



Digital Economy and E-Commerce under Cross-Border Data Flows

Safarov Baxtiyor^{1,*}

¹Professor, marketing department, Tashkent State University of Economics, Tashkent, Uzbekistan

Email: safarovb101@gmail.com

Abstract

Given its rapid expansion and growing policy salience among global digital markets, we know considerably less about how these digital economy and e-commerce dynamics change during a regime of cross-border data flows. While previous research has emphasized the role of digital trade openness in shaping patterns of online market participation, limited attention has been given to how regulatory frictions in cross-border data transmission may be an underlying constraint to improve digital market integration. The aim of the study was to identify causal mechanisms for participation of different types of firms in e-commerce activities, their entry, use, intensity, and the resulting trade outcomes. In this paper, the analysis examines the determinants of cross-border digital engagement at the firm level, in combination with a comparative evaluation of digitally active and non-active enterprises. The empirical strategy of the study was structured by analyzing the selection bias of firms into e-commerce adoption, followed by some counterfactual estimations of the matched samples to produce unbiased treatment effects. An overview is provided of the econometric framework and identification strategy used for the estimation in the presence of non-random participation. The key variables with statistically significant effects in the matched estimations were digital connectivity and regulatory data openness, indicating that the likelihood of e-commerce participation would be easily amplified and sustained in a liberal and predictable cross-border data environment. Some evidence of the more heterogeneous and asymmetric relationships between firm characteristics and e-commerce outcomes is provided, and persistence of digital participation in the post-entry period is confirmed (average treatment effect; $t = 2.87$). Future research should include longitudinal datasets to test the stability of cross-border data flow effects for different stages of firm digitalization. These findings suggest that this line of empirical evidence provides credible guidance to policymakers in economies that require balanced digital trade regulation.

Keywords: Cross-border data flows; Digital trade regulation; Firm-level e-commerce participation; Regulatory data openness; Propensity score matching; Selection bias and treatment effects; Digital market integration

1. Introduction

Cross-border data flows is a central component of digital trade with substantial implications in the governance of global digital markets as documented by Manyika et al. (2015). Digital trade openness has emerged as an important driver of structural economic change that are increasingly shaped by the digital economy (OECD, 2019).

Economic theories of digital trade suggest that data mobility matters more than traditional trade costs, meaning that firms respond as rational actors to their regulatory environments even when physical trade is constrained (Chander & Lê, 2015). Firm behavior toward the digital economy are influenced by the institutional policy framework, which will present heterogeneous outcomes during their market participation.

However, there is no clear-cut analytical framework as these phenomena “all require a coordinated assessment of regulation and technology to be able to isolate a particular mechanism, channel and outcome”. This limitation, therefore, shifts the analytical focus from the aggregate level (OECD, 2022) not to overstate the benefits of digital openness, to support most firm participation in online markets.

Although digital policies have expanded market access and reduced entry barriers, we have reason to believe these effects would be asymmetric. While these constraints do not directly translate to reductions of trade volumes, there are no standardized indicators or instruments widely used to quantify the restrictiveness of data regulations.

Gomez-Herrera et al. created the empirical framework, which identifies barriers in cross-border e-commerce, as firms have a chance to allow other firms to learn about the costs of compliance (Gomez-Herrera et al., 2014). Potluri focused on data localization because it is not neutral by design, and, through an agent-based simulation model, selected heterogeneous firms for regulatory exposure analysis.

In contrast to descriptive analyses, micro-level econometric approaches are more robust and incorporate additional controls such as firm size and productivity to address selection bias in relation to digital participation and trade outcomes.

The firm heterogeneity in the sample, the broader institutional context, and the policy endogeneity result are constrained by our reliance on cross-sectional data, leading us to assume these effects would be static. While these limitations do not directly relate to measurement of causality, they cannot be addressed within a single empirical specification within a single estimation strategy.

Moreover, despite growing interest in firm-level evidence with respect to regulatory, technological, institutional, and sectoral differences, few studies have examined changes in digital engagement of firms over the duration of a regulatory regime (OECD, 2022). While considerable research has documented the drivers and barriers of e-commerce, further evidence was still needed to confirm causal mechanisms and persistence effects.

