No 2025-11 – September 2025

Working Paper

CEPII

Deep Trade Agreements and Heterogeneous Firms' Exports

Matteo Neri-Lainé, Gianluca Orefice & Michele Ruta

Highlights

- The signature of Deep Trade Agreement boosts the export of large and GVC firms.
- Small and less productive firms suffer the increased competition at destination.
- The re-allocation of resources from small to large firms may imply positive welfare effect for developing countries signing Deep Trade Agreements.

Abstract

This paper studies the effect of regional trade agreements on firms' exports using detailed information on the content of trade agreements and firm-level exports for 31 developing countries between 2000 and 2020. Moving from shallow to deep trade agreements - *i.e.* agreements that regulate border and behind-the-border policies - boosts firms' exports, on average, by 3.6 percent. In line with models of trade with heterogeneous firms, this effect is stronger for large firms and firms involved in global value chains and (weakly) negative for small firms, suggesting a competition effect of deep trade agreements with significant welfare consequences for signatory countries. An Instrumental Variable strategy and a battery of robustness tests confirm the causal interpretation of the results.

Keywords

Regional Trade Agreements, Exports, Firm Heterogeneity, Developing Countries.

JEL

F13, F14, F15.

Working Paper

CEPII

CEPII Working Paper
Contributing to research in international economics

CEPII (Centre d'Études Prospectives et d'Informations Internationales) is a French institute dedicated to producing independent, policy-oriented economic research helpful to understand the international economic environment and challenges in the areas of trade policy, competitiveness, macroeconomics, international finance and growth.

Editorial Director:
Antoine Bouët

Visual design and production:
Laure Boivin

ISSN 2970-491X

September 2025

To subscribe to
The CEPII Newsletter:
www.cepii.fr/KeepInformed

Centre d'études prospectives
et d'informations internationales
20, avenue de Ségur
TSA 10726
75334 Paris Cedex 07

contact@cepii.fr
www.cepii.fr – @CEPII_Paris
Press contact: presse@cepii.fr

Deep Trade Agreements and Heterogeneous Firms' Exports*

MATTEO NERI--LAINÉ† GIANLUCA OREFICE‡ MICHELE RUTA§

February 20, 2025

1 Introduction

Since the early 2000s, many developing countries have signed Regional Trade Agreements (RTAs) in an attempt to better integrate their economies into regional and global markets and improve the export performance of their firms – see Freund and Ornelas (2010); Limão (2016); Fernandes et al. (2021). In particular, out of the 190 RTAs signed during the period 2000-2020, 175 involved at least one developing country as a member. Over the same period, the content of RTAs has widely changed (Hofmann et al., 2017; Mattoo et al., 2020). The average RTA signed by a developing country at the end of the period covered roughly 30 percent more policy areas (provisions) than at the beginning of the period, with behind-the-border issues like technical and sanitary standards, investment, and intellectual property rights protection being added to the more traditional areas covered by RTAs such as tariff liberalization and the reduction of other border barriers. We refer to these more complex trade agreements that cover behind-the-border policies as "deep RTAs" or "Deep Trade Agreements" (DTAs).

Starting with the seminal work by Baier and Bergstrand (2007), the effects of regional trade

*We would like to thank Ana Fernandes, Swati Dhingra, Andrea Lassmann, Nuno Limao, Luca Macedoni, Mathieu Parenti, Philip Sauré, Chahir Zaki and seminar participants at World Bank, Indiana University, University of Mainz, Aarhus University, Paris School of Economics, University of Granada, the World Trade Forum 2023 and the European Trade Study Group for helpful comments and suggestions and to Ana Fernandes for help with the data. The views expressed in this paper are the authors' and do not necessarily reflect those of the institutions they belong to.

†Paris School of Economics, 48 boulevard Jourdan 75014 Paris, France.

