shape State Coercion." *International Security* 44(1):42–79.

Ferracane, Martina F. 2017. Restrictions on cross-border data flows. Technical report ECIPE working paper.
URL: *https://papers.ssrn.com/sol3/papers.cfm?abstract$_i$d* = 3089956

Ferracane, Martina F. 2021. "The costs of data protectionism." *Big data and global trade law* pp. 63–82.

Ferracane, Martina F, Bernard Hoekman, Erik van der Marel and Filippo Santi. 2026. "Digital trade, data protection and the EU adequacy club." *Economica* 93(369):230–259.

Ferracane, Martina F. and Erik van der Marel. 2021. "Do Data Policy Restrictions Inhibit Trade in Services?" *Review of World Economics* 157:727–776.

Ferracane, Martina F and Erik van Der Marel. 2025. “Governing personal data and trade in digital services.” *Review of International Economics* 33(1):243–264.

Ferracane, Martina F, Simón González Ugarte and Erik van der Marel. 2025. “The Brussels effect in Africa: is it beneficial for intra-regional trade in digital services?” *Journal of International Economic Law* 28(1):1–22.

Ferracane, Martina F., Simón González Ugarte and Tomás Rogaler. 2026. “Global trends in digital trade policies and practices: Evidence from the Digital Trade Integration Database.” *World Trade Review* 25(1):127–151.

Ferracane, Martina Francesca. 2024. Data Localization. In *Elgar Encyclopedia of International Economic Law*. Edward Elgar Publishing Limited.

Ferracane, Martina Francesca, Janez Kren and Erik Van Der Marel. 2020. “Do data policy restrictions impact the productivity performance of firms and industries?” *Review of International Economics* 28(3):676–722.

Gilbert, Abigail, Noam Shemtov, Philip Treleaven, Patricia Shaw, Steven Rolf, Christopher Foster, Edward Challis, Carlos Balcazar, Swee Leng Harris, Joseph Baines, Julian Germann and Tom Peters. 2026. Economic Sovereignty and the Question of Post-Deployment Training. Technical report Institute for the Future of Work.

Girard, Tyler. 2024. “National identity and the limits of platform power in the global economy.” *International Studies Quarterly* 68(3):sqae090.

Girard, Tyler and Alexander Wilhelm. 2025. “Local Data Policies, Global Data Politics: How Citizens Evaluate Data Localization Policies and Political Responses.” *Foreign Policy Analysis* 21(4):oraf007.

Gohdes, Anita R. 2024. *Repression in the Digital Age: Surveillance, Censorship, and the Dynamics of State Violence*. New York: Oxford University Press.

Goldfarb, Avi and Catherine Tucker. 2019. “Digital economics.” *Journal of Economic Literature* 57(1):3–43.

Greenleaf, Graham. 2019. “Global Data Privacy Laws 2019: 132 National Laws and Many Bills.” *Privacy Laws and Business International Report* 157:14–18.

Hulvey, Samantha B. and Beth A. Simmons. 2025. “Cyber Borders: Exercising State Sovereignty Online.” *International Studies Quarterly* . Forthcoming; early view.

Imbens, Guido W. 2003. "Sensitivity to exogeneity assumptions in program evaluation." *American Economic Review* 93(2):126–132.

Jensen, J. Bradford, Dennis P. Quinn and Stephen Weymouth. 2015. "The Influence of Firm Global Supply Chains and Foreign Currency Undervaluations on U.S. Trade Disputes." *International Organization* 69(4):913–947.

Jia, Jian, Ginger Zhe Jin and Liad Wagman. 2021. "The short-run effects of the general data protection regulation on technology venture investment." *Marketing Science* 40(4):661–684.

Ker, Daniel and Emanuele Mazzini. 2020. "Perspectives on the value of data and data flows." *OECD Digital Economy Papers* .

Kim, In Song. 2017. "Political Cleavages within Industry: Firm-level Lobbying for Trade Liberalization." *American Political Science Review* 111(1):1–20.

Kim, In Song, Helen V. Milner, Thomas Bernauer, Iain Osgood, Gabriele Spilker and Dustin Tingley. 2019*a*. "Firms and Global Value Chains: Identifying Firms' Multidimensional Trade Preferences." *International Studies Quarterly* 63(1):153–167.

Kim, In Song, Helen V Milner, Thomas Bernauer, Iain Osgood, Gabriele Spilker and Dustin Tingley. 2019*b*. "Firms and global value chains: Identifying firms' multidimensional trade preferences." *International Studies Quarterly* 63(1):153–167.

King, Gary, Jennifer Pan and Margaret E. Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107(2):326–343.

Little, Andrew T and Anne Meng. 2024. "Measuring democratic backsliding." *PS: Political Science & Politics* 57(2):149–161.

Liu, Lizhi. 2021. "The rise of data politics: Digital China and the world." *Studies in comparative international development* 56(1):45–67.

Lorentzen, Peter L. 2014. "China's Strategic Censorship." *American Journal of Political Science* 58(2):402–414.

Mattoo, Aaditya and Joshua P. Meltzer. 2018. "International Data Flows and Privacy: The Conflict and Its Resolution." *Journal of International Economic Law* 21(4):769–789.