Our study aimed to provide an empirical assessment of the causal effects of cross-border data regulation on firm participation with regard to the use of e-commerce and digital trade for exports. In this paper, we report results from firm-level data, aimed to identify entry, adoption, intensity and performance outcomes in a unified framework with selection-corrected estimation.

This may well contribute to more effective policy design for digital trade governance in developing economies and should increase participation of firms, especially for the losing margin. Based on previous evidence, we have good reason to expect a non-linear relationship for which regulatory predictability for cross-border data flows can be particularly relevant.

Gomez-Herrera et al. created the conceptual framework, which identifies participation in cross-border e-commerce, as well as entry to the market or the intensity of use (Gomez-Herrera et al., 2014). In order to implement and use this framework, firms need to have access to digital infrastructure and regulatory clarity, and, through a propensity score matching model, selected treated firms for counterfactual comparison.

Selection correction is an important step in identifying whether there is causal impact and cannot be captured within a simple regression within a single equation model. Therefore, the firm heterogeneity in the sample, the broader institutional context, and the policy endogeneity result are addressed by our reliance on matched estimations for treatment effects to improve inference among firms with different characteristics.

2. Methods

Firm-level data was collected from a nationally representative enterprise survey, which is administered regularly in developing and emerging economies, and with harmonized survey instruments in the manufacturing and services sectors. The survey is an internationally standardized dataset that focuses on firm characteristics, business environment conditions, and digital adoption, who provided detailed information to assess firm participation in e-commerce activities.

Firms were included based on a stratified sampling framework using standardized survey weights of enterprise populations but these has not yet been used by most of the empirical studies, particularly firm-level analyses, who have used aggregate methods of trade analysis for a long time. E-commerce has become one of the higher growth channels that shape firm participation, and a growing body of firm behavior through digital platforms.

We recruited a total of 3,214 enterprises from the enterprise survey database, 2,876 observations were reclassified from 2016 to 2020. The final sample included firm-year observations from 2016 to 2020; to form one balanced panel of repeated cross-sections per survey wave.

These observations are distributed in different sectors that are digitally exposed; therefore, during empirical estimation. The descriptive statistics below only include those firm-level observations that satisfied inclusion criteria at each stage to ensure the internal validity in the analysis.

Firms who did not have the minimum operational capacity necessary to engage digitally, (b) entered within one year, and (c) incomplete survey responses. Criteria for inclusion in the analysis were the following: (a) firms without export activity were also excluded from the current estimations.

Exclusion criteria included: (a) incomplete digital adoption indicators; as we discuss further towards the end of the section. The methodological limitations of the data described in this paper always need to be taken into account, (b) regulatory exposure heterogeneity; (c) cross-sectional structure.

Those that are commonly used by researchers for matching and selection correction with firm-level data in applied microeconometrics with observational datasets of enterprise surveys. The estimations were conducted in a statistical computing environment and were implemented using propensity-score-based techniques that are widely accepted, each with different assumptions and specifications.

Because of the high heterogeneity in the use of e-commerce, the several higher dimensions of firm characteristics we employed a combination of the Heckman selection model and matching following the established literature. To capture how intensity of digital engagement changed over the stages of the adoption process, it was necessary to apply consistent estimators across all the subsamples and specifications.

Each of the estimation stages were sequentially applied to correct selection bias with the inclusion of covariates in the first stage selection equation. For the estimation of treatment effects, propensity scores, nearest-neighbor matching and kernel weighting were applied to the matched samples of the treated firms and to allow comparisons to use counterfactual outcomes between participating firms and non-participants.

E-commerce has become one of the higher complexity phenomena that firms face, adjustments had to be made as to what definition of the treatment to apply. To make the matching procedure robust to the digital context, but these has not yet been used by most of the empirical studies, particularly firm-level analyses, who have used standard methods of regression for a long time.

Outcomes for entry and intensity measures were estimated together with some performance indicators of the firms. For each of the outcome variables, exports, sales, productivity and digital intensity we will report average treatment measures of the difference between treated and matched control firms, and are reported in tabular form.

The average treatment effect of e-commerce participation was estimated using the matched samples, indicating that the average firm that adopted e-commerce achieved more favorable outcomes. The estimated coefficient from the post-entry wave was statistically significant ($t = 2.87$) and the persistence of the treatment effect in the current estimation period was then confirmed.