‡University Paris-Dauphine PSL, Place du Maréchal de Lattre de Tassigny Paris, France. *Corresp. author.*

§International Monetary Fund, 700 19th Street N.W., Washington, D.C. 20431, USA.

agreements on trade have been largely studied. In part motivated by the changing depth of RTAs, a number of papers have empirically investigated the effects of different types of trade agreements on international trade and other related outcomes. A key finding of this literature is that the heterogeneity of trade agreements matters (Baier et al., 2018, 2019), and that the depth of RTAs and the specific provisions therein can help to explain different trade outcomes (Orefice and Rocha, 2014; Mattoo et al., 2022; Dhingra et al., 2018; Breinlich et al., 2022).

At the same time, there remain important gaps in the literature on RTAs. In particular, theoretical trade models have emphasized that a common reduction in trade costs affects different types of firms heterogeneously. Larger and more productive firms can more easily take advantage of a reduction in trade costs than smaller and less productive firms *via* a productivity channel (Melitz, 2003) and/or a heterogeneous mark-up and pricing-to-market behaviour (Melitz and Ottaviano, 2008; Atkeson and Burstein, 2008). As deep RTAs can reduce trade costs among members by eliminating tariffs and by reducing other behind-the-border frictions due to regulatory differences and/or policy uncertainty, an important unexplored question is how these agreements affect firms with different characteristics.[1] Specifically, if firms are heterogeneous, we should expect deep RTAs to favor large and highly productive firms' exports at the expense of low-productivity firms. This may lead to a strong *competition effect* that can potentially imply the exit of small and less productive firms – as shown in Mrázová and Neary (2017; 2020) and Crowley et al. (2024).[2] In this case, as resources reallocate from low to high productivity firms, developing economies signing a deep RTA would experience adjustment costs and an improvement in average productivity. Hence, studying heterogeneous firms' exports and the competition effect of deep RTAs is of primary interest to fully understand the welfare consequence of trade agreements for developing countries.

As a first step into this new research question, this paper empirically investigates the effect of RTAs' depth (here approximated by the *number* of provisions therein) on the export performance of heterogeneous firms. The analysis combines firm-level export data for 31 developing countries from the World Bank Exporter Dynamics Database (Fernandes et al., 2016) and information on the detailed content of more than 300 RTAs from the World Bank Deep Trade Agreements database (Hofmann et al., 2017). The richness of the data allows us to precisely take into account the heterogeneous consequences of deep RTAs on the exports of different firms. We use a decomposition à la Berman et al. (2012) to study the extent to which the effects of

[1]The effect of Deep Trade Agreement can be also heterogeneous across different types of signatory countries. Parenti and Gonzague (2022) show that regulatory cooperation in trade agreement is more beneficial for countries with intermediate differences in regulatory preferences.

[2]In line with this argument, recent contributions by Blanga-Gubbay et al. (2021) and Blanga-Gubbay et al. (2024) show that lobbying on RTAs is dominated by large multinational firms, which are in favor of these agreements and lobby for the inclusion of deep provisions in RTAs.

deep RTAs on trade come from the intensive and extensive margins of trade. Importantly, we carefully address the endogeneity of trade and deep RTAs through different methods. We use an Instrumental Variable approach based on the domino effect of RTAs (Baldwin and Jaimovich, 2012), a plausible exogeneity test à la Conley et al. (2012), and we employ an event study approach to check the parallel trend assumption and confirm the validity of the key results and their causal interpretation.

To assess the impact of RTAs' depth on the export performance of firms, we first adapt the standard gravity model for trade (Head and Mayer, 2014) to firm-level analysis and include a variable capturing the depth of RTAs. Using information from the Deep Trade Agreements database, we construct different measures of RTA depth based on the number (and type) of policy areas covered by the agreements and their legal enforceability. Our specification allows us to identify the impact on firms' exports of a *change* in the depth of RTAs (i.e. newly signed RTAs or amendment of pre-existing ones) between two countries, controlling for any firm-year and country-specific factor that may affect the export performance of firms. As in Crowley et al. (2024), the structure of fixed effects included in our estimations allows to interpret the results (also) in terms of the impact of deep RTAs on *market-shares*, and hence on the re-allocation of market shares across firms of different size and productivity. The baseline results show that one additional legally enforceable policy area in RTAs boosts the exports (and the market-share) of the average firm by 0.3%; corresponding to a 3.6% increase in firms' exports when moving from *shallow* to *deep* RTAs.[3] Importantly, this effect adds to the more standard tariff-cut effect of RTAs on trade (here captured by applied tariff variable), and it is larger for policy areas in RTAs that are not regulated by the World Trade Organization (WTO-extra), such as investment and competition provisions. One additional legally enforceable WTO-extra provision increases the firm's exports by 0.7%.

