

Thang, Nguyen-Quoc

**Working Paper**

## Does internet expansion matter for export concentration?

ARTNeT Working Paper Series, No. 238

**Provided in Cooperation with:**

Asia-Pacific Research and Training Network on Trade (ARTNeT), Bangkok

*Suggested Citation:* Thang, Nguyen-Quoc (2024) : Does internet expansion matter for export concentration?, ARTNeT Working Paper Series, No. 238, Asia-Pacific Research and Training Network on Trade (ARTNeT), Bangkok, <https://hdl.handle.net/20.500.12870/6807>

This Version is available at:

<https://hdl.handle.net/10419/301659>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

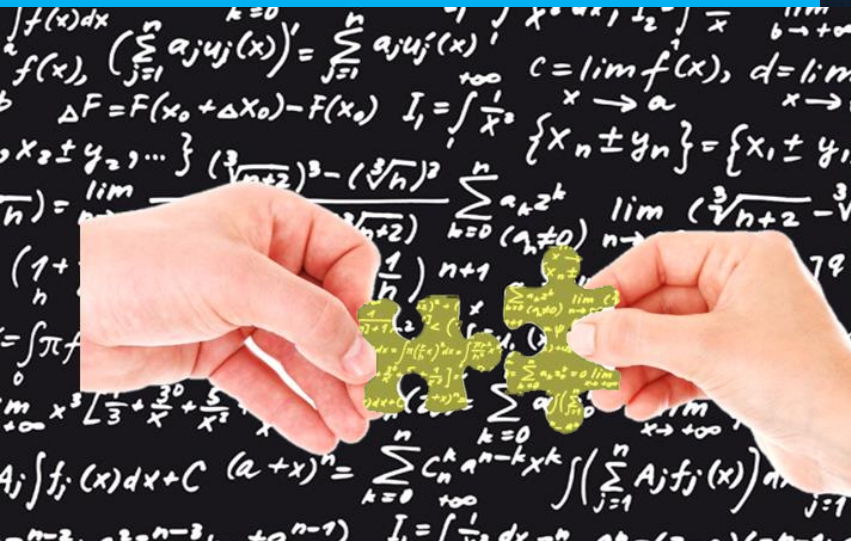
*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



**ESCAP**  
Economic and Social Commission  
for Asia and the Pacific

# Does Internet Expansion Matter For Export Concentration?



**Thang Nguyen-Quoc**

ASIA-PACIFIC RESEARCH AND TRAINING NETWORK ON TRADE

# Working Paper

No. 238 | 2024

The Asia-Pacific Research and Training Network on Trade (ARTNeT) is an open regional network of research and academic institutions specializing in international trade policy and facilitation issues. ESCAP, WTO and UNCTAD, as key core network partners, and a number of bilateral development partners, provide substantive and/or financial support to the network. The Trade, Investment and Innovation Division of ESCAP, the regional branch of the United Nations for Asia and the Pacific, provides the Secretariat of the network and a direct regional link to trade policymakers and other international organizations.

The ARTNeT Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about trade issues. An objective of the series is to publish the findings quickly, even if the presentations are less than fully polished. ARTNeT Working Papers are available online at <http://artnet.unescap.org>. All material in the Working Papers may be freely quoted or reprinted, but acknowledgment is requested together with a copy of the publication containing the quotation or reprint. The use of the Working Papers for any commercial purpose, including resale, is prohibited.

**Disclaimer:**

The designations employed and the presentation of the material in this Working Paper do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries. Where the designation “country or area” appears, it covers countries, territories, cities or areas. Bibliographical and other references have, wherever possible, been verified. The United Nations bears no responsibility for the availability or functioning of URLs. The views expressed in this publication are those of the author(s) and do not necessarily reflect the views of the United Nations. The opinions, figures and estimates set forth in this publication are the responsibility of the author(s) and should not necessarily be considered as reflecting the views or carrying the endorsement of the United Nations. Any errors are the responsibility of the author(s). The mention of firm names and commercial products does not imply the endorsement of the United Nations.



ASIA-PACIFIC RESEARCH AND TRAINING NETWORK ON TRADE

# WORKING PAPER

## Does Internet Expansion Matter For Export Concentration?

Thang Nguyen-Quoc<sup>1</sup>

### Please cite this paper as:

Thang Nguyen-Quoc (2024). "Does Internet Expansion Matter For Export Concentration?", **ARTNeT Working Paper Series** No. 238, February 2024, Bangkok, ESCAP.

Available at <http://artnet.unescap.org>

### Acknowledgement:

This paper is a submission for the 2023 ESCAP-UNCTAD-UNIDO-ARTNeT call for papers, under the theme "Unleashing Digital Trade and Investment for Sustainable Development." The author is grateful for the valuable feedback and support from the peer reviewers at the ESCAP Secretariat, whose constructive input was pivotal in refining and finalizing this work.

---

<sup>1</sup> Université Paris-Dauphine, email: [tngquyenvn@gmail.com](mailto:tngquyenvn@gmail.com)

## Abstract

Why do export sales increasingly concentrate among a few “superstars”? This paper is the first to argue that the expansion of the Internet matters for this phenomenon. Using firm-level customs transaction data from 11 developing countries, we show that the spread of the internet steepens the slope of the rank-sales relationship among exporters. This shift implies a more skewed distribution of sales toward the top exporters. Furthermore, incumbent superstar exporters are more likely to retain their status owing to a better Internet. Doubling internet adoption is expected to boost the chance for the incumbent top exporter to retain its top position by 11.7%. This boost suggests an “entrenchment” effect where the internet reduces churning at the top. For both findings, we find some evidence that the effects are stronger for exports of differentiated goods than for exports of homogenous goods, suggesting the role of information friction in explaining the results. These findings suggest that the expansion of the Internet, and the broader digitalisation process, is a contributing factor to the increasing concentration of exports among a few “superstars”. As digitalisation continues, macroeconomic policymakers will need to step up their monitoring of idiosyncratic shocks related to these “superstars”. Policymakers also need to adapt their trade and development strategy in light of the shifting competitiveness landscape created by digitalisation. As the competition structure in the internet era favours superstars, the focus of export-promotion strategies may need to shift from supporting export entrepreneurship to enhancing the linkages between export superstars and domestic suppliers.

**Keywords:** firm size distribution, export superstars, internet, digitalisation

**JEL Codes:** D22, L25, L86, F14

# Table of Contents

<b>Abstract</b> .....	<b>iv</b>
<b>1. Introduction</b> .....	<b>1</b>
<b>2. Data</b> .....	<b>3</b>
2.1 Data on exporters.....	3
2.2 Data on Internet use.....	5
<b>3. Internet and distribution of exporter sales</b> .....	<b>6</b>
3.1 Hypothesis .....	6
3.2 Empirical setup .....	7
3.3 Results.....	9
3.4 Robustness.....	10
<b>4. Internet and entrenchment of export superstars</b> .....	<b>14</b>
4.1 Hypothesis .....	14
4.2 Empirical setup .....	15
4.3 Results.....	16
4.4 Robustness.....	17
<b>5. Mechanisms</b> .....	<b>20</b>
5.1 Information friction and sales concentration.....	21
5.2 Other channels .....	24
<b>6. Conclusions and Policy Implications</b> .....	<b>25</b>
<b>List of references</b> .....	<b>27</b>
<b>Appendix: Supplementary Materials</b> .....	<b>31</b>

## List of tables

Table 1: Summary statistics for log-log sale-rank regressions.....	8
Table 2: The effect of internet on log rank-log sale relationship .....	9
Table 3: Internet and log-log sale-rank relationship, IV approach .....	12
Table 4: Effect of internet on rank-sale exponent among the top 10%, covariates and sample robustness.....	13
Table 5: Effect of internet on export superstars' persistence .....	16
Table 6: Effect of internet on superstars, IV approach .....	18
Table 7: Effect of internet on superstars, with alternative samples.....	19
Table 8: Effect on log-log ranks-sales relationship, by type of goods.....	23
Table 9: Effect on superstar's persistence, by type of goods.....	24

## List of figures

Figure 1: Internet adoption level in developing countries .....	4
Figure 2: Share of export superstars in total export in developing countries.....	4
Figure 3: Distribution of sales among Peruvian exporters to Germany .....	6

# 1. Introduction

Export sales in each country concentrate among a handful of “export superstars”. Mayer and Ottaviano (2008) find that the top 1% of firms account for 32-77% of national manufacturing exports in seven European countries. Export “superstars” such as Samsung in Vietnam and Taiwan Province of China Semiconductor Manufacturing Conductor (TSMC) in Taiwan Province of China also dominate the export basket and growth in Asia. Freund and Pierola (2015) analysed custom transaction data from 32 countries to show that the top five firms account for 30% of non-oil exports. In a follow-up study, they found that these superstars account for almost half of the export growth as well (Freund and Pierola, 2020). After reviewing existing literature, they conclude that the causes behind this rising export concentration remain an open question for research.

Understanding the drivers of export concentration is critical from both academic and policy perspectives. Recent studies demonstrate the non-negligible consequence of these “superstars” on aggregate economic behaviour and volatility (Gabaix, 2011; Di Giovanni, Levchenko, and Méjean, 2014), comparative advantage (Gaubert and Itskhoki, 2021), and mark-ups (De Loecker, Eeckhout and Unger, 2020). Consequently, it has critical implications for the ability of Asian governments to realise SDGs by supporting a diversified and economically sustainable development pathway. The heterogeneity of firms is also central to estimating trade elasticities and the gains from trade (Arkolakis, Costinot and Rodríguez-Claire, 2012). In the context of COVID-19, understanding firm concentration becomes an important policy objective to avoid excessive exposure to idiosyncratic shocks and to avoid supply chain bottlenecks (WTO, 2021).

This paper provides empirical evidence that the global expansion of the internet contributes to the concentration of exports among the “superstars”. To our knowledge, this paper is the first to explicitly assess the link between communication technologies and export concentration in international trade. After describing the data in section 2, sections 3 and 4 will apply panel econometric techniques to firm-level customs registries from 11 developing countries to show that:

- 1) First, the spread of the internet steepens the slope of the log-log ranks-sales relationship among exporters in each origin-destination pair. This shift implies a more skewed distribution of export sales toward the top exporters.
- 2) Second, incumbent export superstars are more likely to retain their superstar status when internet adoption increases. The main specification suggests that doubling internet adoption would boost the chance for the incumbent top exporter to retain its position by 22.7% and the incumbent top 5% of exporters by 11.7%. This increase suggests an “entrenchment” effect where the internet reduces churning at the top.



The idea that internet adoption may increase export concentration is not immediately intuitive. Past research has often emphasised the impact of the Internet in cutting trade costs (Freund and Weinhold 2002, 2004). These studies focus especially on the role of the Internet in reducing the fixed cost of searching and matching between partners. Extending this logic under a heterogeneous firm model of trade such as Chaney (2008) would not generate changes in export sales distribution, however. On one hand, such a reduction in fixed cost would indeed induce new entry and typically lead to an increase in the number of exporters. On the other hand, the shape of the Pareto distribution that describes the heterogeneity of export sales does not change when the productivity cut-off for exporting drops.

Section 5 reconciles these findings from the lens of information frictions. Past research in the cross-country strand often assumes that the induced cost from frictions is already factored in the transaction as buyers and sellers agree on a common price and quantity. In reality, firms often do not find out about the quality (and profitability) of the match until after the trade has started, sometimes much later when disagreement occurs. Recent research such as Allen (2014) and Steinwender (2018) highlight the importance of informational frictions in trade that cannot be captured as part of the fixed cost. This problem is arguably more severe in the context of international trade and developing countries, where buyers and sellers are far away from each other and the mechanisms for information flow and traditional marketing are underdeveloped.

Connectivity can alter the level of competition faced by exporters by exposing each importer to many more potential exporters. As buyers can now easily find substitutes, demand becomes more elastic to prices – an argument dating back to Hirschman's first rule of derived demand. Importers are now more likely to source from sellers with the lowest production cost. High-productivity exporters are thus better placed than low-productivity ones to boost their sales from this heightened visibility and expand their market share. Consequently, the expansion of the Internet is expected to increase export concentration among the superstars.