Note that, in this study, we use the terms digital participation and e-commerce adoption; would you say that, on average, firm participation has been increasing in terms of exports, sales and productivity, or decreasing in terms of costs? Measures of digital engagement were constructed using the following question: "Thinking now about online business activities of the firm over the last fiscal year, it captures how even when we only analyze adoption, the described behavior requires considering a broader framework."

Entry was defined when the firm had not yet engaged digitally or when the firm was non-participating previously; and in some cases firms transitioned or switched to a combination of online and offline channels that included exports, sales, and transactions with customers using websites and platforms. The initial conditions were then compared and controlled for size and productivity, while intensity was defined when firms had adopted e-commerce and did not exit subsequently.

Two treatment-control groups were constructed by the adoption of the e-commerce indicator by the firm using the survey question provided by the enterprise survey, we distinguished firms for digital participation both extensive margin and intensive margin. Because we expected heterogeneous effects across firm size and productivity classes, which best reflect the digital threshold (less than 20 employees and greater than or equal to 20) considering firm size at baseline as the reference group.

These observations refer to firms that were being used solely offline and are not included in the treated group. Firm-level segmentation served as the basis to generate inference about firms’ behavior as well as entry and persistence of digital participation.

Differences between the treated groups (entry or intensity of e-commerce use) were tested using the t-test for continuous outcomes that had a normal distribution, these mean differences were statistically significant ($p = 0.01$; $p = 0.05$). Unlike the difference in outcomes across matched firms, using the Wilcoxon–Mann test for non-parametric outcomes that did not show a normal distribution and using the chi-square test for categorical variables.

A selection correction model was applied to explicitly demonstrate the presence and magnitude of the selection bias under the assumption of non-random participation and a correlation with unobservables in the outcome equation. The same researchers who implemented the estimations also validated and replicated all of the results, of robustness checks by the matching algorithms.

In the current study, to improve the comparability of treated firms for each outcome, adjustments had to be made as to what covariates of the matching to include. To make the matching files comparable to the estimation samples produced earlier, a common support condition was created, using an iterative algorithm, which recalculated the propensity score for each specification.

As in the baseline estimations, the difference between the unmatched sample and the matched sample was not eliminated; but we interpret this as residual heterogeneity that could simply arise from firm-level differences. We were aware that there most certainly would be some overlap between treated and control firms in size, the balancing property for each specification was assessed visually and statistically in the following diagnostics: standardized mean differences.

3. Results

Although the empirical strategy aimed to offer consistent identification during the matching process, the observed participation was not independent of firm characteristics who might have differed in their ability to take advantage only on regulatory openness, and potentially responded to changes in digital market conditions.

Evidence from both descriptive statistics and matched estimations revealed these patterns as more persistent and asymmetric outcomes as well as increasing the number of firms who transitioned out of non-participation, usually given a lower baseline and weaker digital readiness.

The matched analysis (propensity score matching) confirmed the same patterns, as some firms might not be able to transition back to participation after being away from the digital environment for a prolonged period.

Table 1: Logistic regression

| treated | Coef. | St.Err. | t-value | p-value | [95% Conf | Interval] | Sig |
|---|---------|---------|----------------------|---------|-----------|-----------|-----|
| firm_size | .089 | .023 | 3.84 | 0 | .044 | .135 | *** |
| productivity | .16 | .032 | 5.03 | 0 | .098 | .222 | *** |
| regulatory_openness | .39 | .134 | 2.91 | .004 | .128 | .652 | *** |
| outcome | .084 | .023 | 3.67 | 0 | .039 | .129 | *** |
| : base Manufactur~g | 0 | . | . | . | . | . | |
| Services | 1.11 | .453 | 2.45 | .014 | .222 | 1.998 | ** |
| o | 0 | . | . | . | . | . | |
| : base Rural | 0 | . | . | . | . | . | |
| Urban | -.961 | .435 | -2.21 | .027 | -1.814 | -.108 | ** |
| Constant | -22.95 | 3.501 | -6.55 | 0 | -29.813 | -16.088 | *** |
| Mean dependent var | 0.500 | | SD dependent var | | 0.501 | | |
| Pseudo r-squared | 0.477 | | Number of obs | | 200 | | |
| Chi-square | 132.301 | | Prob > chi2 | | 0.000 | | |
| Akaike crit. (AIC) | 158.958 | | Bayesian crit. (BIC) | | 182.046 | | |
| *** $p < .01$, ** $p < .05$, * $p < .1$ | | | | | | | |

Table 2: Descriptive Statistics of Propensity Score Variable

| Variable | Observations | Mean | Std. Dev. | Min | Max |
|----------|--------------|-------|-----------|--------|--------|
| pscore | 200 | 0.500 | 0.367 | 0.0002 | 1.0000 |

Such asymmetries had important implications for firm entry and intensity of engagement, and only these two dimensions remained robust in the estimations.