Meltzer, Joshua Paul. 2015. "The I nternet, Cross-Border Data Flows and International Trade." *Asia & the Pacific Policy Studies* 2(1):90–102.

Nebbiai, Matteo. 2026. "Data-Driven Lobbying: When Firms Want Data Regulation." *Available at SSRN 6710898* .

Newman, Abraham L. 2008. *Protectors of Privacy: Regulating Personal Data in the Global Economy*. Ithaca: Cornell University Press.

Nunn, Nathan and Diego Puga. 2012. "Ruggedness: The blessing of bad geography in Africa." *Review of Economics and Statistics* 94(1):20–36.

OECD. 2023. *Handbook on Measuring Digital Trade*. Paris: OECD Publishing. Geneva 10/WTO, Geneva.

OECD. 2025. *Digital Trade for Development*. Paris: OECD Publishing. Geneva 10/WTO, Geneva.

Oppenheimer, Harry. 2025. "Digital Interdependence and Power Politics." *British Journal of Political Science* 55:e93.

Osgood, I. 2021. "Vanguards of Globalization: Organization and Political Action among America's Pro-trade Firms." *Business and Politics* 23(1):1–35.

Osgood, Iain. 2017. "The Breakdown of Industrial Opposition to Trade: Firms, Product Variety, and Reciprocal Liberalization." *World Politics* 69(1):184–231.

Osgood, Iain. 2018. "Globalizing the supply chain: Firm and industrial support for US trade agreements." *International Organization* 72(2):455–484.

Polyakova, Alina and Chris Meserole. 2019. "Exporting Digital Authoritarianism: The Russian and Chinese Models." *Brookings Policy Brief* .

Rentzhog, Magnus. 2015. *No Transfer, No Production – A Report on Cross-Border Data Transfers, Global Value Chains, and the Production of Goods, 3rd edn*. Kommerskollegium.

Rentzhog, Magnus and Henrik Jonströmer. 2014. *No Transfer, No Trade: The Importance of Cross-Border Data Transfer for Companies Based in Sweden*. Kommerskollegium.

Roberts, Margaret E. 2018. *Censored: Distraction and Diversion Inside China's Great Firewall*. Princeton, NJ: Princeton University Press.

Schwartz, Paul M. 2019. "Global Data Privacy: The EU Way." *New York University Law Review* 94(4):771–818.

Weymouth, Stephen. 2023. *Digital Globalization: Politics, Policy, and a Governance Paradox*. Cambridge: Cambridge University Press.

Weymouth, Stephen. 2025. "Digital Disintegration: Techno-Blocs and Strategic Sovereignty in the AI Era." *International Organization* 79(S1):S57–S70.

Whittle, Devon. 2024. "Digital: The WTO E-Commerce JSI's Stabilised Text - what does it deliver, who couldn't support it, and what are its next steps?" *Trade Notes* .
URL: *https://tradenotes.substack.com/p/digital-the-wto-e-commerce-jsis-stabilised*

World Trade Organization. 2025*a*. "Digital trade is key to boosting growth in developing economies." accessed May 2025 https://www.wto.org/english/blogs_e/ce_ralph_ossa_e/blog_ro_15dec23_e.htm.

World Trade Organization. 2025*b*. *Economic Implications of Data Regulation: Balancing Openess and Trust*. Geneva: World Trade Organization.

Zhang, Hongxia, Heeho Kim and Sangkil Moon. 2025. "The effects of data restriction policies and institutions on digital financial service and trade." *Emerging Markets Finance and Trade* 61(9):2745–2762.

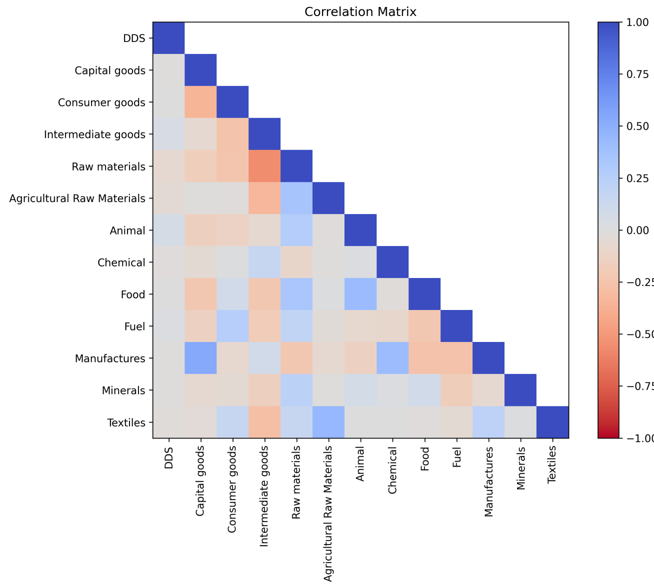
Zuboff, S. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Profile Books.

.tocmtappendix mtchapternone mtappendixsubsection mtappendixsubsubsection

Online Appendix

Table of Contents

Figure A1: Correlations matrix for revealed comparative advantages (2005-2024)



Note: The first column of the matrix shows that the correlation between the comparative advantage in DDS and the comparative advantage in other sectors is almost zero.