As discussed above, this average effect of RTAs' depth on the export performance of firms may hide substantial heterogeneity across firms with different characteristics: large and highly productive firms are expected to benefit the most from deeper RTAs at the expenses of low-productivity firms. Moreover, in the presence of variable mark-ups or heterogeneous elasticity of demand, the competition effect induced by deep RTAs may imply the re-allocation of destination-specific market shares from less to more productive exporters (see Atkeson and Burstein 2008, Mayer et al. 2021, Dhingra and Morrow 2019, Crowley et al. 2024 and Edmond et al. 2015). Our second set of estimations therefore tests the effect of RTAs' depth on heterogeneous ex-

[3] We consider *shallow* those RTAs including only the two tariff-related provisions that are always included in RTAs (i.e. tariff cut on agriculture and industrial sectors), and *deep* those RTAs including 14 provisions (i.e. the 75 percentile in the depth of RTAs).

porters.[4] Results point to strong heterogeneous effects of RTAs' depth on firms with different characteristics. While large and high-productivity firms benefit from deep RTAs (i.e. from a larger number of provisions therein), small and less productive firms suffer the increased within-origin competition at destination (i.e. re-allocation of sales and market shares from small to large firms). On average, including *one* additional legally enforceable provision in RTAs stimulates the exports of large firms by 0.4%-0.6%, while it reduces the exports of small firms by 0.4%-0.5%.[5] Interestingly, and in line with the argument in Limão (2016),[6] high-productivity firms participating in Global Value Chains (GVC) benefit the most from deep RTAs. For these firms, an additional legally enforceable provision in RTAs stimulates exports by 0.7%. The competition effect of deep trade agreements, through the re-allocation of market shares (and hence resources) from low to high-productivity firms, shows that using firm-level data is key to understand the welfare effects of deep trade agreements in developing countries.

The heterogeneous trade effects of deep trade agreements have also interesting consequences at the aggregate level. First, the re-allocation of market shares from low- to high-productivity firms implies the exit of low-productivity firms from the destination market. Accordingly, we show that an increase in the depth of trade agreements reduces the number of firms that keep exporting at destination, but increases their average exports. Second, if only high-productivity firms benefit from deeper RTAs, origin countries with a more dispersed distribution of firms' size (i.e. a larger share of high-productivity firms) will benefit relatively more from deep RTAs. Consistently with this line of reasoning, we show that origin countries having a less skewed distribution of firm size benefit more (less) on the intensive (extensive) margin of trade.

We next investigate how deep trade agreements affect the margin of trade at the firm level. Deep RTAs could reduce the variable and fixed trade costs between member countries. This, in turn impacts on the margin of trade that is affected. Previous work by Baier et al. (2014) uses aggregate trade data to study how different types of RTAs affect the intensive and extensive trade margin. Here, we use a decomposition approach à la Berman et al. (2012), to disentangle the average export sales effect of deep RTAs into intensive *versus* extensive (firm-based) margin component. Namely, we use the firm-specific export behavior to impute the extensive *versus* intensive margin effect of deep RTAs. We show that the extensive margin channel has only

[4] For our large sample of developing countries, we do not have data on the balance sheet of firms to compute direct measures of productivity, so we resort to several proxies discussed in detail in section 4.2.