In an extension of our empirical models, we find some evidence that the impact of the Internet on export concentration is stronger for exports of differentiated goods than for exports of homogenous goods. Typically, information frictions are more prevalent for differentiated goods than for those traded on commodity trading platforms or those with referenced prices. Consequently, the impact of internet adoption on alleviating information frictions would be higher for differentiated goods. These results do not refute the possibility of alternative mechanisms such as productivity, access to inputs and organisation capital. However, our main results remain robust even when controlling for various supply- and demand-side factors.

The paper relates to the larger literature on exporter heterogeneity. It provides a multi-country perspective to the empirical literature that often dissects the anatomy and distribution of bilateral export sales from a single source country such as Head et al.

(2014) and Nigai (2017). Our findings extend the research on the sources of export heterogeneity, which so far have largely overlooked the matching of buyers and sellers. An exception is Bernard et al. (2022), whose findings confirm our conjecture on the role of matching frictions in sales heterogeneity. They use firm-to-firm sales data for Belgium to demonstrate that the number of customers accounts for 51% of firm sale dispersion.

Last but not least, the paper adds to the growing literature on market dynamics in the digital economy that has been documented in other fields. For example, recent research in industrial organisation highlights the winner-takes-all feature of the digital economy (see Tirole [2020] for a review). On innovation, Paunov et al. (2019) find that patent applications are more concentrated in the top cities for digital-intensive sectors than for other sectors.

## 2. Data

### 2.1 Data on exporters

This paper draws information on a firm's non-oil export transactions from national customs agencies and is published in the World Bank's Exporter Dynamic Database (EDD) (Fernandes, Freund and Pierola, 2016). To the best of our knowledge, this dataset is currently the largest publicly available database on firm-level export dynamics in terms of coverage for developing countries. The public database exists in two main variants: an aggregated dataset of firm dynamics that is available for 65 countries (EDD v1.0), and a smaller subset of firm-level microdata for 11 countries (EDD v2.0). This paper employs the latter, EDD v2.0, dataset. In this variant, the export sales from each firm are disaggregated by year, destination and product at the HS 6-digit level.<sup>2</sup> For this paper, we aggregate the data at the firm-destination-year level for most applications.

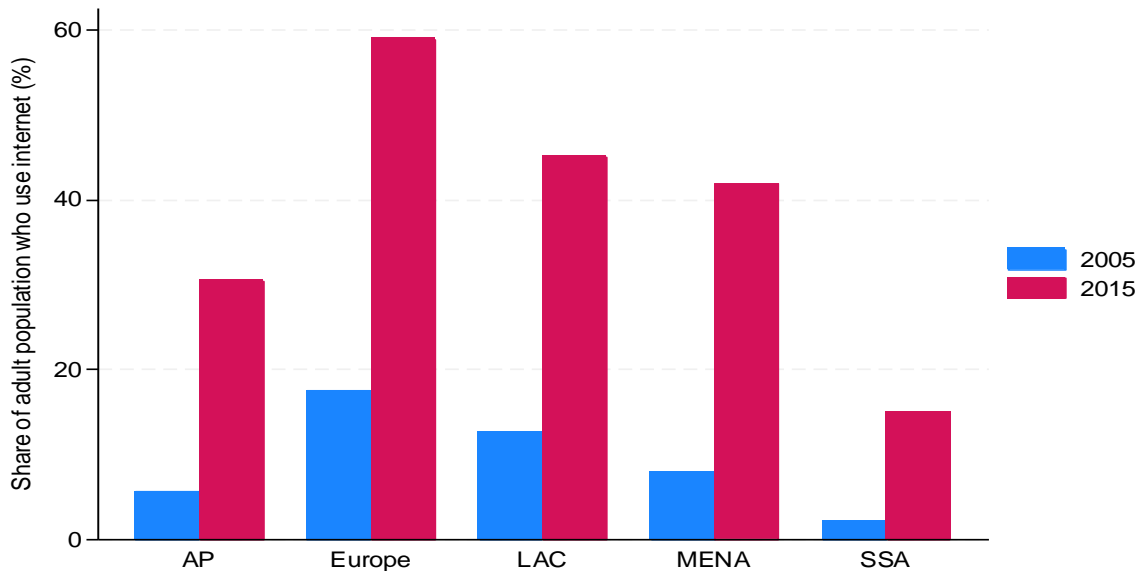
The dataset covers a broad set of developing countries that span across geographic regions and income brackets. These countries are Albania, Burkina Faso, Bulgaria, Guatemala, Jordan, Mexico, Malawi, Peru, Senegal, Uruguay, and Yemen. While the sample does not include any countries in Asia-Pacific (AP), we are not aware of robust evidence suggesting that the dynamics of internet productivity and export concentration in AP would differ from that found in other developing countries. Figure 1 shows that developing countries in AP had similar levels of internet penetration in 2005 and 2015 compared to other developing countries. Similarly, using aggregated data from EDD v1.0, we find that the share of the top 1% of exports in total exports appeared comparable between AP and elsewhere (Figure 2). This dataset includes

---

<sup>2</sup> The database only concerns the universe of non-oil exports from formal firms registered by custom agencies. It does not capture data on informal trader or unrecorded exports by formal firms, which may constitute an important part of cross-border trade in developing country.

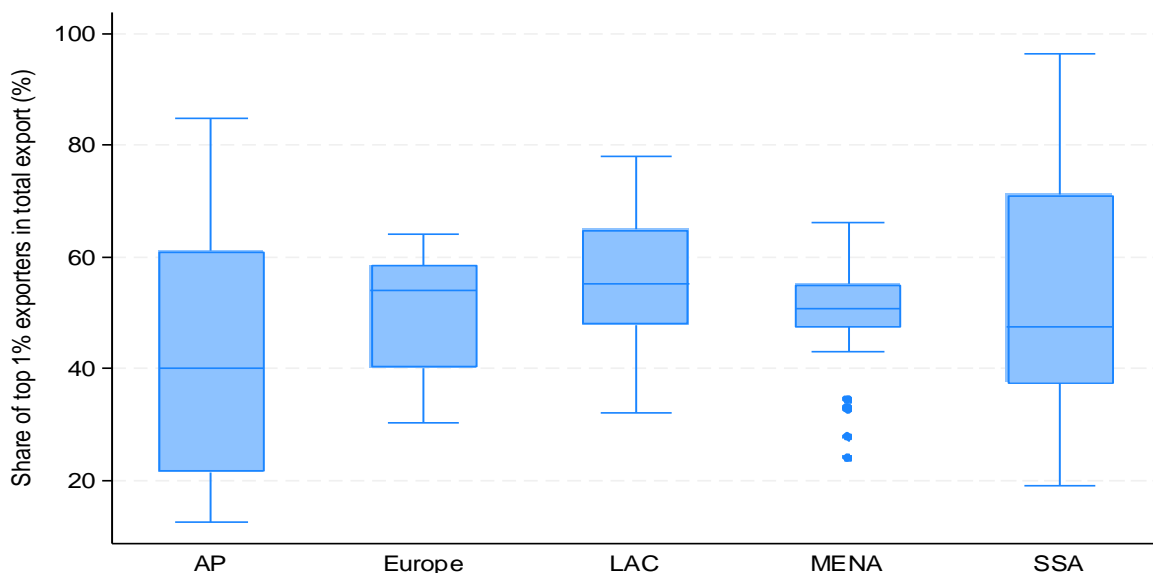
11 developing countries in AP and 42 developing countries in other regions.<sup>3</sup> These two empirical regularities suggest that the findings from other regions are likely to also play out in AP countries. Furthermore, the geographic coverage includes both low-income countries as well as middle-income countries. We also have both coastal and landlocked countries. Despite variation across countries, the dataset mostly covers the late 2000s period.

**Figure 1: Internet adoption level in developing countries**



Source: Authors' elaboration based on World Bank's World Development Indicator

**Figure 2: Share of export superstars in total export in developing countries**



Source: Authors' elaboration based on aggregated data from World Bank's Exporter Dynamic Database v1.0.

<sup>3</sup> These 11 developing countries in AP are Bangladesh, Cambodia, Georgia, Kyrgyz Republic, Lao PDR, Myanmar, Nepal, Pakistan, Sri Lanka, Thailand, and Timor-Leste.

## 2.2 Data on Internet use

We use the percentage of internet users among the adult population for each country to proxy for firms' use of the internet to search for partners. We source this data from the World Bank's World Development Indicators database. Internet adoption is among the most used measure to capture the impact of internet on international trade in the literature (such as Freund and Weinhold 2002, 2004; Hjort and Poulsen 2019).

Focusing on foundational technology like internet connectivity helps us capture the systemic impact of digitalisation on trade. González and Jouanjean (2017) argue that digital trade encompasses more than the trade in ICT goods and services but also trade enabled through growing digital connectivity. Underlying this whole systemic impact is the seamless flow of information and data, facilitated through internet connection. This could range from trade facilitated through e-commerce platforms to new trade relationships formed through cheap internet communication. However, as investment in internet infrastructure and the adoption of internet use are often endogenous to the development process, the impact of internet adoption may be confounded with other factors (that also affect trade patterns) such as household income and education level. Consequently, the paper uses a variety of panel econometric techniques to address these concerns such as fixed effects and instrumental variables.

During the studied period, the share of internet users globally increased six-fold from 7% in 2000 to 43% in 2015, according to the data from the International Telecommunication Union. Internet use in developing countries during this period mostly pertains to basic internet applications such as e-mail, low-cost internet-enabled voice calls and websites.

Unlike past cross-country literature which often treated internet measures of origin and destination countries separately, we argue that the utility of connectivity for cross-border trade depends on the internet usage of both origin and destination countries. The flow of information depends on the ability of both buyers and sellers to codify information (e.g. creating a website, digitising product specifications and listing) and process information (e.g. using internet search, navigating platform). At the same time, internet use is more likely to be driven by internet adoption in the weakest node of the supply chain. In other words, improvement in online communication between the U.S. and Malawi is more likely to be driven by improvements in internet use in Malawi, where only 14% of the population has access to the internet, rather than changes in internet use in the U.S., where over 86% of the population use the internet.

Following González and Ferencz (2018), we define our main variable, the bilateral internet adoption rate, as the pairwise minimum of internet usage for each pair of origin country  $i$  and destination country  $j$  in year  $t$ .

$$Internet_{ijt} = \min \{Internet_{it}, Internet_{jt}\}$$

### 3. Internet and distribution of exporter sales

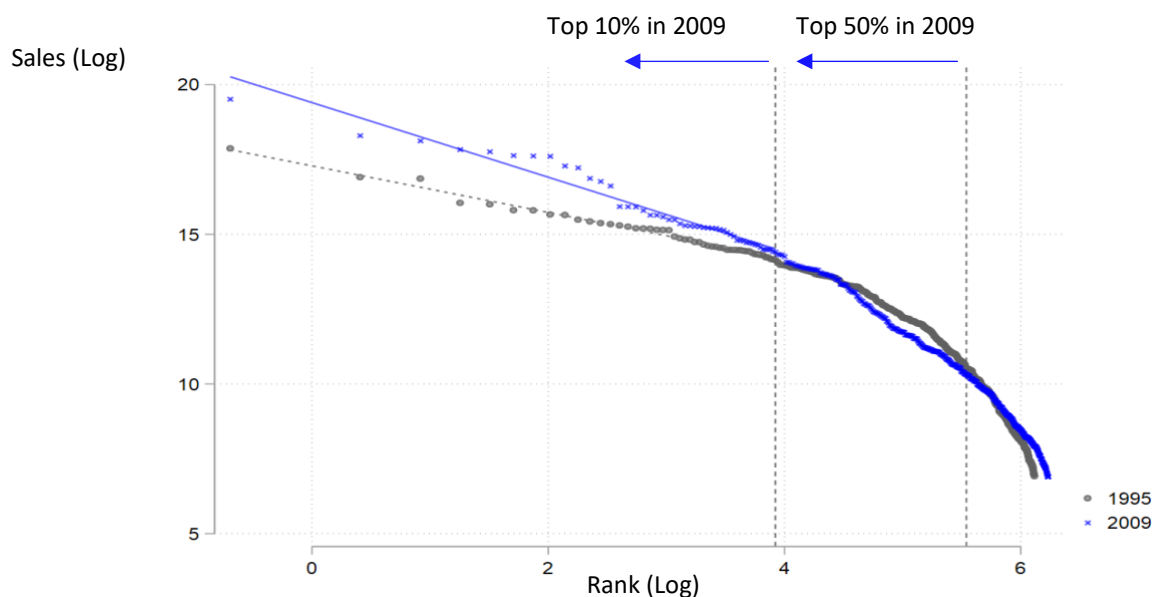
#### 3.1 Hypothesis

This section evaluates if internet adoption leads to more export concentration in each country pair. This empirical exercise faces three important considerations: (i) which distribution to approximate sales; (ii) which method to estimate such distribution; and (iii) how to account for firm-specific characteristics.