Table A1: Data-driven weights obtained from MCA

Indicator	Criteria	Score
6.1. Ban on transfer and local processing requirement	1) Ban/local processing for an entire sector or horizontal or personal data, or more than one measure in category (2).	0.57
	2) Ban/local processing for specific data or transfer prohibited to one country.	-0.25
	3) No restrictions.	-3.61
6.2. Local storage requirement	1) Local storage for entire sector or horizontal or personal data, or more than one measure in category (2).	1.39
	2) Local storage for specific data.	0.08
	3) No restrictions.	-1.53
6.3. Infrastructure requirement	1) Infrastructure requirement.	0.33
	2) No restrictions.	-5.43
6.4. Conditional flow regime	1) Conditions for personal data/entire sector.	0.177
	2) Conditions for specific data/light conditions.	1.52
	3) No restrictions.	-0.33

Notes: Scores are obtained using the first coordinate after Multiple Correspondence Analysis (MCA). Scores satisfy ordering consistency.

Table A2: Revealed data intensity measures

IPD-E code	Sector description	EBOPS 2010	ISIC Rev. 4	Dig 1	Dig 2	Dig 3	Dig 4	Dig. Deliv.	S/L ratio	Share PS
162	Telecom, computer, and information	SI+SK1	J	x	x	x	x	Yes	4.1	2.59
163	Other business services	SJ excl SJ34	M+N	x	x	x	x	Yes	1.08	0.08
159	Insurance and pension services	SF	K(60%)		x	x	x	Yes	2.88	0.05
160	Financial services	SG	K(40%)		x	x	x	Yes	2.87	0.05
164	Heritage and recreational services	SK23	R			x	x	Yes	0.54	0.09
169	Trade-related services	SJ34	G			x	x	Yes	0.51	0.02
161	Charges for the use of IPR	SH	-				x	Yes	-	-
165	Health services	SDB1+SK21	Q				x	Yes	0.22	0.04
166	Education services	SDB2+SK22	P				x	Yes	0.28	0.11
170	Other personal services	SK24	S				x	Yes	0.09	-
154	Manufacturing services	SA	-					No	-	-
155	Maintenance and repair	SB	-					No	-	-
156	Transport	SC	H					No	0.58	-
157	Travel	SDA+SDB3	I					No	0.04	0.08
158	Construction	SE	F					No	0.06	0
167	Government goods and services	SL	-					No	-	-
168	Services not allocated	SN	-					No	0.18	-

B Robustness

Some relevant concerns may arise from the empirical design: i) The possibility that confounders that are not accounted for may overturn the results, ii) The possibility that any single observation is driving the results, and iii) Sensitivity to measurement. We explore these potential issues next:

Sensitivity to unobserved confounding. We check the sensitivity of the estimated results with respect to deviations from the conditional exogeneity assumption; i.e. if there are unobserved variables that affect assignment to treatment and the outcome variable simultaneously that the estimated coefficients may not be robust to. We explicitly relax the exogeneity assumption by allowing for a limited amount of correlation between treatment and unobserved components of the outcome (Imbens, 2003). We find that an unobservable confounder that could potentially overturn the main results needs to exhibit a higher partial R^2 vis-á-vis than the confounders already included (Figure B1).

Parameter stability to countries. To further corroborate that there's no one country observation driving the results, we carry out a robustness tests wherein we drop on congressional district at a time with replacement (*á la* Jackknife). We find that the treatment indicator is quite stable and statistically significant for each permutation (Figure B2).

Alternative measures of accountability. The are various ways to measure political accountability, which can provide contrasting conclusions (e.g., Little and Meng 2024). We seek within the V-DEM project for similar concepts/measures to the one we use in our main results; we refrain from using PolityIV and Freedom house due to their more limited time and country coverage. Specifically, we consider `egaldem` (Figure B3, panel a) which measures the distribution of de jure and de facto power, and `eqaccess` (Figure B3, panel v) wich measures how equally citizens can access de jure power. This are measures of the equal distribution of political influence. All in all, Figure B3 shows that our results hold to using these measures instead of participatory democracy (`partipdem`).

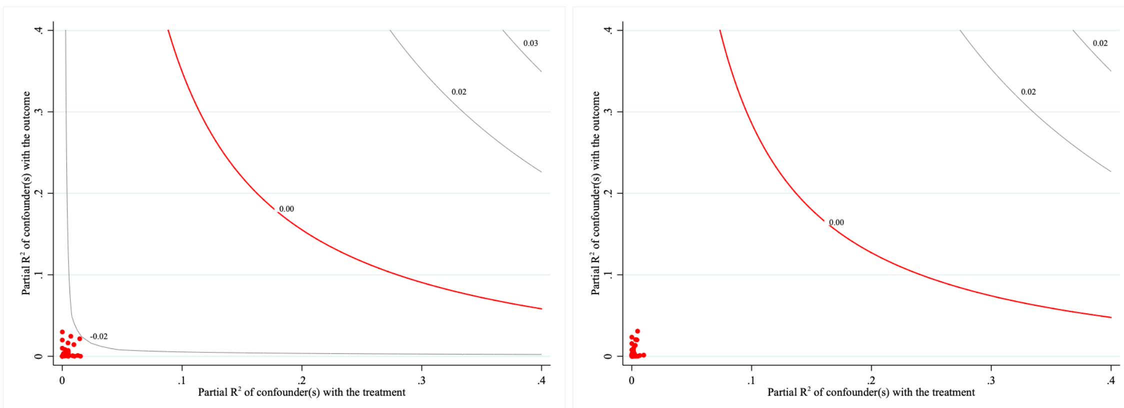
Alternative measures of digital trade barriers. We now use an alternative digital restrictiveness index based on weights drawn from consulting with multiple policy makers, both in governments and international organizations, which measures are more restrictive. Then we assign scores from 0 to 1 to the most restrictive measures for each indicator, and a weight. (See Table B2.) This was done as part of the Digital Trade Integration project; the index takes values between zero and one.