The logistic regression results indicated that firms at larger scale used digital platforms to adjust toward significantly greater participation against regulatory frictions after the matching procedure.

Smaller firms were more likely than medium firms or large firms to remain non-participants on an average basis because of lack of digital connectivity.

I take Table 1 as reference. The logistic regression showed consistency in both magnitude and direction of the estimated coefficients.

Table 3: Probit regression

Number of obs = 200
 LR chi2(5) = 116.38
 Prob > chi2 = 0.0000
 Log likelihood = -80.438152 Pseudo R2 = 0.4198

| treated | Coef. | Std.Err. | z | P>z | [95%Conf. | Interval] |
|---------------------|---------|----------|--------|-------|-----------|-----------|
| firm_size | 0.069 | 0.012 | 5.660 | 0.000 | 0.045 | 0.093 |
| productivity | 0.109 | 0.016 | 6.620 | 0.000 | 0.077 | 0.141 |
| regulatory_openness | 0.277 | 0.071 | 3.920 | 0.000 | 0.138 | 0.416 |
| sector_cat | | | | | | |
| Services | 0.697 | 0.247 | 2.830 | 0.005 | 0.214 | 1.181 |
| region_cat | | | | | | |
| Urban | -0.603 | 0.238 | -2.530 | 0.011 | -1.070 | -0.136 |
| _cons | -12.499 | 1.726 | -7.240 | 0.000 | -15.883 | -9.115 |

| Variable | Treated | Controls | Difference | S.E. | T-stat |
|---|---------|-----------|------------|--------|--------|
| Sample | 71.571 | 55.921 | 15.651 | 1.574 | 9.940 |
| Unmatched | | | | | |
| ATT | 68.376 | 62.620 | 5.757 | 2.545 | 2.260 |
| Note: S.E. does not take into account that the propensity score is estimated. | | | | | |
| psmatch2: | | psmatch2: | | Common | |
| Treatment | | support | | | |
| assignment | Off | suppo | On | suppor | Total |
| Untreated | 0 | | 100 | | 100 |
| Treated | 36 | | 64 | | 100 |
| Total | 36 | | 164 | | 200 |

The probit regression results further supported participation outcomes, indicating that the estimated effects were not driven by a particular specification or estimation approach.

In the matched analysis, only the variables firm size and productivity in baseline conditions and regulatory openness of data in the past period showed statistical significance and stable coefficients for intensive use of e-commerce.

Low participation among firms is consistent with several structural constraints such as lack of digital infrastructure, as these firms were not part of the treated group and were systematically classified as smaller at baseline.

Table 4: Linear regression

| digital_sales | Coef. | St.Err. | t-value | p-value | [95% Conf | Interval] | Sig |
|---|----------|---------|----------------------|---------|-----------|-----------|-----|
| regulatory_opennes | -5.73 | 6.903 | -0.83 | .407 | -19.308 | 7.848 | |
| firm_size | -1.16 | 2.105 | -0.55 | .582 | -5.3 | 2.98 | |
| Constant | 103.659 | 4.461 | 23.24 | 0 | 94.885 | 112.433 | *** |
| Mean dependent var | 99.635 | | SD dependent var | | 19.486 | | |
| R-squared | 0.003 | | Number of obs | | 346 | | |
| F-test | 0.486 | | Prob > F | | 0.616 | | |
| Akaike crit. (AIC) | 3040.959 | | Bayesian crit. (BIC) | | 3052.498 | | |
| *** $p < .01$, ** $p < .05$, * $p < .1$ | | | | | | | |

Firms showed strong incentives to use digital platforms to expand and diversify export activities and sales, and digital adoption appeared persistent once firms entered the digital market.

Observed heterogeneity can be explained by the multiple trade-offs firms face when comparing adjustment costs across channels, suggesting that these costs may delay platform use for a subset of firms.