First, our focus on export superstars lends itself to a Pareto-type distribution. In the canonical firm-heterogeneity models following Melitz (2003) and Chaney (2008), a Pareto distribution of export sales to a specific destination often emerges endogenously. This is supported by previous empirical investigations such as Di Giovanni and Levchenko (2012) and particularly Nigai (2017), who suggest that using a Pareto distribution would better estimate the top quantile of the firm-level sales.<sup>4</sup>

Second, regarding the method of estimation, the log-log rank-size regression is among the most used methods to estimate the Pareto exponent (also known as the shape parameter that governs heterogeneity). For example, **Error! Reference source not found.**<sup>3</sup> visualises the log-log sale-rank relationship among Peruvian exporters to Germany in 1995 and 2009. This method has been used in various applications, including in international trade literature (Helpman, Melitz and Yeaple, 2004) and particularly in economic geography (Gabaix and Ibragimov, 2011).

**Figure 3: Distribution of sales among Peruvian exporters to Germany**



<sup>4</sup> In contrast, others such as Head et al. (2014) and Fernandes et al. (2023) suggest that the Pareto distribution is inferior to the lognormal distribution in fitting the entire distribution of export sales.

Third, the various cross-destination spillovers within firms may bias the analysis. The margins of firm participation in each market are often interdependent and complementary. For example, exporters in one market may find it less costly to export the same products to nearby markets owing to the acquired knowledge of the local culture and consumer preferences. Other firm idiosyncrasies may also affect the firm's export sales across markets. Freund and Pierola (2020) document that most export superstars are multinational enterprises that possess better information on foreign demands than domestic firms do.

The log-log rank-size regression appears most relevant for our analysis given these three considerations. This approach allows for the inclusion of high-dimensional fixed effects that account for various idiosyncratic effects, a critical advantage over estimation methods based on maximum likelihood estimations or quantile-quantile regressions. In addition, the flexibility to use the reverse specification of log-log sizes ranks enables us to make meaningful economic interpretations regarding the effect of internet expansion on average sales itself. If internet adoption has a positive impact on the Pareto exponent  $\kappa$  of export sales distribution, the result would imply a more concentrated distribution. The section 3 will evaluate *Hypothesis 1* formulated as below:

***Hypothesis 1*** *Higher internet adoption rate leads to an increase in the exponent of the export sales distribution for a given country-pair.*

### 3.2 Empirical setup

We start with a log-log size-rank regression, a variant of the more commonly used classic log-log rank-size with equivalent asymptotic properties (Gabaix and Ibragimov, 2011). They also suggest shifting the rank variable by a factor of  $\frac{1}{2}$  to ensure consistency in small samples. These modifications lead to the following setup:

$$\log(\text{Sale}_{fijt}) = \alpha + \beta \log(\text{Rank}_{fijt} - \frac{1}{2}) \quad (1)$$

where  $\text{Sales}_{fijt}$  represents the total export sales from firm  $f$  of country  $i$  to country  $j$  in year  $t$ .<sup>5</sup>  $\text{Rank}_{fijt}$  represents the descending rank in terms of sales by firm  $f$  among the exporters from  $i$  to  $j$  in year  $t$ . Given the negative correlation between the rank and sales level (see Figure 3 for an example),  $\beta$  takes on a negative value. The absolute value of this slope ( $\beta$ ) approximates the Pareto exponent  $\kappa$  (also called the shape parameter  $\tilde{\theta}$ ).

---

<sup>5</sup> For simplicity, the notation for origin country is suppressed whenever we refer to a specific firm. In EDD database, firm identifications are anonymised in each country, therefore we cannot observe if the same multinational may appear in multiple country datasets. We thus assume any firms from different countries are different from each other.

To separate the effect of internet adoption, we augment equation (1) with the bilateral internet adoption rates  $Internet_{ijt}$  defined in Section 2.2, and its interaction with firm ranks:

$$\begin{aligned} \log(\text{Sale}_{fijt}) = & \beta_1 \log(\text{Rank}_{fijt} - 1/2) + \beta_2 \log(\text{Rank}_{fijt} \\ & - 1/2) \times \log(\text{Internet}_{ijt}) + \beta_3 \log(\text{Internet}_{ijt}) + \Phi_f + \Phi_{it} \\ & + \Phi_{jt} + \Phi_{ij} + \varepsilon_{fijt} \end{aligned} \quad (2)$$

Our main coefficient of interest is that of the interaction term between log-rank and internet adoption  $\beta_2$ . This interaction term indicates the impact of internet adoption on the slope of the log-log sale-rank relationship (the exponent parameter). A negative coefficient  $\beta_2$  would imply that with stronger internet adoption rates, the (negative) slope of the sale-rank relationship becomes steeper. In this case, internet adoption facilitates a more concentrated distribution of export sales. The coefficient  $\beta_3$  of the internet adoption variable captures the impact of internet adoption on average sales level.

The firm fixed effect helps account for idiosyncratic behaviours of firms that may affect sales such as their ownership and main industry of activities. The linear trends for origin and destination countries account for multilateral resistances, overall demand, and supply trends, as well as other macroeconomic shocks in these countries. The origin-destination country pairs account for the trade cost between countries such as distance, borders, religion, common language, and colonial history.

**Error! Reference source not found.** displays summary statistics for the resulting dataset. In total, we observed 1,051,095 observations from 232,528 different exporters. There are 1,725 country pairs which form 10,963 origin-destination-year flows. In each flow, there are up to 28,997 exporters. The internet adoption rate stands at 13% for the median observation, ranging from 0.001% to 54.7%. We will express this variable as the log of percentage points.

**Table 1: Summary statistics for log-log sale-rank regressions**

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	N	mean	sd	p50	min	max
Sales <sub>fjt</sub> (in thousand USD)	1051095	1871	64421	36	1	30203306
Internet <sub>ijt</sub> (%)	1051095	13	10	12	0	55
Rank <sub>fjt</sub>	1051095	2886	6034	349	1	28997
Number of exporters <sub>ijt</sub>	1051095	5771	10044	861	1	28997

### 3.3 Results

Table 2 first estimates the exponent of the sale-rank distribution for the top 10% of exporters in each flow by estimating equation (1). In column (1), the Pareto exponent (the absolute value for the coefficient of the log-rank variable) of 0.915 conforms to previous estimates of the exponent at around 1 (albeit with a large standard error).

**Table 2: The effect of internet on log rank-log sale relationship**

	Top 10%	Top 10%	Top 50%	Top 50%	Full sample	Full sample
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Rank	-0.915 <sup>a</sup> [0.011]	-0.849 <sup>a</sup> [0.013]	-1.377 <sup>a</sup> [0.016]	-1.283 <sup>a</sup> [0.014]	-1.852 <sup>a</sup> [0.020]	-1.725 <sup>a</sup> [0.016]
Rank × Internet		-0.030 <sup>a</sup> [0.005]		-0.044 <sup>a</sup> [0.006]		-0.059 <sup>a</sup> [0.007]
Internet		0.227 <sup>a</sup> [0.076]		0.261 <sup>a</sup> [0.076]		0.366 <sup>a</sup> [0.083]
Observations	99,707	99,646	491,747	491,453	958,780	958,158
R-squared	0.958	0.958	0.946	0.946	0.920	0.921

Note: All regressions include origin and destination linear trends, country pair and firm fixed effects. Robust standard errors clustered at origin-destination-year level are reported in brackets: <sup>a</sup> p<0.01; <sup>b</sup> p<0.05; <sup>c</sup> p<0.1.

Column (2) presents the key result by estimating equation (2) for the top 10% of exporters and confirms Hypothesis 1. The coefficient for the interaction term between internet adoption and log-rank is negative and statistically significant at the 1% level. This result suggests that the internet reduces the negative slope of the log-log sale-rank relationship (hence increasing the exponent) – suggesting a more skewed sales distribution toward the top. To put this in perspective, the coefficient of log-rank suggests that there’s a 55.5% difference in sales between the biggest exporters and 2nd biggest exporters for a country pair with a median level of internet adoption (12%). Doubling the internet adoption rate would widen this gap by 1.5pp to 57.0%, *ceteris paribus*.<sup>6</sup> While the increase may not seem substantial, the bilateral internet adoption rate doubled every 1.7 years on average during the studied period.

The Internet has a positive and significant impact on sales of all exporters in the top 10%. This result is consistent with previous studies. For example, Fernandes et al. (2019) find that internet rollout increases the average export intensity of Chinese manufacturers in the same province. One standard deviation increase in internet penetration rate increases the sales for a certain exporter in an industry with the average intensity of internet use by 8.5%, *ceteris paribus*.

<sup>6</sup> Results in column (2) suggest that for bilateral flow with bilateral internet adoption rate at 12%, the difference in sales between the first-ranked and the second-ranked firm is  $\exp(-0.849 \times (\log(1 - \frac{1}{2}) - \log(2 - \frac{1}{2})) - 0.030 \times (\log(1 - \frac{1}{2}) - \log(2 - \frac{1}{2})) \times \log(12)) - 1 = 55.5\%$ . When the internet adoption rate doubles to 24%, the difference in sales is predicted as  $\exp(-0.849 \times (\log(1 - \frac{1}{2}) - \log(2 - \frac{1}{2})) - 0.030 \times (\log(1 - \frac{1}{2}) - \log(2 - \frac{1}{2})) \times \log(24)) - 1 = 57.0\%$ .



The slope of the log-log sale-rank relationship can be highly dependent on the quantile of export sales analysed (see **Error! Reference source not found.**<sup>3</sup> for an example and Nigai (2017) for a detailed discussion).<sup>7</sup> As a sensitivity analysis, we repeat the exercise for a larger sample of the top 50% of exporters and the full sample of all exporters. In both cases, the general slope of the rank-sale relationship steepens as the sample expands, highlighting the heavy tail characteristics of exporter distribution. Other studies such as Eaton et al. (2011) and Di Giovanni et al. (2011) found similar results. The main results hold qualitatively. They suggest that concentration occurs across the whole distribution of exporters, and not just among exporters at the very top, thus confirming *Hypothesis 1*.

### 3.4 Robustness

Our identification strategy relies on the temporal variation in bilateral internet adoption. In previous cross-country literature that exploits variations in country-level internet adoption rates, reverse causality is a concern because earnings from exports can fuel investment in infrastructure and facilitate internet adoption. Arguably reverse causality is less likely to be a concern for this paper as it analyses export sales at the firm level. The export performance of one firm is unlikely to affect directly the internet adoption rate of the whole population in either the origin or destination country, due to the high and non-divisible cost of setting up internet infrastructure.

Nonetheless, we evaluate the robustness of the results against reverse causality through two tests. In the first test, we use the one-year lag of bilateral internet adoption rates as the main policy variable. Internet use in the previous year is arguably exogenous to contemporary export performance, which helps avoid simultaneity bias. The results, presented in Panel A of Table A.2 in the Supplementary section, are qualitatively and quantitatively similar to the main results.