The index obtained from using expert weights has a high correlation coefficient, of 0.9, with the index obtained from applying MCA. Therefore we expect similar results, as we do obtain (Table B1, columns (1) and (2)). We also show that our results hold for the data flows restrictiveness index, which captures indicators 6.1-6.4, in columns columns (3) and (4).

Figure B1: Sensitivity analysis to unobserved confounding

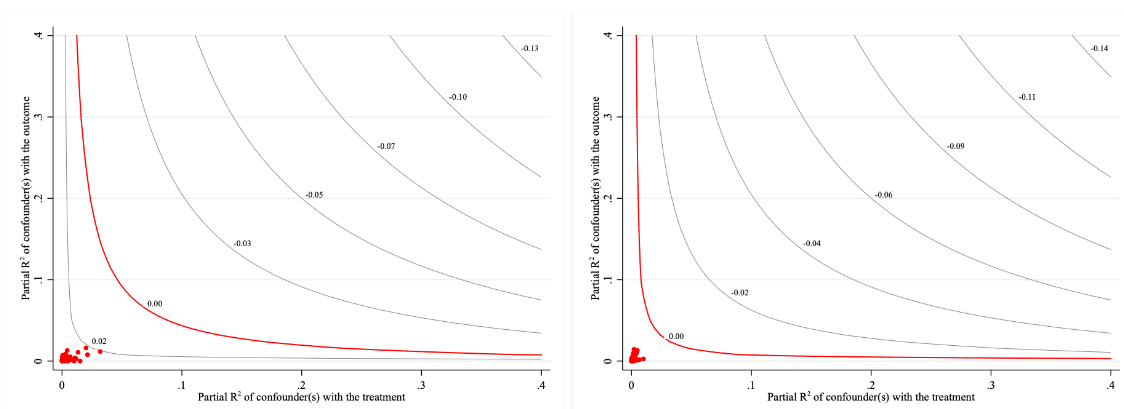
(a) Effect for low accountability, strict localization

(b) Differential effect for political accountability, strict localization



(c) Effect for low accountability, conditional flow regime

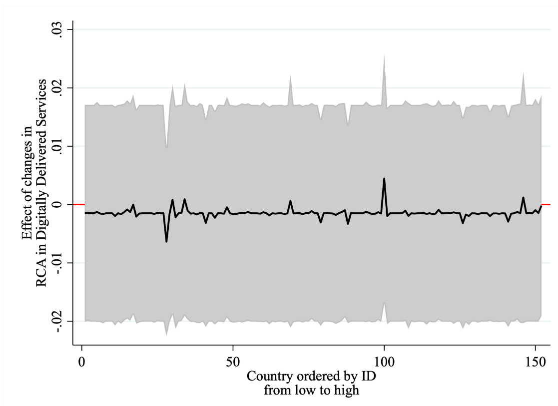
(d) Differential effect for political accountability, conditional flow regime



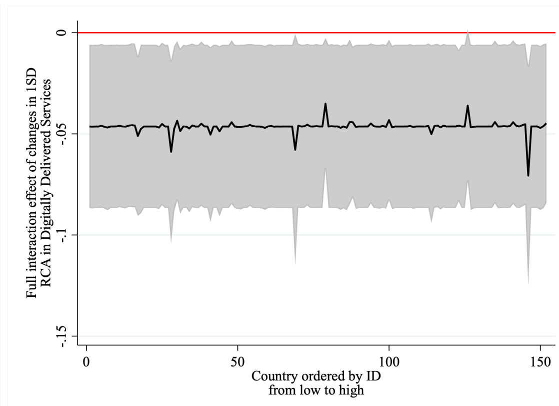
Note: Test clustering by country. Controls include year fixed effect, country fixed effects, and region-year fixed effects. They also include includes lagged measures of a host of socioeconomicc confounders including cybersecurity capacity, digital authoritarianism, political accountability, GDP per capita, population, urbanization, employment, scientific production, and other. Estimates are heterogeneity robust.