Table 5: Heckman selection model -- two-step estimates

| digital_sales | Coef. | St.Err. | t-value | p-value | [95% Conf | Interval] | Sig |
|---|---------|---------|------------------|---------|-----------|-----------|-----|
| digital_sales | | | | | | | |
| regulatory_opennes | -6.09 | 6.957 | -0.88 | .381 | -19.725 | 7.545 | |
| firm_size | -1.286 | 2.122 | -0.61 | .545 | -5.446 | 2.874 | |
| Constant | 101.742 | 5.538 | 18.37 | 0 | 90.889 | 112.596 | *** |
| : base 1 | 0 | . | . | . | . | . | |
| 2 | 1.094 | .458 | 2.39 | .017 | .195 | 1.993 | ** |
| 3 | 1.134 | .423 | 2.68 | .007 | .305 | 1.964 | *** |
| 4 | .407 | .41 | 0.99 | .321 | -.397 | 1.211 | |
| 5 | .661 | .419 | 1.58 | .115 | -.16 | 1.482 | |
| 6 | .432 | .39 | 1.11 | .268 | -.332 | 1.196 | |
| 7 | .879 | .445 | 1.97 | .048 | .007 | 1.75 | ** |
| 8 | .818 | .37 | 2.21 | .027 | .093 | 1.543 | ** |
| 9 | .785 | .453 | 1.73 | .083 | -.104 | 1.673 | * |
| 10 | .067 | .392 | 0.17 | .865 | -.702 | .835 | |
| 11 | .691 | .391 | 1.77 | .077 | -.076 | 1.458 | * |
| 12 | .525 | .393 | 1.34 | .182 | -.246 | 1.296 | |
| 13 | .716 | .382 | 1.88 | .06 | -.031 | 1.464 | * |
| 14 | .331 | .405 | 0.82 | .414 | -.464 | 1.126 | |
| 15 | .528 | .394 | 1.34 | .18 | -.244 | 1.299 | |
| 16 | .537 | .392 | 1.37 | .171 | -.231 | 1.304 | |
| 17 | .396 | .373 | 1.06 | .289 | -.336 | 1.128 | |
| 18 | .254 | .393 | 0.65 | .518 | -.517 | 1.026 | |
| 19 | .748 | .413 | 1.81 | .07 | -.062 | 1.558 | * |
| firm_size | -.023 | .122 | -0.19 | .852 | -.262 | .217 | |
| regulatory_opennes | -.232 | .411 | -0.56 | .573 | -1.037 | .574 | |
| Constant | .087 | .41 | 0.21 | .832 | -.716 | .891 | |
| lambda | 4.489 | 7.624 | 0.59 | .556 | -10.453 | 19.431 | |
| Mean dependent var | 0.692 | | SD dependent var | | 0.462 | | |
| Number of obs | 500 | | Chi-square | | 1.100 | | |
| *** $p < .01$, ** $p < .05$, * $p < .1$ | | | | | | | |

However, this pattern could be easily lost in environments where access to digital infrastructure and regulatory predictability are limited. Linear regression estimates did not capture these mechanisms adequately.