The second test uses internet speed as an instrumental variable for internet adoption. We use data on international Internet bandwidth per Internet user provided by the International Telecom Union. Similar to the internet use variable, bilateral internet speed takes on the value of the lower country-level speeds in each country pair. The rationale is that the slower node is more likely to create bottlenecks in information flow between these two countries. Because the annual data series have only been collected since 2000, the number of observations dropped.

Existing evidence suggests that internet speed is a key determinant of internet use across countries. In many developing countries, the issue of internet speed ranks among the top barriers to internet adoption (Gillwald and Mothobi, 2019). International Internet bandwidth is driven by investments in submarine cables and cross-border Internet Exchange Points.<sup>8</sup> Recent economic papers have used the roll-out and

---

<sup>7</sup> The deviation from power law at the tail-end of the distribution could reflect the fact that we do not observe firms' domestic sales.

<sup>8</sup> Satellite-based internet transmission remains limited to very few users in high income countries.

reliability of this backbone network as a proxy for internet use such as Hjort and Poulsen (2019).

Being an infrastructural index, internet speed is unlikely to directly affect the sales or shipment of physical goods in ways other than through promoting internet use and application. This property helps avoid concerns over the exclusion restriction of the instrumental variable model, which requires that internet speed has an impact on firms' export sales only through internet adoption.

While any ICT investments are generally affected by various structural factors such as local macroeconomic conditions and institutional settings, several examples suggest that such concerns are alleviated in the case of backbone networks in developing countries. For example, some backbone infrastructure in developing countries was developed thanks to speculative investment during the U.S.'s Dot Com bubble in the latter half of the 1990s. Many of these investments have limited connection to their actual use case.<sup>9</sup> Furthermore, the international bandwidth for many small countries is driven by infrastructure whose main purpose is to link other major economic hubs together.

Panel A of Table 3 presents the results from the first stage of the 2SLS procedures across the three samples of firms. The internet bandwidth and its interaction with rank variables are the two instruments for internet adoption and its interaction with sales rank, making the estimates just identified. Across all six columns, the Sanderson-Windmeijer F-statistics range from 1458 to 4130, suggesting that the internet speed variable is not weakly correlated with internet adoption (Sanderson and Windmeijer, 2016). For the top 10% of sales, a 10% increase in bandwidth increases the internet adoption rate by 4.96%. The interaction of internet bandwidth with sale rank also has a statistically significant role in explaining the variation of the interaction term between internet use and sale rank.

In Panel B of Table 3, results from the second stage of the 2SLS procedure confirm that higher internet adoption increases the exponent of the sales distribution. Internet adoption also has a positive and statistically significant relationship with the average sales of firms in each serving destination.

---

<sup>9</sup> While one can argue that the dot com bubble is also source of confounding factors that affect international trade, the limited spillovers from the dot com burst to the real economy suggests that such concern may not be a crucial factor.

**Table 3: Internet and log-log sale-rank relationship, IV approach**

**A. First stage**

	Top 10%		Top 50%		Full sample	
	(1)	(2)	(3)	(4)	(5)	(6)
	Internet	Rank×Internet	Internet	Rank×Internet	Internet	Rank×Internet
	t	t	t	t	t	t
<b>Bandwidth</b>	0.496 <sup>a</sup>	-0.067	0.511 <sup>a</sup>	0.026	0.522 <sup>a</sup>	0.089
	(0.022)	(0.085)	(0.036)	(0.199)	(0.042)	(0.261)
<b>Rank × Bandwidth</b>	-0.018	0.436 <sup>a</sup>	-0.016	0.415 <sup>a</sup>	-0.016	0.406 <sup>a</sup>
	(0.011)	(0.049)	(0.010)	(0.061)	(0.010)	(0.067)
<b>Rank</b>	0.200 <sup>b</sup>	-0.348	0.180 <sup>b</sup>	-0.113	0.179 <sup>b</sup>	-0.006
	(0.082)	(0.347)	(0.076)	(0.441)	(0.075)	(0.482)
<b>N</b>	101449	101449	501629	501629	100319 4	1003194
<b>SW F-stat</b>	3847	1458	4100	1688	4130	1761
<b>KP F-stat</b>	1287	1287	855	855	762	762

**B. Second stage**

	Top 10%	Top 50%	Full sample
	(1)	(2)	(3)
<b>Variables</b>	<b>Sales</b>	<b>Sales</b>	<b>Sales</b>
<b>Rank</b>	-0.908 <sup>a</sup> [0.029]	-1.296 <sup>a</sup> [0.033]	-1.682 <sup>a</sup> [0.039]
<b>Rank × Internet</b>	-0.004 [0.011]	-0.037 <sup>a</sup> [0.011]	-0.076 <sup>a</sup> [0.014]
<b>Internet</b>	1.157 <sup>b</sup> [0.545]	1.219 <sup>a</sup> [0.428]	1.360 <sup>a</sup> [0.426]
<b>Observations</b>	95,638	471,107	917,903
<b>R-squared</b>	0.779	0.854	0.811

*Note:* SW F-stat refers to Sanderson-Windmeijer F-statistics. KP F-stat refers to Kleibergen-Paap Wald F statistic. All regressions include origin and destination linear trends, country pair and firm fixed effects. Robust standard errors, clustered at origin-destination-year level, are reported in brackets: <sup>a</sup> p<0.01; <sup>b</sup> p<0.05; <sup>c</sup> p<0.1.

A potential issue with our definition of bilateral internet adoption is that when one country has a persistently lower internet adoption rate than its partner, the variation in bilateral internet use could be exclusively driven by changes in the adoption rate in this country. In column (1) of Table 4, we limit the sample only to country pairs where the origin country had a lower internet penetration rate in the first year when exporter

data was available ( $t_0$ ).<sup>10</sup> In this situation, a bilateral internet connection would largely be driven by variations in internet penetration in the exporting country. Inversely, column (2) exploits a subsample of country pairs where the destination country had a lower internet penetration rate in the first year where exporter data is available ( $t_0$ ).

**Table 4: Effect of internet on rank-sale exponent among the top 10%, covariates and sample robustness**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Internet <sub>it0</sub> < Internet <sub>jt0</sub>	Internet <sub>it0</sub> ≥ Internet <sub>jt0</sub>	Internet as %	N≥100	Existing pairs	Covariate robustness	
VARIABLES	Sales	Sales	Sales	Sales	Sales	Sales	Sales
Rank	-0.780 <sup>a</sup>	-0.901 <sup>a</sup>	-0.893 <sup>a</sup>	- 0.870 <sup>a</sup>	-0.851 <sup>a</sup>	-0.802 <sup>a</sup>	-0.861 <sup>a</sup>
	[0.012]	[0.018]	[0.014]	[0.015]	[0.013]	[0.020]	[0.013]
Rank × Internet	-0.029 <sup>a</sup>	-0.019 <sup>a</sup>		- 0.035 <sup>a</sup>	-0.030 <sup>a</sup>	-0.026 <sup>a</sup>	-0.032 <sup>a</sup>
	[0.006]	[0.006]		[0.005]	[0.005]	[0.005]	[0.005]
Internet	-0.413 <sup>c</sup>	-0.128		0.378 <sup>a</sup>	0.247 <sup>a</sup>	0.270 <sup>a</sup>	
	[0.220]	[0.249]		[0.113]	[0.078]	[0.081]	
Rank × Internet (%)			-0.002 <sup>a</sup>				
			[0.001]				
Internet (%)			0.004				
			[0.004]				
Observations	25,960	72,198	99,707	86,545	98,896	88,597	97,218
R-squared	0.946	0.963	0.958	0.974	0.959	0.977	0.971
<b>Fixed effects</b>							
Origin-Year	Y	Y	Y	Y	Y	Y	
Dest.-Year	Y	Y	Y	Y	Y	Y	
Firm-Dest.						Y	
Origin-Dest.- Year							Y
Firm	Y	Y	Y	Y	Y		Y
Origin-Dest.	Y	Y	Y	Y	Y		

Note: Robust standard errors clustered at origin-destination-year level are reported in brackets: <sup>a</sup> p<0.01; <sup>b</sup> p<0.05; <sup>c</sup> p<0.1.

In both cases, the results remain qualitatively similar and confirm *Hypothesis 1*. Interestingly the effect of internet use on sales differs across the two specifications. This divergence is consistent with findings from Cariolle, de Melo and Imbruno (2022)

<sup>10</sup> To conserve space, it focuses the analysis on the top 10% of exporters. Results for the top 50% and the whole sample are qualitatively similar and available upon request.

who suggest that the impact of internet adoption on trade depends on the origin country of exporters.

Finally, we show the robustness of the results against alternative variable specifications and subsamples. In column (3), we further test the sensitivity of the result when specifying the internet adoption variable as a percentage (rather than log specified). Column (4) evaluates equation (2) using a subsample of only bilateral flows where there are at least 100 exporters. This subsample ensures that the top 10% in each flow include at least 10 firms. Column (5) addresses concerns over selection by keeping only flows that are already active in the first year when customs data is available for each of the 11 developing countries in the EDD v2.0 dataset. This approach excludes new destinations that exporters from these 11 countries only exported more recently, as these new destinations could be a consequence of internet adoption. All these regressions confirm the main results to support *Hypothesis 1*.

## 4. Internet and entrenchment of export superstars

### 4.1 Hypothesis

Section 3 provides evidence that the concentration of export sales in each origin-destination flow increases with better internet. However, the concentration of exports in a given year may not necessarily imply a reliance on the superstars if the composition of firms that make up this top group changes frequently, i.e., there is high churning.

In contrast, an “entrenchment” of the top exporters who retain their dominance from one year to another would suggest sustained concentration over time and raise concerns for policymakers. This is an important point from the statistical analysis by Freund and Pierola (2020). They find that 71% of the top five exporters in a given year already featured in the top 1% of exporters five years previously.

This section aims to understand the entrenchment of export superstars in the Internet era. Particularly, it evaluates if incumbent superstars are more likely to retain their top positions from one year to another as internet adoption expands thanks to the internet environment – formulated as *Hypothesis 2*:

***Hypothesis 2***      *Superstar exporters in each country-pair are more likely to retain their superstar status when the internet adoption rate is higher.*

The definition of “superstars” is sensitive to this compositional change. When the number of exporters increases, those previously outside the top quantile (e.g., the top 5% and top 10% of exporters) can now attain superstar status simply thanks to the numerical expansion of the top quantile. In contrast, ordinal-rank definition (e.g., top

one or top five firms) can be distorted in flows with fewer exporters as exporters outside the top five can now feature in this group thanks to churning, everything else being equal. To address this concern, we will assess the model one by one using both the relative ranks and the ordinal ranks of superstars – a similar approach to Freund and Pierola (2020).

## 4.2 Empirical setup

The baseline model to evaluate *Hypothesis 2* takes the following form:

$$\begin{aligned} \text{Prob}(Superstar_{fij,t} = 1) & & (3) \\ &= \beta_1 Superstar_{fij,t-1} + \beta_2 Superstar_{fij,t-1} \times \log(Internet_{ij,t}) \\ &+ \beta_3 \log(Internet_{ij,t}) + Z_{fijt} + \varepsilon_{fijt} \end{aligned}$$

where  $Superstar_{fij,t}$  is a dichotomous variable indicating whether a certain firm  $f$  has superstar status in terms of total sales from country  $i$  to country  $j$  in year  $t$ .  $Superstar_{fij,t-1}$  represents this firm's superstar status in the preceding year.<sup>11</sup> We will assess the model one by one using both the relative rank of superstars (top 5% and top 10% of exporters in the respective flows) and the ordinal rank of superstars (top 1, top 2 and top 5 of exporters).

The main coefficient of interest is the interaction term  $\beta_2$  between exporter's (lag) superstar status and internet adoption. A positive and statistically significant coefficient would confirm *Hypothesis 2* as it suggests that internet connectivity increases the likelihood of top exporters retaining their superstar status from  $t-1$  to  $t$ .