Figure B2: Parameter stability to excluding one country with replacement

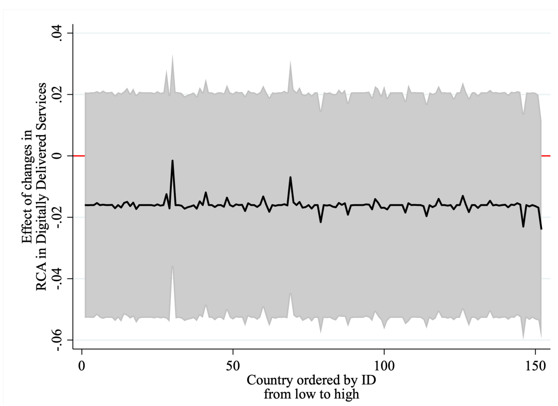
(a) Effect for low accountability, strict localization



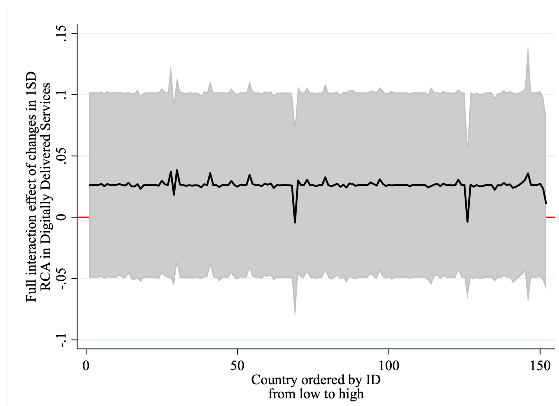
(b) Effect for high accountability, strict localization



(c) Effect for low accountability, conditional flow regime

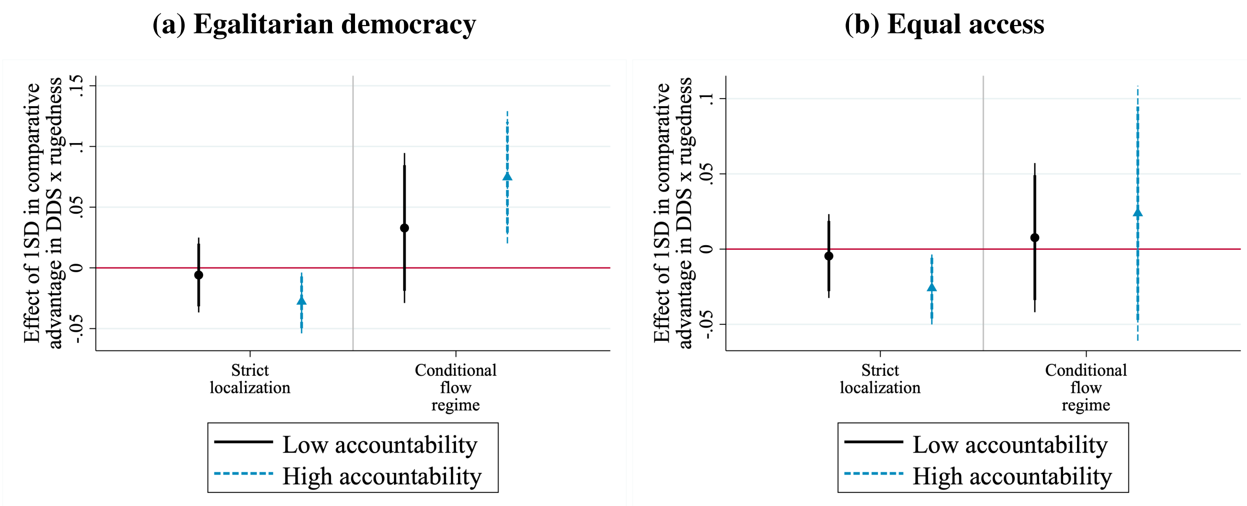


(d) Effect for high accountability, conditional flow regime



95% confidence bands clustered by country in gray. Controls include year fixed effect, country fixed effects, and region-year fixed effects. They also include includes lagged measures of a host of socioeconomicc confounders including cybersecurity capacity, digital authoritarianism, political accountability, GDP per capita, population, urbanization, employment, scientific production, and other. Estimates are heterogeneity robust. Estimates are heterogeneity robust.

Figure B3: Effect of revealed comparative advantage on cross-border barriers to data flows



Notes: 95% confidence bands clustered by country in gray; 95% (90%) confidence intervals clustered by country in thin (thick) spikes. Controls include year fixed effect, country fixed effects, and region-year fixed effects. They also include includes lagged measures of a host of socioeconomicc confounders including cybersecurity capacity, digital authoritarianism, political accountability, GDP per capita, population, urbanization, employment, scientific production, and other. Estimates are heterogeneity robust. Estimates are heterogeneity robust.

Table B1: Effect of revealed comparative advantage on cross-border barriers to data flows, using expert weights

	Using expert weights		Using MCA	
	Strict measures	Data flows restrictiveness index	Data flows localization index	
	(1)	*(2)*	*(3)*	*(4)*
A. Differential effect				
Differential treatment effect of comparative advantage in DDS	-0.022** (0.010)		-0.018** (0.008)	
B. Differential effect by accountability				
Differential effect for of comparative advantage in DDS low political accountability		-0.008 (0.013)		-0.003 (0.011)
Differential effect for of comparative advantage in DDS high political accountability		-0.042** (0.013)		-0.045** (0.015)
IMD specification tests				
Linearity test for IM	0.001	0.001	-0.001	-0.001
	(0.006)	(0.006)	(0.006)	(0.006)
Significance test for IM	0.047**	0.047**	0.047**	0.047**
	(0.021)	(0.021)	(0.021)	(0.021)
Observations	2354	2354	2354	2354
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Confounders	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by country in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Controls include year fixed effect, country fixed effects, and region-year fixed effects. They also include includes lagged measures of a host of socioeconomicc confounders including cybersecurity capacity, digital authoritarianism, political accountability, GDP per capita, population, urbanization, employment, scientific production, and other. Estimates are heterogeneity robust. Estimates are heterogeneity robust.