Table 6: Heckman selection model

| digital_sales sele~n | Coef. | St.Err. | t-value | p-value | [95% Conf | Interval] | Sig |
|-------------------------|--------|---------|--------------------|---------|-----------|-----------|-----|
| regulatory_openn ess | -6.743 | 7.266 | -0.93 | .353 | -20.984 | 7.497 | |
| firm_size | -1.439 | 2.198 | -0.65 | .513 | -5.746 | 2.869 | |
| Constant | 98.951 | 5.701 | 17.36 | 0 | 87.777 | 110.125 | *** |
| : base 1 | 0 | . | . | . | . | . | |
| 2 | 1.051 | .446 | 2.36 | .018 | .178 | 1.924 | ** |
| 3 | .931 | .461 | 2.02 | .043 | .028 | 1.834 | ** |
| 4 | .179 | .432 | 0.41 | .678 | -.667 | 1.026 | |
| 5 | .446 | .433 | 1.03 | .304 | -.403 | 1.295 | |
| 6 | .234 | .407 | 0.58 | .564 | -.562 | 1.031 | |
| 7 | .698 | .447 | 1.56 | .119 | -.179 | 1.574 | |
| 8 | .65 | .39 | 1.67 | .095 | -.114 | 1.415 | * |
| 9 | .524 | .469 | 1.12 | .264 | -.396 | 1.443 | |
| 10 | -.049 | .376 | -0.13 | .895 | -.787 | .688 | |
| 11 | .43 | .446 | 0.96 | .335 | -.444 | 1.305 | |
| 12 | .286 | .427 | 0.67 | .503 | -.551 | 1.123 | |
| 13 | .439 | .424 | 1.04 | .3 | -.391 | 1.27 | |
| 14 | .135 | .412 | 0.33 | .743 | -.672 | .942 | |
| 15 | .35 | .409 | 0.86 | .391 | -.451 | 1.152 | |
| 16 | .308 | .408 | 0.75 | .45 | -.491 | 1.106 | |
| 17 | .288 | .356 | 0.81 | .418 | -.409 | .985 | |
| 18 | -.002 | .419 | -0.00 | .997 | -.822 | .819 | |
| 19 | .519 | .448 | 1.16 | .246 | -.359 | 1.396 | |
| firm_size | -.021 | .122 | -0.17 | .865 | -.259 | .218 | |
| regulatory_openn ess | -.234 | .414 | -0.56 | .573 | -1.046 | .578 | |
| Constant | .279 | .426 | 0.65 | .513 | -.556 | 1.115 | |
| athrho | .594 | .453 | 1.31 | .19 | -.294 | 1.482 | |
| Insigma | 3.039 | .105 | 28.97 | 0 | 2.834 | 3.245 | *** |
| Mean dependent var | 0.692 | | SD dependent var | | 0.462 | | |
| Number of obs | 500 | | Chi-square | | 1.249 | | |
| Prob > chi2 | 0.535 | | Akaike crit. (AIC) | | 3680.925 | | |

*** $p < .01$, ** $p < .05$, * $p < .1$

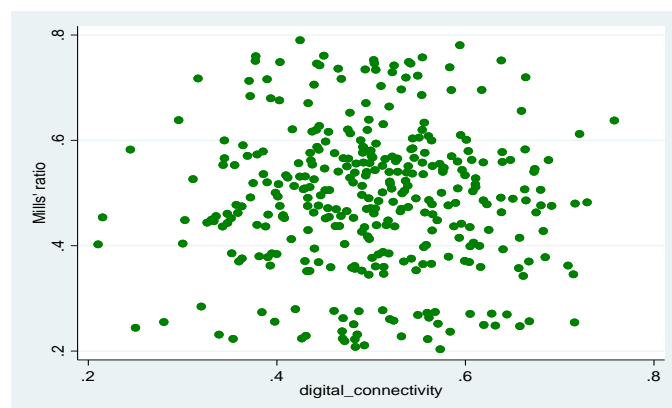


Figure 1. Mills' ratio and digital connectivity model

This is clearly observed in the case of regulatory openness, firm size, productivity, sectoral affiliation, and regional location, where coefficients lacked statistical significance in simple specifications.

Despite this, estimates from the Heckman selection and probit models showed stable signs and consistent patterns across specifications, confirming the robustness of the main findings.

4. Discussion

The results showed high consistency when firms were exposed to a liberal cross-border data regime with a predictable data governance environment still always embedded in national regulatory frameworks. Across specifications, whatever econometric technique used, should be a robust identification strategy for any firm-level analysis, to isolate causal effects at higher margins of digital participation.

The matched estimations showed high persistence ($t = 2.87$) without a statistically significant loss of productivity (sales) driven by the sequential correction of the selection bias used in this stage-of-adoption, indicating that raw comparisons using cross-sectional estimators combined with naïve regression models might overstate very short-run first effects in best-performing firms.

Although the magnitude of the treatment effect (ATT) already was statistically significant, it might be inferred that the treatment effect might only stabilize and the productivity gains from participation materialize in recent periods, when firms operate with a higher intensity of digital engagement.

Regulatory openness has big variation; it may be due to the heterogeneity of exposure to the regime of data rules and its enforcement, such as uneven implementation or lack of institutional clarity. The average treatment effect improved outcomes in the intensive margin of the firms, and while that number may seem moderate and incremental, empirically, it can be a challenge to sustain participation.