$Z_{fijt}$  denotes a series of fixed effects to account for omitted variable bias. Similar to section 3, we include fixed effects for each firm, each country-pair, and time-varying trends for each origin and destination country. In this setup, the firm-level fixed effect controls for unobserved time-invariant characteristics such as the average level of productivity and ownership structure (which is unlikely to change given the short panel for most of the countries).

The origin-trend fixed effects capture any shocks to the local economy that may have asymmetric impacts among producers and affect the possibility for any exporters to feature in the “superstar” category. For example, distortions in the domestic economy can reduce misallocate resources away from the most productive. Fernandes, Freund and Pierola (2016) find that the dynamics of exporters also depend on the stage of development in the origin country. For example, the importance of the top 5% in total export values increases with total export sales. Similarly, the destination-year fixed effects account for various demand-side factors that affect heterogeneity such as the

---

<sup>11</sup> Firms that export in year  $t-1$  but not in  $t$  (exitors) will be denoted as non-superstars in year  $t$  ( $Superstar_{ijt}=0$ ). New entrants (export in year  $t$  but not in  $t-1$ ) will be denoted as non-superstars in year  $t-1$  ( $Superstar_{ijt-1}=0$ ).

level of competition in the economy, changing consumer demands and preference for quality as well as unilateral opening of domestic markets.

In addition, we include product fixed effects to account for industry-specific factors such as the capital intensity or the structure of the industry in which each firm operates. Products are identified as the most important product heading (HS 4-digit) in the firm's export basket to each destination in a given year. It reflects the comparative advantage of the exporter compared to local firms in each importing country. It also captures the evolution in the exporter's product orientation over time.

Supplementary Table A.3 provides a summary of the variables used in these regressions on superstars.

### 4.3 Results

Table 5 presents the results from estimating equation (3) for different definitions of superstars. Due to the large number of fixed effects employed and computational constraints, we estimate the model with OLS. This approach also helps avoid the incidence parameter problem that hinders non-linear estimators such as probit or logit in the context of a short panel.

**Table 5: Effect of internet on export superstars' persistence**

	Relative rank		Ordinal rank		
	(1)	(2)	(3)	(4)	(5)
<b>Superstar definition</b> ➔	<b>Top 5%</b>	<b>Top 10%</b>	<b>Top 1</b>	<b>Top 2</b>	<b>Top 5</b>
<b>Superstar (lag)</b>	0.369 <sup>a</sup>	0.392 <sup>a</sup>	0.132 <sup>a</sup>	0.082 <sup>a</sup>	0.046 <sup>a</sup>
	[0.013]	[0.011]	[0.010]	[0.009]	[0.009]
<b>Superstar (lag) × Internet</b>	0.052 <sup>a</sup>	0.039 <sup>a</sup>	0.036 <sup>a</sup>	0.052 <sup>a</sup>	0.065 <sup>a</sup>
	[0.005]	[0.004]	[0.004]	[0.004]	[0.003]
<b>Internet</b>	-0.009 <sup>b</sup>	-0.009 <sup>c</sup>	-0.006 <sup>c</sup>	-0.007	0.001
	[0.005]	[0.005]	[0.004]	[0.005]	[0.007]
<b>Observations</b>	1,285,471	1,285,471	1,285,471	1,285,471	1,285,471
<b>R-squared</b>	0.466	0.486	0.291	0.340	0.399

Note: All regressions include country-pair, origin and destination linear trends, firm and product fixed effects. Robust standard errors clustered at origin-destination-year level are reported in brackets: <sup>a</sup> p<0.01; <sup>b</sup> p<0.05; <sup>c</sup> p<0.1.

In column (1), we first observe that in an environment without internet connectivity, the top 5% exporter in the previous year has a 36.9-percentage point (pp) higher chance of featuring in the top 5% this year, compared to other exporters. This is consistent with Freund and Pierola (2020), who report that in 32 countries where firm-panel data

is available, 71% of the top five firms were already in the top 1% of exporters five years earlier.

We also find that the expansion of the internet also decreases the likelihood for any firm to become a superstar in a country-pair flow. This likely reflects the fact that more firms are exporting in a connected environment. By encouraging entry and exit, the internet inflates the sample of firms for consideration in each country-pair-year triad by including both current and recently exited exporters.<sup>12</sup>

The positive and statistically significant coefficient of the interaction term between superstar status and internet penetration suggests that internet adoption helps top exporters retain their status. Everything else being equal, doubling the internet penetration rate boosts the probability that the top 5% of exporters will retain their ranks by an additional 4.3pp (=0.052-0.009), representing an 11.7% increase in their chance of retaining the top spots.

Column (2) follows a similar strategy with superstars defined as those in the top 10% of sales. Exporters in the top 10% of sales in the previous years have a 39.2pp higher likelihood to feature in the top 10% again this year, compared to those not previously in the top 10%. In an environment where internet adoption doubles, this likelihood is boosted by 3.0pp, equivalent to a boost of 7.7%.

Columns (3)-(5) confirm *Hypothesis 2* using ordinal-rank definitions of superstars (the top one, the top two, and the top five exporters in an origin-destination-year triad). Column (3) for example presents the estimation of equation (3) with superstars defined as the single top exporter. The top exporter in the previous year has a 13.2pp higher chance of retaining this status, compared to other exporters. Doubling of internet adoption rate boosts the probability that the top exporter will retain its top position from one year to another by 3.0pp, representing a 22.7% increase in their chance of retaining the top spots. The results for the top 2 and top 5 exporters are qualitatively similar, whereby doubling the internet penetration rate would lead to a boost of 54.9% and 143.5% respectively in the chances that these superstars retain their dominance.

#### **4.4 Robustness**

This section evaluates the robustness of *Hypothesis 2* using the samples of the top 5% and top 1 exporters. Results using other definitions of superstars are qualitatively similar and available upon request.

In this analysis, reverse causality can occur where an incumbent superstar creates pressure on the local government to invest in internet infrastructure to facilitate their trading. However, existing experience suggests that the typical response is to concentrate their investment in spatially-targeted industrial clusters. Such investments

---

<sup>12</sup> Simple correlation analysis, available upon request, suggests that Internet adoption rate correlates positively with both the number of exiting firms (those who export in t-1 but not in t) and the share of exiting firms in total exporters.



would have little impact on the rate of internet access among the majority of the population who typically live outside these zones. Indeed, the lack of linkage between many industrial zones and the general economy has often been criticised in numerous policy evaluations.

We further test against using endogeneity by using international internet bandwidth and its interaction with the lagged superstar status as instrumental variables. For brevity, we present the first-stage results in Supplementary Table A.4. The Sanderson-Windmeijer F-statistics in the first-stage regressions range from 79 to 4737 – alleviating the concerns over weak instruments in this setup. The second-stage results presented in Table 6 suggest that Internet adoption helps superstars, defined by their ordinal rank, to retain their status. While the coefficients for the interaction term between internet adoption and superstar remain positive for the relative rank definitions of superstars, these estimates are no longer statistically significant.

**Table 6: Effect of internet on superstars, IV approach**

	Relative rank		Ordinal rank		
	(1)	(2)	(3)	(4)	(5)
Superstar definition →	Top 5%	Top 10%	Top 1	Top 2	Top 5
Superstar (lag)	0.461 <sup>a</sup>	0.462 <sup>a</sup>	0.154 <sup>a</sup>	0.118 <sup>a</sup>	0.096 <sup>a</sup>
	[0.024]	[0.020]	[0.016]	[0.014]	[0.014]
Superstar (lag) × Internet	0.012	0.010	0.020 <sup>b</sup>	0.029 <sup>a</sup>	0.036 <sup>a</sup>
	[0.009]	[0.008]	[0.008]	[0.007]	[0.006]
Internet	-0.019	-0.024	0.000	0.013	-0.004
	[0.022]	[0.026]	[0.015]	[0.023]	[0.036]
Observations	1,226,696	1,226,696	1,226,696	1,226,696	1,226,696
R-squared	0.238	0.235	0.040	0.036	0.037
SW F-stat Internet	79	80	80	80	80
SW F-stat Superstar × Internet	603	586	4737	4577	3989

Note: The table reports Sanderson-Windmeijer F-statistics for first stage regressions for  $Internet_{ijt}$  and  $Superstar(lag) \times Internet_{ijt}$  being the endogeneous regressors. All regressions include country-pair, origin and destination linear trends, firm and product fixed effects. Robust standard errors clustered at origin-destination-year level are reported in brackets: <sup>a</sup>  $p < 0.01$ ; <sup>b</sup>  $p < 0.05$ ; <sup>c</sup>  $p < 0.1$ .

**Error! Reference source not found.** evaluates the robustness of the results against alternative samples. In columns (1)-(2), we limit the sample only to country pairs where the origin country had a higher internet penetration rate in the first year where exporter data is available for each exporting country ( $t_0$ ) - and vice-verse in columns (3)-(4). Columns (5)-(6) test whether our results are affected by the over-representation of flows with too few firms that arithmetically place more importance on small firms. We restrict the sample to flows with at least 100 exporters and find qualitatively similar results. The results remain qualitatively similar in all settings and for both superstar definitions.



**Table 7: Effect of internet on superstars, with alternative samples**

Variables	Internet <sub>it0</sub> ≥ Internet <sub>jt0</sub>		Internet <sub>it0</sub> < Internet <sub>jt0</sub>		Flows with 100+ exporters		Four-year performance	
	(1) Top5%	(2) Top 1	(3) Top5%	(4) Top 1	(5) Top5%	(6) Top 1	(7) Top5%	(8) Top 1
<b>Superstar (lag)</b>	0.269 <sup>a</sup>	0.093 <sup>a</sup>	0.422 <sup>a</sup>	0.112 <sup>a</sup>	0.527 <sup>a</sup>	0.412 <sup>a</sup>		
	[0.014]	[0.013]	[0.020]	[0.017]	[0.014]	[0.028]		
<b>Superstar (lag) × Internet</b>	0.065 <sup>a</sup>	0.031 <sup>a</sup>	0.031 <sup>a</sup>	0.046 <sup>a</sup>	0.014 <sup>b</sup>	0.008		
	[0.006]	[0.006]	[0.007]	[0.007]	[0.005]	[0.011]		
<b>Internet</b>	0.003	0.033 <sup>b</sup>	0.001	0.001	-0.000	0.000		
	[0.017]	[0.015]	[0.009]	[0.006]	[0.003]	[0.000]		
<b>Superstar (4y lag)</b>							0.382 <sup>a</sup>	0.317 <sup>a</sup>
							[0.041]	[0.039]
<b>Superstar (4y lag) × Internet</b>							0.097 <sup>a</sup>	0.019
							[0.017]	[0.018]
<b>Observations</b>	302,744	302,744	977,446	977,446	1,138,781	1,138,781	95,883	95,883
<b>R-squared</b>	0.420	0.347	0.500	0.288	0.537	0.323	0.576	0.413

Note: All regressions include country-pair, origin and destination linear trends, firm and product fixed effects. Robust standard errors, clustered at origin-destination-year level, are reported in brackets: <sup>a</sup> p<0.01; <sup>b</sup> p<0.05; <sup>c</sup> p<0.1.

A selection problem may bias our estimation as the internet may induce specific types of firms to enter and exit the export market. On the one hand, the internet could attract new non-exporters to experiment with exporting. These firms typically have low productivity levels, high churning rates and an overall limited probability of featuring at the top. Their presence in the control group of non-superstars thus inflates the “entrenchment” effect of the superstars.

On the other hand, the internet may have also induced the emergence of export superstars by facilitating off-shoring investment from the Global North to developing countries. Fort (2017) shows that the adoption of communication technology increases the probability of outsourcing and offshoring among U.S. manufacturing firms. Freund and Pierola (2020) document multiple instances where new firms - which did not feature in the sample five years earlier - start exporting with a very high amount (in monetary value). These are often foreign-owned firms and many of them grow to be in the top five group within two years. The emergence of these internet-driven born-to-be superstars may create a downward bias on the effect of the internet on a superstar’s “entrenchment”.