[5] This finding is the average effect of deep trade agreements on firms' exports. It is still possible that individual provisions in DTAs have effects that are more favorable to small firms. For instance, Fernades et al. (2021) find that for a sample of Latin American countries, firms' exports increase significantly in destination markets with RTAs that promote regulatory cooperation and that the effect is stronger for smaller firms. They find that this effect is driven by entry into new product markets and increases in the export quality of smaller firms.

[6] In discussing the economic consequences of deep RTAs, Limão (2016) highlights that "*Reducing NTBs can be particularly important when firms rely heavily on intermediates and/or can rearrange their production structure across borders [...].*"

a slightly larger contribution (53%) than the intensive margin channel (46%). The further contribution of this paper is therefore to provide novel, firm-level evidence on the margins through which deep RTAs boost aggregate exports.

Finally, we study the dynamic effects of deep trade agreements on firms' exports. By adopting an event study approach in the vein of Fajgelbaum et al. (2020), we show that a *change* in the depth of RTAs makes firms exporting more during the three years after the shock. In line with intuition, the effect of the change in depth of RTAs increases over time. The event study approach also shows the validity of the parallel trend assumption and reinforces our baseline results' causal interpretation.

This paper falls at the intersection of two strands of literature in international trade and trade policy. The first is the literature on heterogeneous firms in trade. Existing research shows theoretically that only high-productivity firms benefit from larger market access at destination (Melitz, 2003; Atkeson and Burstein, 2008; Melitz and Ottaviano, 2008). A large number of studies provide empirical evidence of such theoretical prediction (Bernard et al., 2007, 2009, 2018; Freund and Pierola, 2015). The second branch of the literature is the one that investigates the trade effects of regional trade agreements (Baier and Bergstrand, 2007; Baier et al., 2019; Nagengast and Yotov, 2024). Recent research shows that RTAs, especially deeper trade agreements that cover behind-the-border policies, provide an anchor to lower trade barriers and reduce policy frictions, leading to a strong impact on trade flows (Dhingra et al., 2018; Fernades et al., 2021; Crowley et al., 2024; Mattoo et al., 2022). By exploiting firm-level data for a large number of developing countries and detailed information on the content of trade agreements, this paper contributes to these two branches of the literature by showing that firms with different characteristics are affected differently by deeper trade agreements. Specifically, we uncover that deep trade agreements promote more the exports of large and more productive firms and GVC firms relative to small and less productive firms, leading to a competition effect that has relevant consequences for welfare, and it is documented for the first time in this paper.

The rest of the paper is organized as follows. Section 2 presents the data used in the analysis and provides some descriptive evidence. Section 3 discusses the empirical strategy. Section 4 shows the baseline results and the heterogeneity tests. In section 5 we address the potential endogneneity problem. Section 6 tests the dynamic effect of deep RTAs. Section 7 shows the effect of DTAs across different sectors. Concluding remarks follow.

2 Data and descriptive evidence

Our empirical analysis is based on two main World Bank data sources: (i) the Exporter Dynamics Database (Fernandes et al., 2016) providing firm-level exports for 55 developing countries in the period 1996-2020,[7] and (ii) the Content of Deep Trade Agreements (Hofmann et al., 2017) on the policy areas—i.e. the provisions—covered in RTAs in force and notified to the WTO for the period 2000-2020. Moreover, we complete our dataset by including gravity-related variables (such as distance, common border, language, etc.) from the CEPII gravity database, and data on the effectively applied tariffs faced by each exporter at the destination from the MacMap (CEPII) database.[8]

We build our final dataset in several steps. First, we reshape the original DTA database into a country pair-year-specific dataset. Indeed, since a given pair of countries (origin and destination) may have two (or more) RTAs contemporaneously in force with possibly different coverage of provisions, we consider the maximum value of each provision dummy across multiple RTAs (if any) within each country pair-year combination.