Table B2: Expert weights

Indicator	Criteria	Score	Weight
6.1. Ban on transfer and local processing requirement	1) Ban/local processing for an entire sector or horizontal or personal data, or more than one measure in category (2). 2) Ban/local processing for specific data or transfer prohibited to one country. 3) No restrictions.	1 0.5 0	1
6.2. Local storage requirement	1) Local storage for entire sector or horizontal or personal data, or more than one measure in category (2). 2) Local storage for specific data. 3) No restrictions.	1 0.5 0	0.3
6.3. Infrastructure requirement	1) Infrastructure requirement. 2) No restrictions.	1 0	0.8
6.4. Conditional flow regime	1) Conditions for personal data/entire sector. 2) Conditions for specific data/light conditions. 3) No restrictions.	1 0.5 0	0.3

C Dealing with missingness

One of the important limitations dealing with country-level data, is that oftentimes the information is not measured for many countries and (or) several periods of time, creating a problem of missingness. This missingness can reduce the power of our tests leading to higher probability of type II error (Arel-Bundock et al., 2026). Imputation is also not desirable as it is well-established that it can bias estimates in regression analysis because variables are function of other variables plus a random element, hence not allowing us to properly adress confounding (e.g., Balcazar and Malis 2022; Barnum 2022).

To deal with the previous problem we consider that missingness can be of three types: i) Not all variables are measured for every country, hence using the variable would necessarily mean dropping a country of the sample, ii) Not all variables are measured in every year, hence using the variable means dropping an entire year of the sample, and iii) There are variable that are not measured in every year for every country, therefore using the variable implies creating an unbalanced panel. Since we will be using fixed effects to exploit variation within countries, we want to maximize that amount of country-years in the sample—hence we want to have the most amount of information, and the data to be as-close-to a balanced panel as we can.

We consider multiple country-level datasets such as the World Development Indicators, Varieties of Democracy, Correlates of War, and other, encompassing about 526 variables that could be potential confounders. With these data we generate a score of the amount of missingness between country for every variable: the percentage of year between 2005-2024 with missing multiplied by the range of years that the variable covers. Then, we obtain the average score by year for that variable. This provides us with a score for average within-country missingness that scales linearly with time, implying that we give a higher weight to more recent information. The (reciprocal) cumulative distribution of this score is given by Figure C1 below.

We establish the cut-off using the density of this variable, dropping those above 1.3, which are those variable that comprise few countries but have many missing—noted by the elbow in the reciprocal of the cumulative density function. Using the "elbow" or "knee of a curve" as a cutoff point is a common heuristic in mathematical optimization to choose a point where diminishing returns are no longer worth the additional cost. The intuition is that increasing the number of variables will naturally improve the fit (explain more of the variation), since there are more parameters to use, but that at some point this reduces the sample size and can induce over-fitting over a sample that falls in size with the number of variables included. This allow us to drop those variables that generate substantial information loss.

The descriptive statistics of the variables chosen are presented in Table C1 below. We observe that we select a subset of variables from Varieties of Democracy and the World Development Indicators after our procedure given their wide country and temporal coverage.