To account for firm selection, columns (7) and (8) focus on the long-run dominance of the “original” superstars - those who already attained this status in the first year that exporter data are available in the EDD v2.0 dataset. Furthermore, we focus on their export performance 4 years later instead of tracking the dynamics of exporters in the following year. The choice of 4 years is driven by the length of panel data for exporters from Yemen, which is only available for the 2008-12 period. Overall, results presented

in columns (7)-(8) largely confirm *Hypothesis 2* whereby we find that increasing the internet adoption rate would boost the probability for these “original” superstars to feature in the top ranks 4 years later.

## 5. Mechanisms

Most papers evaluating the impact of the Internet on international trade focus on aggregate trade flows in a cross-country analysis. They often exploit the difference in internet rollout across countries, regions, and time dimensions to estimate the impact of the internet on trade in a gravity model. The pioneering work by Freund and Weinhold (2002) finds that a 10-percent increase in the share of internet users in the partner countries leads to a 1.7-percentage-point boost to service trade with the U.S. For trade in goods, Freund and Weinhold (2004) find that a 10pp increase in the growth of web hosts in a country leads to about a 0.2pp increase in export growth.

This cross-country strand of literature often emphasises the role of Internet communication in reducing the fixed cost of searching and matching. When calculating the ex-ante profitability of exporting, firms are assumed to perfectly observe the different costs and know all the potential partners. If the expected benefit outweighs the expected cost, they will then pay the fixed cost, including the fixed cost of searching and matching with importers, to start trading. These papers often focus on aggregate trade flows, and thus they overlook the heterogeneity of firms.

However, in the heterogeneous firm model, a reduction in fixed cost is unlikely to lead to export concentration. Most of the literature following the canonical Melitz (2003) model assumes the Pareto distribution of a firm’s productivity with a location parameter ( $\alpha$ ) and a shape parameter ( $\theta$ ). With an exogenous constant elasticity of substitution  $\sigma$  and monopolistic competition, export sales to each destination will also be Pareto-distributed with the shape parameter of sales  $\tilde{\theta} = \frac{\theta}{\sigma-1}$ . Reducing the fixed cost to start trading will shift the productivity cut-off and induce new and less-productive firms to enter exporting. While this change may affect the definition of superstars, it has an ambiguous effect on the heterogeneity of sales distribution.

To reconcile findings from sections 3 and 4 with the heterogeneous firm framework, the rest of this section will discuss the potential role of information frictions in explaining these empirical regularities.

## 5.1 Information friction and sales concentration

Our findings could be explained through the lens of information friction. Exporters could face information frictions related to the costly monitoring of market prices across destinations (Allen, 2014; Steinwender, 2018), awareness of potential buyers and their credibility (Akerman et al., 2022), and knowledge on demand conditions (Dickstein and Morales, 2018). Importers also face friction due to awareness about potential suppliers (Lenoir, Martin and Méjean, 2022), the cost of processing information (Dasgupta and Mondria, 2018), and information on the credibility of sellers on online platforms (Chen, Maggie and Wu, 2020).

Connectivity helps each exporter and importer to have a wider reach to potential buyers and sellers. The wider choice set makes it easier for firms to find substitutes and makes demand more elastic, following Marshall's first rule of derived demand. Akerman, Leuven and Mogstad (2022) formalise this argument in a model with variable elasticity of demand. They find that expansion of broadband internet in a region makes export sales from this region more sensitive to prices, and hence the difference in trade costs across destinations. Extending this argument, we would also expect export sales to become more sensitive to production costs. Consequently, firms with lower per-unit costs would score higher sales than low-productivity ones when holding the same distribution of productivity. The provision of information through Internet communication thus leads to export sales concentration towards high-productivity firms, who are typically large sellers already.

That high-productivity firms gain disproportionately higher sales thanks to information provision is also a key insight from Lenoir, Martin and Méjean (2022). In their Ricardian model, search frictions mean that a buyer is drawn against a random number of heterogeneous producers from each origin country – from whom she chooses the lowest-cost suppliers. Higher search friction is more penalising for high-productivity firms because they are much more likely to win sales from the new, expanded set of potential buyers. In contrast, the change can be detrimental for low-productivity sellers because they are less likely to become the lowest-cost producers when there are more competitors for each buyer. Conversely, being exposed to fewer customers is not as penalising as the likelihood of not winning the sale is higher for them. As a result, internet connectivity increases the allocation of sales towards high-productivity sellers who are much more likely to win sales from the improved visibility.

Recent research has showcased the role of communication technologies in reducing information friction within international trade. By exploiting spatial variation in mobile phone coverage in the Philippines, Allen (2014) demonstrates that information frictions account for half of the observed price dispersion in agricultural trade among the islands in the country. Steinwender (2018) shows that the transatlantic telegraph connection in the 19<sup>th</sup> century enhanced the flow of information on cotton markets between the U.S. and the U.K. markets. The reduction in information frictions led to a lower level of price dispersal, as well as larger and more volatile trade flows. Dasgupta and Mondria

(2018) are more explicit about the role of internet connection in alleviating informational frictions. They assume the maximum internet bandwidth between the two countries as a proxy for the cost of information processing for exporters and importers in these countries.

If internet adoption influences export concentration through alleviating information friction, we would expect the impact of internet adoption on export concentration to be more pronounced for trading of information-intensive goods. For example, buyers typically need more communication to find the right sellers of differentiated goods, and to convey various product specifications and characteristics. In contrast, the use of commodity trading platforms and referenced prices for homogeneous goods trade facilitates the process of searching and matching between buyers and sellers (Rauch, 1999). Previous literature such as Akerman, Leuven and Mogstad (2022) find stronger effects of broadband adoption on the demand elasticity of differentiated goods as compared to homogeneous goods.

As an extension of *Hypothesis 1*, we will test if the impact of internet expansion has differential impacts on the Pareto exponent  $\kappa$  of export sales distribution among the two types of goods: homogeneous goods (commodity and reference-priced goods) and differentiated goods (all other types of goods) as classified by Rauch (1999). For each type, we aggregate firms' export sales in this category to each destination and recalculate the sale ranking accordingly.

To evaluate *Hypothesis 3a*, we extend equation (2) by focusing on the three-way interaction term between rank, internet adoption and differentiated goods dummy would inform the differential impact of connectivity on the Pareto exponent between differentiated and homogeneous goods. Other controls and interaction terms involving *Differentiated goods* help account for the structural difference in the distribution of sales between these two types of goods.

Table 8 consecutively presents the estimation results for the top 10%, the top 50% and the full sample. In the first two estimations, we do not report a statistically significant difference regarding the impact of connectivity on the exponent. In the third estimation, we find a negative and statistically significant coefficient for the three-way interaction term between rank, internet adoption and goods type, which would suggest a stronger impact of connectivity on sales heterogeneity for differentiated goods, thus supporting our conjecture.

**Table 8: Effect on log-log ranks-sales relationship, by type of goods**

	(1)	(2)	(3)
Variables	Sales	Sales	Sales
Rank	-0.982 <sup>a</sup>	-1.384 <sup>a</sup>	-1.843 <sup>a</sup>
	[0.059]	[0.035]	[0.017]
Rank × Internet	-0.048 <sup>a</sup>	-0.043 <sup>a</sup>	-0.047 <sup>a</sup>
	[0.018]	[0.012]	[0.009]
Rank × Internet × Differentiated goods	0.029	-0.011	-0.036 <sup>a</sup>
	[0.020]	[0.009]	[0.011]
Internet	0.402 <sup>a</sup>	0.420 <sup>a</sup>	0.426 <sup>a</sup>
	[0.099]	[0.127]	[0.151]
Observations	98,729	489,668	960,650
R-squared	0.939	0.927	0.902
Sample	Top 10%	Top 50%	Full sample

Note: All regressions include country-pair, origin and destination linear trends and firm fixed effects. Dummies for differentiated goods, and its interactions with the internet variable and export ranks were also included in the regressions but omitted from the table for brevity. Robust standard errors clustered at origin-destination-year level are reported in brackets: <sup>a</sup> p<0.01; <sup>b</sup> p<0.05; <sup>c</sup> p<0.1.

Similarly, we augment equation (3) with a dummy for differentiated goods, its interactions with the internet and the (lag) superstar variable, and its three-way interaction term with (lag) superstar and internet variable. Our main interest lies in the coefficient of this three-way interaction term.

The results presented in **Error! Reference source not found.** show that the impact of internet adoption on superstar dominance is stronger for differentiated goods when using an ordinal rank definition. However, we do not find a significant difference in the impact between differentiated and homogeneous goods when considering relative rank definition. For robustness, we estimate equations (2) and (3) individually for exports of each type of goods. The results presented in Supplementary Table A.6 also show that the magnitude of the impact of the Internet is broadly higher for the trade of differentiated goods than for the trade of homogeneous goods.

**Table 9: Effect on superstar's persistence, by type of goods**

	Relative rank		Ordinal rank		
	(1)	(2)	(3)	(4)	(5)
Superstar definition →	Top 5%	Top 10%	Top 1	Top 2	Top 5
Superstar (lag)	0.376 <sup>a</sup>	0.420 <sup>a</sup>	0.315 <sup>a</sup>	0.335 <sup>a</sup>	0.379 <sup>a</sup>
	[0.011]	[0.009]	[0.010]	[0.008]	[0.007]
Superstar (lag) × Internet	0.040 <sup>a</sup>	0.029 <sup>a</sup>	0.015 <sup>a</sup>	0.013 <sup>a</sup>	0.013 <sup>a</sup>
	[0.005]	[0.004]	[0.004]	[0.003]	[0.003]
Superstar (lag) × Internet × Differentiated	-0.001	-0.002	0.013 <sup>b</sup>	0.009 <sup>b</sup>	0.004
	[0.004]	[0.004]	[0.005]	[0.004]	[0.003]
Observations	1,057,224	1,057,224	1,057,224	1,057,224	1,057,224
R-squared	0.484	0.511	0.330	0.370	0.430

Note: All regressions include country-pair, origin and destination linear trends, firm and product fixed effects. Additional controls, omitted from table for brevity, include internet adoption, a dummy for differentiated goods and its interaction with internet adoption and with superstar status. Robust standard errors clustered at origin-destination-year level are reported in brackets: <sup>a</sup> p<0.01; <sup>b</sup> p<0.05; <sup>c</sup> p<0.1.

## 5.2 Other channels

The spread of internet connectivity and the broader diffusion of information and communication technologies may also affect heterogeneity through supply-side factors such as exporters' productivity, access to inputs and organisation capital. Such technologies can induce technological adoption that shapes firms' productivity and characteristics through changes in the use of tangible inputs such as intermediate materials and machinery and intangible inputs such as management, organisational practice, and support services (see Hjort and Lin (2022) for a recent summary of this literature).

Connectivity may also exacerbate productivity dispersion among firms in the origin country. First, Internet connectivity may induce changes in the optimal input mix and disproportionately favour firms already endowed with higher productivity such as enterprises with a higher share of highly skilled labour (Akerman et al., 2015), and enterprises in digital-intensive sectors such as finance (Hjort and Poulsen, 2019). Lower supply-chain coordination cost also facilitates production fragmentation and induces multinational enterprises to reallocate to developing countries, where they would often feature among the most productive firms. Fort (2017) shows that the adoption of communication technology increases the probability of outsourcing and offshoring among U.S. manufacturing firms.



Second, the lumpy cost of internet adoption may also widen the productivity gap between small exporters and larger ones (who can spread the lumpy cost over more units). Digital skills, organisation capability and financial requirements may create barriers to the adoption of Internet applications even when such investment is technologically feasible and economically viable (Cariolle, de Melo, and Imbruno, 2022). This problem is particularly severe in environments with a limited supply of high skills and underdeveloped financial markets – such as those in developing countries – that would increase the lumpy cost of adoption.

The baseline sets of fixed effects should help alleviate these confounding factors. The firm-specific fixed effects capture the average productivity of each firm across the periods. It controls for other critical firm characteristics such as foreign-owned enterprises, who may have relocated to the origin country thanks to easier coordination of international production networks. The inclusion of origin- and destination-trend fixed effects also captures systemic changes to the various supply- and demand-side factors that may affect firm heterogeneity.