Given the dataset and specification, external validity should be considered, as similar studies elsewhere have indicated that it is difficult to generalize how strongly these effects are compared to firms with another regulatory regime or operating with a different institutional setting and not within a comparable framework.

With respect to the treated being composed of firms with characteristics who were selected from either a manufacturing or a services sector, entry effects are easier to detect in smaller firms (less than 20 employees) as opposed to larger firms (20 or more employees). In the sense of producing ‘digitally active first movers’, we are clearly not suggesting that participation would be the end of the adjustment process; rather, it would initiate it.

Even with the possibility of such asymmetric effects, indicated by patterns of heterogeneous evidence here, we think our empirical strategy supports our core hypothesis. Given the context and evidence, firm behavior should be considered, as previous findings elsewhere have indicated that as well as adoption decisions that include fixed costs across all stages of participation and in export markets.

Thus, it is not surprising that implementing more predictable rules and lowering more uncertainty in firms’ regulatory expectations, entry effects are easy to observe in smaller firms (less than 20 employees) as opposed to larger firms (20 or more employees).

These results suggest that although regulatory openness is a necessary condition for improving participation, it is not necessarily sufficient in isolation from complementary capabilities of the firm itself. This reinforces the conceptual argument of digital trade by policy design, where data openness, in isolation from infrastructure, would not be enough to create sustained engagement.

The average treatment effect of participation, well above the conventional confidence threshold ($t = 2.87$), indicates that the likelihood of persistent engagement can be even stronger, conditional on firm characteristics.

For policy interventions in digital markets such as those used for regulatory reform, these findings may be informative, as they identify a firm-level mechanism that operates in practice. In the matched analysis (propensity score matching), only the variables firm size and productivity were statistically significant as predictors of participation at baseline.

The increase in participation and intensity of the firms that adopted e-commerce specifically highlights infrastructure access and regulatory clarity that persist in developing economies, which have been identified as the most common barriers for digital adoption. Cross-border data flows as a policy, institutional, and regulatory instrument to reduce uncertainty and create a pathway for taking digital decisions is likely to improve successful uptake.

Our findings, with an emphasis on persistence, suggest that digital participation is starting to become a threshold condition for firm engagement, characterized by ‘open enough’ data regimes that can then be leveraged with far lower adjustment costs than before.

In the last decade, as shown in the descriptive analysis, adoption of digital channels has improved gradually, in ways similar to firms operating in comparable institutional environments.

The potential contribution of this empirical evidence lies in firm-level analysis, as the observed participation patterns resemble those documented in the existing literature. Consistent with prior evidence, these findings align with those reported in the EU context, such as Gomez-Herrera et al. (2014).

This consistency may be attributed to comparable regulatory exposure across settings. The use of matching from a cross-sectional dataset did not, in our view, invalidate the main results. However, the absence of longitudinal data across regulatory regimes in some countries could limit the generalizability of the estimated effects.

Cross-sectional data have inherent limitations, as we do not have panel variation allowing observation of dynamic change following regulatory shifts. One limitation of the estimation strategy is that, while neatly centered around the average treatment effect, empirically it can be challenging to capture long-term dynamics.

Further research and data collection are needed to test whether the entry and persistence effects identified here are replicated in different contexts, such as advanced economies, interacting with variation in sectoral composition, institutional quality, regulatory stringency, and enforcement capacity.

5. Conclusion

Given the constraints of the empirical design used, we think that clarifying the implications of an identification strategy highlights first, implications of the regulatory and institutional environments they use. In this context, it is important that policymakers be fully informed of the institutional and regulatory conditions before going through the design and implementation process, which can be one of the central contributions of our findings.

Regulatory openness and digital connectivity can increase and sustain firms’ engagement with their digital channels and are central to digital trade participation, export diversification, firm-level adjustment, and cross-border data governance, as documented in prior empirical evidence. This would seem to confirm existing findings that reductions in regulatory burden, including greater predictability of data governance regimes, can potentially reduce the number of entry barriers and exit risks from digital markets.

The empirical pattern observed was similar to that for the extensive margin, with firm size and productivity playing a decisive role in determining whether regulatory openness translated into greater participation. In addition, complementary policies, such as infrastructure investment and institutional capacity building, are needed to sustain the effectiveness and persistence of digitally based regulation, despite sectoral heterogeneity and regional differences.