Figure C1: Missingness scores

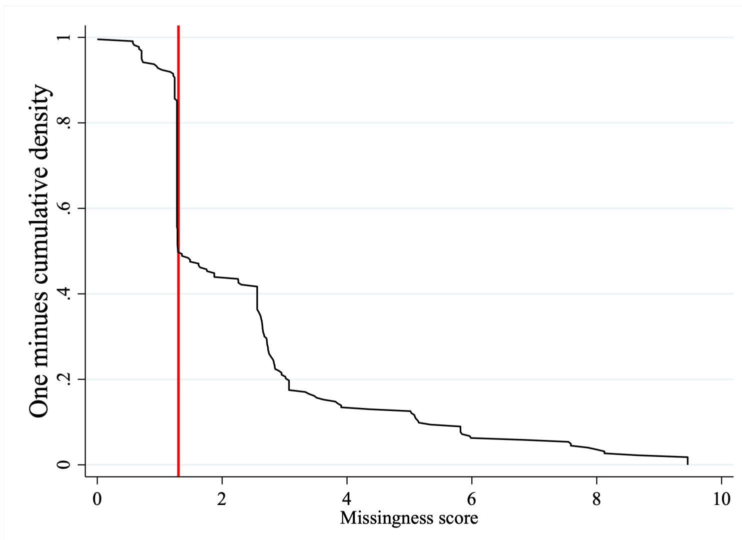


Table C1: Summary statistics

Variable	Average	Std. dev.	Min	Max	N	Quartiles of cross border data index Q1	Q2	Q3	Q4
Index of cross-border data barriers	**0.20**	**0.18**	**0.00**	**1.00**	**3081.00**	**0.10**	**0.15**	**0.21**	**0.45**
Strict localization	0.22	0.18	0.00	1.00	3081.00	0.13	0.15	0.22	0.45
Ban on transfer and local processsing	0.16	0.33	0.00	1.00	3081.00	0.00	0.00	0.30	0.58
Local storage requirement	0.13	0.31	0.00	1.00	3081.00	0.06	0.00	0.41	0.30
Infraestructure requirement	0.06	0.24	0.00	1.00	3081.00	0.00	0.00	0.00	0.25
Conditional flow regime	0.41	0.49	0.00	1.00	3081.00	0.09	1.00	0.00	0.61
A. Digitally delivered services									
RCA in DDS	2.09	1.51	0.00	13.08	3152.00	2.07	2.31	2.28	1.92
DDS exports (mill. USD)	13522.26	48469.40	0.00	741043.00	3563.00	8516.21	17871.61	16124.44	24589.55
DDS imports (mill. USD)	11984.60	35687.61	-13.68	454992.44	3563.00	6603.25	18767.01	12009.96	21089.48
Exports (mill. USD)	66256.20	226955.86	0.00	2957001.00	3319.00	35545.88	101562.72	104345.83	125462.63
Imports (mill. USD)	64452.21	193458.17	0.00	2539368.00	3319.00	40654.07	79218.51	100578.80	134864.19
C.Institutions									
Participatory democracy index	0.34	0.20	0.01	0.81	3828.00	0.30	0.44	0.27	0.35
Egalitarian democracy index	0.39	0.24	0.02	0.88	3828.00	0.35	0.52	0.31	0.41
Equal protection index	0.64	0.25	0.02	0.99	3828.00	0.59	0.74	0.49	0.66
Government cyber security capacity	-0.10	1.32	-3.60	3.71	3828.00	-0.50	0.11	-0.10	0.70
Government Internet shut down in practice	0.65	1.27	-3.90	2.10	3828.00	0.50	0.97	0.48	0.64
Government social media monitoring	0.22	1.43	-3.59	2.99	3828.00	0.19	0.80	-0.42	-0.17
Government social media shut down in practice	0.64	1.30	-3.83	2.12	3828.00	0.50	1.04	0.50	0.54
Foreign governments dissemination of false information	0.32	1.23	-3.88	3.21	3828.00	0.38	0.54	0.21	0.11
D. Macroeconomics									
FDI, net inflows (mill. USD	10224.27	41883.97	-343402.81	733826.50	3603.00	5697.14	13578.39	18636.07	22995.49
GDP per capita (cUSD)	15933.36	26065.43	147.23	288001.44	3719.00	11789.32	19788.76	10824.32	16054.79
Population density (people/sq. km)	287.56	1394.48	0.14	18822.89	3576.00	139.83	278.74	370.82	173.00
Population (mill.)	38.30	142.53	0.01	1450.94	3774.00	22.17	16.96	86.67	114.29
Unemployment rate	7.71	5.94	0.10	37.32	3651.00	7.47	8.15	5.97	6.96
Employment in agriculture	25.35	21.75	0.08	90.61	3651.00	30.84	19.55	26.56	20.34
Employment in services	48.63	14.78	10.59	92.61	3651.00	46.21	51.80	47.74	51.40
Access to state business opportunities by socio-economic position	0.75	1.28	-2.22	3.80	3476.00	0.40	1.39	0.34	1.00

D Firm level data

Table D1: CCES survey questions

Question	Value
Percentage of sales through e-payments	0-100
Percentage of purchases through e-payments	0-100
As a percentage of a typical transaction, how much does it cost to accept payments, using [insert the e-payment method to receive a payment identified in BEE.K.6]?	0-100
As a percentage of a typical transaction, how much does it cost to accept payments, using [insert the e-payment method to receive a payment identified in BEE.K.6]?	0-100