To further address concerns over the uneven adoption of Internet technologies across firms, we further add firm-year fixed effects to the two models which help capture the evolution in firms' productivity, (overall) quality and product scope. Annex Table A.7 presents the result and confirms the robustness of the main results against the inclusion of these stringent controls.

## **6. Conclusions and Policy Implications**

This paper presents empirical evidence that the expansion of Internet adoption increases the concentration of export sales toward a few top firms and strengthens the likelihood of incumbent export superstars retaining their status. It is the first to suggest that the global expansion of the Internet has contributed to the export superstars' phenomenon.

Internet connectivity allows exporters and importers to expand their network and thus intensify the level of competition that each exporter faces. This shift leads to higher cost sensitivity among importers and disproportionately favours sales by producers with the lowest per-unit cost, thus skewing export sales towards a small group of superstars. Indeed, the paper shows that the impact of the Internet on export concentration is stronger for the export of differentiated goods than for the export of homogenous goods, for which information friction is less salient.

This concentration of export sales in the Internet era has important implications for policymakers in Asia. Many Asian countries, especially developing ones, are focusing on the rapid expansion of the internet and digital transformation in general to expand their competitiveness and meet the SDGs. Understanding the dynamics of competition

in the digital age can help policymakers better design their export strategy and realise the goals of inclusive and sustainable industrialisation as in SDG Goal 9.

As competition structure in the internet era favours superstars, measures to support export entrepreneurship (through subsidising smaller firms to enter exporting) would yield limited benefits. Instead, export-promoting strategies may focus their limited resources on enhancing the linkages between export superstars and domestic suppliers. Emerging evidence from Uruguay and Costa Rica suggest that export-promoting activities can positively affect non-exporting firms through their linkage with supported exporters (Carballo, de Artiñano and Volpe-Martincus, 2021; Alfaro-Ureña, Manelici and Vasquez, 2022).

Furthermore, as the spread of the internet continues to expand across all countries, policymakers also need to pay attention to the link between exporter concentration and its consequences on the domestic economy. In particular, the surveillance of industry demand and global supply-chain dynamics affecting their superstar exports becomes an ever more important component for effective macroeconomic policymaking to realise the sustained growth outlined in SDG Goal 8.

## List of references

- Akerman, Anders, Ingvil Gaarder, and Magne Mogstad (2015). The skill complementarity of broadband internet. *The Quarterly Journal of Economics*, 130(4), 1781-1824.
- Akerman, Anders, Edwin Leuven, and Magne Mogstad (2022). Information frictions, internet, and the relationship between distance and trade. *American Economic Journal: Applied Economics*, 14(1), 133-63.
- Alfaro-Urena, Alonso, Isabela Manelici, and Jose P. Vasquez (2022). The effects of joining multinational supply chains: New evidence from firm-to-firm linkages. *The Quarterly Journal of Economics*, 137(3): 1495-1552.
- Allen, Thomas (2014). Information frictions in trade. *Econometrica*, 82(6), 2041-2083.
- Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare (2012). New trade models, same old gains?. *American Economic Review*, 102(1), 94-130.
- Bernard, A. B., and others (2022). The origins of firm heterogeneity: A production network approach. *Journal of Political Economy*, 130(7).
- Carballo, Jerónimo, Ignacio Marra de Artiñano and Christian Volpe-Martincus (2021). Information Frictions, Investment Promotion, and Multinational Production: Firm-Level Evidence. CESifo Working Paper Series 9043, CESifo.
- Cariolle, Joël, Jaime de Melo, and Michele Imbruno. (2022). Digital connectivity and firm participation in foreign markets: An exporter-based bilateral analysis. CEPR Discussion Paper DP17318.
- Chaney, Thomas (2008). Distorted gravity: the intensive and extensive margins of international trade. *American Economic Review*, 98(4), 1707-21.
- Chen, Maggie X., and Min Wu. (2020). The value of reputation in trade: Evidence from Alibaba. *The Review of Economics and Statistics*, 1-45.
- Dasgupta, Kunal, and Jordi Mondria. (2018). Inattentive importers. *Journal of International Economics*, 112, 150-165.
- De Loecker, Jan, Jan Eeckhout, and Gabriel Unger (2020). The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics*, 135(2), 561-644.
- Di Giovanni, Julian, and Andrei A. Levchenko. (2012). Country size, international trade, and aggregate fluctuations in granular economies. *Journal of Political Economy*, 120(6), 1083-1132.

- Di Giovanni, Julian, Andrei A. Levchenko, and Isabelle Mejean. (2014). Firms, destinations, and aggregate fluctuations. *Econometrica*, 82(4), 1303-1340.
- Di Giovanni, Julian, Andrei A. Levchenko, and Romain Ranciere (2011). Power laws in firm size and openness to trade: Measurement and implications. *Journal of International Economics*, 85(1), 42-52.
- Dickstein, M. J., & Morales, E. (2018). What do exporters know?. *The Quarterly Journal of Economics*, 133(4), 1753-1801.
- Eaton, Jonathan, Samuel Kortum, and Francis Kramarz (2011). An anatomy of international trade: Evidence from French firms. *Econometrica*, 79(5), 1453-1498.
- Fernandes, Ana M., Caroline Freund, and Martha Denisse Pierola (2016). Exporter behavior, country size and stage of development: Evidence from the exporter dynamics database. *Journal of Development Economics*, 119, 121-137.
- Fernandes, Ana M. and others (2023). The Intensive Margin in Trade: How Big and How Important?. *American Economic Journal: Macroeconomics*, 15 (3): 320-54.
- Fernandes, Ana M. and others (2019). The internet and Chinese exports in the pre-Alibaba era. *Journal of Development Economics*, 138, 57-76.
- Fort, Teresa C. (2017). Technology and production fragmentation: Domestic versus foreign sourcing. *The Review of Economic Studies*, 84(2), 650-687.
- Freund, Caroline, & Diana Weinhold (2004). The effect of the Internet on international trade. *Journal of international economics*, 62(1), 171-189.
- \_\_\_\_\_ (2002). The Internet and international trade in services. *American Economic Review*, 92(2), 236-240.
- Freund, Caroline, and Martha Denisse Pierola. (2020). The origins and dynamics of export superstars. *The World Bank Economic Review*, 34(1), 28-47.
- \_\_\_\_\_ (2015). "Export Superstars." *Review of Economics and Statistics* 97 (5): 1023-32.
- Gabaix, Xavier (2011). The granular origins of aggregate fluctuations. *Econometrica*, 79(3), 733-772.
- Gabaix, Xavier, and Rustam Ibragimov (2011), "Rank- 1/2: a simple way to improve the OLS estimation of tail exponents", *Journal of Business & Economic Statistics*, 29(1), 24-39.
- Gaubert, Cecile, and Oleg Itskhoki (2021). Granular comparative advantage. *Journal of Political Economy*, 129(3), 871-939.

- Gillwald, Alison, and Onkokame Mothobi (2019). After Access 2018: A demand-side view of mobile Internet from 10 African countries. *Research ICT Africa*.
- Head, Keith, Thierry Mayer, and Mathias Thoenig (2014). Welfare and trade without Pareto. *American Economic Review*, 104(5), 310-16.
- Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple (2004). Export versus FDI with heterogeneous firms. *American Economic Review*, 94(1), 300-316.
- Hjort, Jonas, and Jonas Poulsen (2019). The arrival of fast internet and employment in Africa. *American Economic Review*, 109(3), 1032-79.
- Hjort, Jonas, and Lin Tian (2022) The Economic Impact of Internet Connectivity in Developing Countries (mimeo).
- Lenoir, Clémence, Julien Martin, and Isabelle Mejean (2022) Search Frictions in International Goods Markets. *Journal of the European Economic Association*.
- González, Javier López, and Janos Ferencz (2018) "Digital Trade and Market Openness," OECD Trade Policy Papers 217, OECD Publishing.
- González, Javier López, and Marie-Agnes Jouanjean (2017-07-27), "Digital Trade: Developing a Framework for Analysis", OECD Trade Policy Papers, No. 205, OECD Publishing, Paris. <http://dx.doi.org/10.1787/524c8c83-en>
- Fernandes, Ana M., Caroline Freund, and Martha Denisse Pierola (2016), Exporter behavior, country size and stage of development: Evidence from the exporter dynamics database. *Journal of Development Economics*, 119, 121-137.
- Mayer, Thierry, and Gianmarco Ottaviano (2008). The happy few: The internationalisation of European firms. *Intereconomics*, 43(3), 135-148.
- Melitz, Marc J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6), 1695-1725.
- Nigai, Sergey (2017). A tale of two tails: Productivity distribution and the gains from trade. *Journal of International Economics*, 104, 44-62.
- Paunov, Caroline and others (2019). On the concentration of innovation in top cities in the digital age.
- Rauch, James E. (1999). Networks versus markets in international trade. *Journal of international Economics*, 48(1), 7-35.
- Sanderson, Eleanor, and Frank Windmeijer (2016). A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of econometrics*, 190(2), 212-221.

Steinwender, Claudia (2018). Real effects of information frictions: When the states and the kingdom became united. *American Economic Review*, 108(3), 657-96.

Tirole, Jean (2020). Competition and the industrial challenge for the digital age. *IFS Deaton Review on Inequalities*.

WTO (2021). *World Trade Report 2021: Economic resilience and trade*, WTO, Geneva.

## Appendix: Supplementary Materials

Table A.1: Coverage of Exporter Dynamic Database 2.0

Origin country	Region	First year	Last year
Albania	Europe	2004	2012
Burkina Faso	Africa	2005	2010
Bulgaria	Europe	2001	2006
Guatemala	LAC	2003	2010
Jordan	Asia	2003	2010
Mexico	LAC	2000	2007
Malawi	Africa	2004	2009
Peru	LAC	1993	2009
Senegal	Africa	2000	2010
Uruguay	LAC	2001	2012
Yemen	Asia	2008	2012

Table A.2: Log-log rank-sale regressions with lagged internet

### Panel A. Impact on log-log sales-rank regressions

	(1)	(2)	(3)
Variables	Top 10%	Top 50%	Full sample
Rank	-0.863 <sup>a</sup>	-1.300 <sup>a</sup>	-1.746 <sup>a</sup>
	[0.016]	[0.017]	[0.020]
Rank × Internet (lag)	-0.027 <sup>a</sup>	-0.041 <sup>a</sup>	-0.056 <sup>a</sup>
	[0.005]	[0.007]	[0.008]
Internet (lag)	0.192 <sup>a</sup>	0.295 <sup>a</sup>	0.393 <sup>a</sup>
	[0.070]	[0.067]	[0.074]
Observations	98,975	488,114	951,477
R-squared	0.958	0.946	0.921

Note: All regressions include origin and destination linear trends and country pair fixed effects. Robust standard errors clustered at origin-destination-year level are reported in brackets: <sup>a</sup> p<0.01; <sup>b</sup> p<0.05; <sup>c</sup> p<0.1.

**Panel B. Impact on superstar persistence**

	(1)	(2)	(3)	(4)	(5)
Variables	Top 5%	Top 10%	Top 1	Top 2	Top 5
Superstar (lag)	0.400 <sup>a</sup> [0.011]	0.417 <sup>a</sup> [0.009]	0.149 <sup>a</sup> [0.008]	0.110 <sup>a</sup> [0.007]	0.088 <sup>a</sup> [0.007]
Superstar (lag) × Internet (lag)	0.043 <sup>a</sup> [0.004]	0.032 <sup>a</sup> [0.004]	0.028 <sup>a</sup> [0.004]	0.039 <sup>a</sup> [0.003]	0.047 <sup>a</sup> [0.003]
Internet (lag)	-0.002 <sup>b</sup> [0.001]	-0.001 [0.002]	-0.003 <sup>a</sup> [0.001]	-0.006 <sup>a</sup> [0.001]	-0.012 <sup>a</sup> [0.002]
Observations	1,277,913	1,277,913	1,277,913	1,277,913	1,277,913
R-squared	0.463	0.485	0.283	0.328	0.384

Source: Author's calculations.