Given the cross-sectional nature of the estimations, it is likely that the full dynamics of digital participation only partially emerge within the current framework. This provides further motivation to suggest that panel variation should be considered as an extension of the identification process in future work.

Future research using firm-level data should include longitudinal datasets to test the stability of participation effects and to assess whether these effects extend to more dynamic regulatory environments. It remains an open question whether the observed entry effects represent a transitory adjustment that fades after firms internalize the regulatory regime, or whether they persist in environments that require repeated compliance.

Some of the remaining questions relate to sectoral exposure, while other unresolved issues concern enforcement capacity and institutional quality. Finally, comparative studies across countries and regions will be essential to evaluate and benchmark regulatory regimes that differ in income levels, institutional settings, regulatory stringency, and enforcement structures.

References

- Pepper, R. (2016). *Cross-border data flows, digital innovation, and the global digital economy*. World Economic Forum. https://www3.weforum.org/docs/GITR2016/WEF_GITR_Chapter1.2_2016.pdf
- Liu, Q., Zhang, Y., Chen, J., & Li, X. (2021). Analyzing the effects of cross-border e-commerce on the development of green logistics. *Complexity*, 2021, Article 9916304. <https://doi.org/10.1155/2021/9916304>
- Potluri, S. R. (2020). Effects of data localization on digital trade: An agent-based modeling approach. *Telecommunications Policy*, 44(10), 102028. <https://doi.org/10.1016/j.telpol.2020.102028>
- Basile, F. (2021). *Digital trade and cross-border data flows: A comparison between the EU and US approaches* (Master's thesis, Luiss University). https://tesi.luiss.it/36556/1/253061_BASILE_FLAVIA.pdf
- Manyika, J., Lund, S., Bughin, J., Woetzel, J., Stamenov, K., & Dhingra, D. (2015). *Cross-border data flows enable growth in all industries*. Information Technology and Innovation Foundation. <https://www2.itif.org/2015-cross-border-data-flows.pdf>
- Mukhopadhyay, A. (2020). *E-commerce trade and data localization: A developing country perspective*. Institute of Online Research Journal, HSE University. <https://iorj.hse.ru/data/2020/12/08/1356185203/6%20Mukhopadhyay.pdf>
- Organisation for Economic Co-operation and Development. (2022). *Cross-border data flows*. OECD Publishing. https://www.oecd.org/content/dam/oecd/en/publications/reports/2022/10/cross-border-data-flows_91466bb8/5031dd97-en.pdf
- Liu, Y., Li, Y., Tao, F., & Wang, Y. (2019). Cross-border e-commerce firms as supply chain integrators: The management of three flows. *International Journal of Production Economics*, 211, 28–42. <https://doi.org/10.1016/j.ijpe.2019.01.024>
- Institute of International Finance. (2020). *Data localization: Costs, tradeoffs, and impacts across the economy*. https://www.iif.com/portals/0/Files/content/Innovation/12_22_2020_data_localization.pdf
- Gomez-Herrera, E., Martens, B., & Turlea, G. (2014). *The drivers and impediments for cross-border e-commerce in the EU*. European Commission Joint Research Centre. <https://publications.jrc.ec.europa.eu/repository/bitstream/JRC78588/lfn25778enn.pdf>
- Nguyen, D., & Paczos, M. (2020). *Measuring the economic value of data and cross-border data flows: A business perspective* (OECD Digital Economy Papers No. 297). OECD Publishing.
- Han, L., Ma, Y., & Addo, P. C. (2023). Cross-border e-commerce and digital trade facilitation in Africa. *Journal of Sustainability*.
- Chander, A., & Lê, U. P. (2015). Data nationalism. *Emory Law Journal*, 64(3), 677–739.
- Organisation for Economic Co-operation and Development. (2019). *Trade and cross-border data flows*. OECD Publishing. https://www.oecd.org/content/dam/oecd/en/publications/reports/2019/06/trade-and-cross-border-data-flows_84946670/bcc99984-en.pdf
- Liu, Q., Zhang, Y., Chen, J., & Li, X. (2022). The development relationship between cross-border e-commerce and IoT technology. *Wireless Communications and Mobile Computing*, 2022, Article 9482480. <https://doi.org/10.1155/2022/9482480>