Table D2: Descriptive statistics for World Bank Enterprise Survey data

Variable	Average	Std. dev.	Min	Max	N	Quartiles of D/L Q1	Q2	Q3	Q4
Data use intensity (D/L)	**3450.24**	**4723.08**	**53.50**	**14693**	**27285**	**149.33**	**347.13**	**2338.12**	**11285.70**
Index of cross-border data barriers	**0.33**	**0.24**	**0.00**	**1.00**	**24976**	**0.37**	**0.31**	**0.28**	**0.31**
Strict localization	0.33	0.25	0.00	1.00	24976	0.38	0.30	0.27	0.30
Ban on transfer and local processsing	0.32	0.32	0.00	1.00	24976	0.33	0.34	0.30	0.30
Local storage requirement	0.65	0.44	0.00	1.00	24976	0.56	0.70	0.75	0.68
Infraestructure requirement	0.20	0.40	0.00	1.00	24976	0.30	0.14	0.10	0.15
Conditional flow regime	0.33	0.47	0.00	1.00	24976	0.34	0.34	0.32	0.30
A. E-payments									
% received e-payments	68.89	30.25	0	100	24617	68.22	67.56	73.78	66.36
Cost of receiving e-payment (%)	1.58	2.43	0	100	18233	1.67	1.24	1.59	1.63
% of e-payments	70.08	32.47	0	100	23981	69.17	68.38	74.14	69.01
Cost of e-payment (%)	1.50	4.84	0	100	16538	1.62	1.53	1.46	1.34
B. Obstacles									
Transport	0.12	0.32	0	1	28250	0.11	0.15	0.10	0.12
Customs and trade regulations	0.09	0.28	0	1	26821	0.08	0.08	0.08	0.12
Informality	0.09	0.29	0	1	27473	0.08	0.10	0.09	0.10
Taxrates	0.22	0.42	0	1	28069	0.22	0.26	0.23	0.19
Politcal instability	0.17	0.37	0	1	27865	0.15	0.18	0.18	0.16
Corruption	0.12	0.33	0	1	27489	0.16	0.11	0.08	0.12
C. Opinion on institutions									
Customs and trade regulations	8.41	4.31	1	15	28480	8.25	8.21	8.66	8.50
Workforce	8.68	4.00	1	15	28480	8.81	8.70	8.62	8.55
Labor regulations	8.87	4.01	1	15	28480	8.92	9.02	8.74	8.84
Political instability	8.07	4.46	1	15	28480	8.03	8.08	8.09	8.08
E. Other characteristics									
No. competitors	75.86	1464.06	0	100000	10934	13.63	20.15	255.34	25.98
No. full time employess	19.74	455.44	1.00	30000	25417	16.20	52.55	15.27	8.07
% indirect exports	1.93	9.99	0.00	100	27693	2.10	0.64	2.27	2.22
% direct exports	3.16	13.38	0.00	100	27692	4.01	1.26	3.46	2.82
Experienced internet disruptions	0.23	0.42	0	1	25630	0.23	0.20	0.28	0.22
Hard to switch internet	3.02	0.94	1	4	24195	2.99	3.02	2.96	3.12

E Interaction effect is identified

Proof of Lemma 1. By the FWL theorem

$$\text{plim}\widehat{\beta}_3 = \beta_3 + \frac{Cov(\widetilde{D \times I}, \tilde{\varepsilon})}{Var(\widetilde{D \times I})}.$$

From here onward, $\tilde{D}$, $\tilde{I}$ and $\widetilde{D \times I}$ and $\tilde{\varepsilon}$ denote the residualized counterparts to D, I, $D \times I$ and ε.

Note

$$\begin{aligned}
Cov(\widetilde{D \times I}, \varepsilon) =& Cov\left(D \times I - \frac{Cov(D \times I, \tilde{I})}{Var(\tilde{I})} \cdot I, \varepsilon\right) \\
=& E\left(\left[D - \frac{Cov(D \times I, \tilde{I})}{Var(\tilde{I})}\right] \cdot I \cdot \varepsilon\right) \\
=& \left[E(D) - E(D) \cdot \left(\frac{E(I^2) - E(I)^2}{Var(I)}\right)\right] E(I \cdot \varepsilon) \\
=& 0.
\end{aligned}$$

Thus $\text{plim}\widehat{\beta}_3 = \beta_3$. □

Proof of Proposition 1. If f is linear such that $D = \omega_1 \chi + \omega_2 I + \nu$ with $E(\chi' \varepsilon) \neq 0$, then we can replace D in

$$Y = D\beta_1 + I\beta_2 + D \times I\beta_3 + \varepsilon.$$

to obtain

$$Y = \omega_1 \beta_1 \chi + (\omega_2 \beta_1 + \beta_2) I + \omega_1 \beta_3 \chi \times I + \omega_2 \beta_3 I^2 + (\beta_1 \nu + \varepsilon).$$

Next we aim to estimate

$$Y = D\gamma_1 + I\gamma_2 + D \times I\gamma_3 + I^2\gamma_4 + \vartheta$$

which we can re-write as

$$Y = \omega_1 \gamma_1 \chi + (\omega_2 \gamma_1 + \gamma_2) I + \omega_1 \gamma_3 \chi \times I + (\omega_2 \gamma_3 + \gamma_4) I^2 + (\gamma_1 \nu + \vartheta).$$

Since both Y is a function of the same variables, then it follows that the following equalities must be satisfied:

$$\begin{aligned}
\gamma_1 &= \omega_1 \beta_1 \\
\omega_2 \gamma_1 + \gamma_2 &= \omega_2 \beta_1 + \beta_2 \\
\omega_1 \gamma_3 &= \omega_1 \beta_3 \\
\omega_2 \beta_3 &= \omega_2 \gamma_3 + \gamma_4 \\
\beta_1 \nu + \varepsilon &= \gamma_1 \nu + \vartheta
\end{aligned}$$

Thus

$$\begin{aligned}
\gamma_1 &= \beta_1 \\
\gamma_2 &= \beta_2 \\
\gamma_3 &= \beta_3 \\
0 &= \gamma_4 \\
\varepsilon &= \vartheta
\end{aligned}$$

Therefore estimating

$$Y = D\gamma_1 + I\gamma_2 + D \times I\gamma_3 + I^2\gamma_4 + \vartheta$$

is equivalent to estimating

$$Y = D\beta_1 + I\beta_2 + D \times I\beta_3 + \varepsilon.$$

Therefore by Lemma 1, $\text{plim}\widehat{\gamma}_3 = \text{plim}\widehat{\beta}_3 = \beta_3$. □