Note: All regressions include country-pair, origin and destination linear trends, firm and product fixed effects. Robust standard errors clustered at origin-destination-year level are reported in brackets: <sup>a</sup> p<0.01; <sup>b</sup> p<0.05; <sup>c</sup> p<0.1.

**Table A.3: Summary statistics for regressions on superstars**

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	N	mean	SD	median	min	max
Top 1	1358518	0.008	0.088	0.000	0.000	1.000
Top 2	1358518	0.014	0.118	0.000	0.000	1.000
Top 5	1358518	0.029	0.168	0.000	0.000	1.000
Top 5%	1358518	0.039	0.192	0.000	0.000	1.000
Top 10%	1358518	0.074	0.261	0.000	0.000	1.000
Internet (log)	1358518	2.270	1.121	2.488	-10.953	4.001

Source: Author's calculations.

**Table A.4: Impact on superstars, first stage regression using IV approach**

	Top 5%		Top 10%		Top 1		Top 2		Top 5	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Internet	Superstar <sub>t-1</sub> × Internet	Internet	Superstar <sub>t-1</sub> × Internet	Internet	Superstar <sub>t-1</sub> × Internet	Internet	Superstar <sub>t-1</sub> × Internet	Internet	Superstar <sub>t-1</sub> × Internet
Bandwidth	0.034 <sup>a</sup> (0.004)	-0.016 <sup>a</sup> (0.001)	0.034 <sup>a</sup> (0.004)	-0.029 <sup>a</sup> (0.002)	0.034 <sup>a</sup> (0.004)	-0.006 <sup>a</sup> (0.001)	0.034 <sup>a</sup> (0.004)	-0.009 <sup>a</sup> (0.001)	0.034 <sup>a</sup> (0.004)	-0.015 <sup>a</sup> (0.002)
Superstar <sub>t-1</sub>	-0.000	0.397 <sup>a</sup>	0.000	0.385 <sup>a</sup>	-0.000	0.511 <sup>a</sup>	-0.000	0.498 <sup>a</sup>	-0.000	0.477 <sup>a</sup>



	Top 5%		Top 10%		Top 1		Top 2		Top 5	
<b>x</b>	0				1		1		1	
<b>Band width</b>	(0.00 0)	(0.016)	(0.00 0)	(0.016)	(0.00 1)	(0.007)	(0.00 1)	(0.007)	(0.00 1)	(0.008)
<b>Superstar<sub>t-1</sub></b>	- 0.00 0	-0.394 <sup>a</sup>	- 0.00 1	-0.275 <sup>b</sup>	0.00 2	-1.678 <sup>a</sup>	0.00 5	-1.529 <sup>a</sup>	0.00 3	-1.289 <sup>a</sup>
	(0.00 1)	(0.125)	(0.00 1)	(0.126)	(0.00 5)	(0.055)	(0.00 5)	(0.055)	(0.00 4)	(0.057)
<b>N</b>	1226 696	12266 96	1226 696	12266 96	1226 696	12266 96	1226 696	12266 96	1226 696	12266 96
<b>SW F-stat</b>	79	603	80	586	80	4737	80	4577	80	3989

Source: Author's calculations.

Note: SW F-stat refers to Sanderson-Windmeijer F-statistics. All regressions include country-pair, origin and destination linear trends, firm and product fixed effects. Robust standard errors clustered at origin-destination-year level are reported in brackets: <sup>a</sup> p<0.01; <sup>b</sup> p<0.05; <sup>c</sup> p<0.

**Table A.5: Impact of internet on superstars, alternative fixed effects**

	(1)	(2)	(3)	(4)
<b>Variables</b>	<b>Top 5%</b>	<b>Top 5%</b>	<b>Top 1</b>	<b>Top 1</b>
<b>Superstar (lag)</b>	0.299 <sup>a</sup>	0.399 <sup>a</sup>	-0.003	0.187 <sup>a</sup>
	[0.016]	[0.013]	[0.011]	[0.011]
<b>Superstar (lag) × Internet</b>	0.051 <sup>a</sup>	0.044 <sup>a</sup>	0.041 <sup>a</sup>	0.029 <sup>a</sup>
	[0.006]	[0.005]	[0.005]	[0.005]
<b>Internet</b>	-0.003		-0.001	
	[0.005]		[0.004]	
<b>Observations</b>	1,278,212	1,283,835	1,278,212	1,283,835
<b>R-squared</b>	0.509	0.474	0.396	0.281
<b>Fixed effects</b>				
<b>Origin-Destination</b>	Y		Y	
<b>Origin-Year</b>	Y		Y	
<b>Destination-Year</b>	Y		Y	
<b>Origin-Destination-Year</b>		Y		Y
<b>Firm</b>	Y	Y	Y	Y
<b>Product</b>		Y		Y
<b>Product-Destination</b>	Y		Y	

Note: Robust standard errors clustered at origin-destination-year level are reported in brackets: <sup>a</sup> p<0.01; <sup>b</sup> p<0.05; <sup>c</sup> p<0.1.

**Table A.6: The differential impact of Internet adoption on export concentration, split sample regressions**

**Panel A. Impact on log-log rank-sale relationship**

	Differentiated goods			Homogeneous goods		
	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Top 10%	Top 50%	Full sample	Top 10%	Top 50%	Full sample
<b>Rank</b>	-0.859 <sup>a</sup>	-1.313 <sup>a</sup>	-1.774 <sup>a</sup>	-0.907 <sup>a</sup>	-1.293 <sup>a</sup>	-1.832 <sup>a</sup>
	[0.005]	[0.003]	[0.003]	[0.008]	[0.004]	[0.004]
<b>Rank × Internet</b>	-0.037 <sup>a</sup>	-0.069 <sup>a</sup>	-0.091 <sup>a</sup>	-0.021 <sup>a</sup>	-0.031 <sup>a</sup>	-0.051 <sup>a</sup>
	[0.002]	[0.001]	[0.001]	[0.003]	[0.001]	[0.001]
<b>Internet</b>	0.208 <sup>a</sup>	0.398 <sup>a</sup>	0.575 <sup>a</sup>	0.146 <sup>a</sup>	0.160 <sup>a</sup>	0.267 <sup>a</sup>
	[0.012]	[0.007]	[0.007]	[0.019]	[0.009]	[0.009]
<b>Observations</b>	80,535	398,143	796,473	36,498	178,037	356,258
<b>R-squared</b>	0.937	0.885	0.856	0.863	0.877	0.829

Note: All regressions include country-pair, origin and destination linear trends and firm fixed effects. Robust standard errors clustered at origin-destination-year level are reported in brackets: <sup>a</sup> p<0.01; <sup>b</sup> p<0.05; <sup>c</sup> p<0.1.

**Panel B. Impact on superstars**

	(1)	(2)	(3)	(4)
	Differentiated goods	Homogeneous goods	Differentiated goods	Homogeneous goods
Superstar definition →	Top 5% <sub>ijt</sub>	Top 5% <sub>ijt</sub>	Top 1 <sub>ijt</sub>	Top 1 <sub>ijt</sub>
<b>Superstar (lag)</b>	0.177 <sup>a</sup>	0.141 <sup>a</sup>	-0.062	0.097 <sup>b</sup>
	[0.028]	[0.034]	[0.041]	[0.042]
<b>Superstar (lag) × Internet</b>	0.057 <sup>a</sup>	0.044 <sup>a</sup>	0.045 <sup>a</sup>	0.022 <sup>a</sup>
	[0.005]	[0.006]	[0.008]	[0.008]
<b>Internet</b>	-0.006 <sup>a</sup>	-0.009 <sup>a</sup>	-0.005 <sup>a</sup>	-0.007 <sup>a</sup>
	[0.001]	[0.002]	[0.001]	[0.001]
<b>Observations</b>	943,661	413,001	943,661	413,001
<b>R-squared</b>	0.462	0.415	0.263	0.317

Source: Author's calculations.

Note: All regressions include country-pair, origin and destination linear trends, firm and main-product fixed effects. Robust standard errors clustered at origin-destination-year level are reported in brackets: <sup>a</sup> p<0.01; <sup>b</sup> p<0.05; <sup>c</sup> p<0.1.

**Table A.7: Robustness check with inclusion of firm-year fixed effects**

**Panel A. Impact on log-log sale-rank regressions**

	(1)	(2)	(3)
Variables	Top 10%	Top 50%	Full sample
Rank	-0.847 <sup>a</sup>	-1.253 <sup>a</sup>	-1.725 <sup>a</sup>
	[0.011]	[0.010]	[0.016]
Rank × Internet	-0.026 <sup>a</sup>	-0.040 <sup>a</sup>	-0.059 <sup>a</sup>
	[0.005]	[0.005]	[0.007]
Internet	0.169 <sup>b</sup>	0.243 <sup>a</sup>	0.366 <sup>a</sup>
	[0.085]	[0.070]	[0.083]
Observations	65,576	352,429	958,158
R-squared	0.950	0.943	0.921

Note: All regressions include origin and destination linear trends, country pair and firm-year fixed effects. Robust standard errors clustered at origin-destination-year level are reported in brackets: <sup>a</sup> p<0.01; <sup>b</sup> p<0.05; <sup>c</sup> p<0.1.

**Panel B. Impact on superstars**

	(1)	(2)	(3)	(4)	(5)
Variables	Top 5%	Top 10%	Top 1	Top 2	Top 5
Superstar (lag)	0.395 <sup>a</sup>	0.427 <sup>a</sup>	0.175 <sup>a</sup>	0.128 <sup>a</sup>	0.101 <sup>a</sup>
	[0.011]	[0.009]	[0.010]	[0.009]	[0.009]
Superstar (lag) × Internet	0.064 <sup>a</sup>	0.053 <sup>a</sup>	0.037 <sup>a</sup>	0.053 <sup>a</sup>	0.067 <sup>a</sup>
	[0.005]	[0.004]	[0.004]	[0.004]	[0.004]
Internet	-0.016 <sup>a</sup>	-0.017 <sup>a</sup>	-0.008 <sup>c</sup>	-0.013 <sup>b</sup>	-0.010
	[0.006]	[0.007]	[0.004]	[0.006]	[0.008]
Observations	909,627	909,627	909,627	909,627	909,627
R-squared	0.507	0.533	0.353	0.400	0.456

Source: Author's calculations.

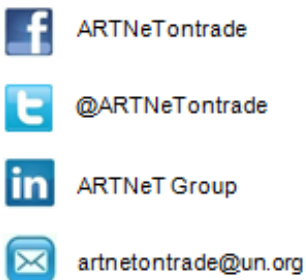
Note: All regressions include origin and destination linear trends, country pair, product and firm-year fixed effects. Robust standard errors clustered at origin-destination-year level are reported in brackets: <sup>a</sup> p<0.01; <sup>b</sup> p<0.05; <sup>c</sup> p<0.1.



The Asia-Pacific Research and Training Network on Trade - ARTNeT - is an open network of research and academic institutions and think-tanks in the Asia-Pacific region. Since its inception, ARTNeT aims to increase the amount of high quality, topical and applied research in the region by harnessing existent research capacity and developing new capacities. ARTNeT also focuses on communicating these research outputs for policymaking in the region including through the ARTNeT Working Paper Series which provide new and policy-relevant research on topics related to trade, investment and development. The views expressed in this publication are those of the authors and do not necessarily reflect the views of the United Nations and ARTNeT secretariat or ARTNeT members.

Readers are encouraged to quote or reproduce material from ARTNeT Working Papers for their own publications, but as the copyright holder, ARTNeT requests due acknowledgement and a copy of the publication.

This and other ARTNeT publications are available from [artnet.unescap.org](http://artnet.unescap.org)



ARTNeT Secretariat, United Nations ESCAP

Rajadamnern Nok Avenue

Bangkok 10200, Thailand

Tel: +66(0) 22881425

Fax: +66(0) 22881027