As a second step, we construct several measures of RTAs' depth based on the *number* (i.e. count) of provisions included in the agreement that each country pair shares in a given year. Specifically, we consider different count variables, differentiating by type of provision: (i) number of provisions (no matter their legal enforceability), (ii) number of legally enforceable provisions (i.e. whose implementation is supported by the strong legal language and by the availability of a dispute settlement mechanism – see Hofmann et al.2017), (iii) number of WTO-plus provisions (WTO+) – i.e. provisions covered by the current mandate of the WTO, (iv) number of WTO-extra provisions (WTO-X) – i.e. provisions not covered by the current mandate of the WTO, (iv) number of core provisions – i.e. those directly related to trade enhancing factors, (v) number of provisions regulating non-tariff measures (SPS and TBT), and (vi) number of Policy provisions – i.e. services, investment and competition policy provisions (Dhingra et al., 2018).[9]

While in our empirical exercise we use all these proxies for the depth of RTAs, our baseline measure of RTA depth is the count of legally enforceable provisions in each RTA.[10] The simple

[7]We thanks the World Bank for kindly providing confidential firm-level EDD data.

[8]MacMap (CEPII) database is available here: `https://www.cepii.fr/CEPII/fr/bdd_modele/bdd_modele_item.asp?id=12`.

[9]See Horn et al. (2010) for definition of WTO+ and WTO-X provisions. Core provisions are defined in Hofmann et al. (2017) and include all WTO+ provisions plus clauses that regulate competition policy, bilateral investment, movement of capital and intellectual property rights. Provisions related to non-tariff measures concern the harmonization or mutual recognition of technical standards (TBT) and/or sanitary and phyto-sanitary measures (SPS). Following Dhingra et al. (2018), policy-related provisions are those concerning services, investment and competition policy.

[10]It should be noted that in this paper we aim to capture the impact of the *overall* depth of RTAs on firms with different characteristics. However, different provisions in trade agreements have different effects on aggregate trade outcomes (Fernandes et al., 2021) and are likely to have heterogeneous effects on firms' exports. For example,

count of provisions in RTAs gives equal weight to each clause included in the agreement, and one may want to give more weight to "rare" provisions (i.e. provisions included less frequently in RTAs and likely to signal stronger market integration). To address this concern, as a robustness check, we also use alternative measures of RTA depth based on the *weighted* sum of provisions (using one minus the frequency to which each provision appears in the 300 RTAs covered by the DTA database as a weight).[11] In Figure A1 we show how frequently each provision is included in Deep Trade Agreements.

The Exporter Dynamics Database (EDD) takes the form of multiple databases: one per country.[12] So, as a last step in the construction of the final data set, we add to each country-specific EDD dataset: (i) the gravity variables from the CEPII gravity database, and (ii) the tariffs data from MacMap (CEPII). Finally, we pool all the 55 country-specific databases to get a single firm-level dataset covering all the 55 developing countries available. This dataset initially contains 29,009,865 observations (firm-destination-product-year specific) spanning from the late 1990s to 2020. However, the DTA database covers only the period 2000-2020; so we keep only firm export data for this period. Also, we exclude highly concentrated Oil and Mineral sectors where monopolistic competition based trade models do not apply.[13] Since our main variable of interest is not sector specific (i.e. the count of provisions in RTAs is country pair-year specific), we first aggregate data at firm-destination-year level and run our baseline estimations using firm-destination-year specific data. However, in the last part of our empirical exercise we resort to firm-destination-*sector*-year level data to explore the possible heterogeneous effect of deep RTAs across different sectors.

As described in Table 1, the majority of exporting countries covered by the EDD database have at least one active RTA in force (i.e. 48 out of 55). However, only a sub-sample of countries (and country-pairs) experienced a change in the depth of their RTAs in the period 2000-2020 (i.e. a newly signed RTA or amendment of pre-existing RTAs). In particular, 31 exporting countries signed new RTAs or amended a pre-existing one, giving *within* variation to our measures of RTAs

provisions that lower fixed entry costs on the destination market, such as provisions aiming at improving trade facilitation or reducing regulatory divergence, can make it easier for small firms to export. Fernades et al. (2021) finds some preliminary evidence of this, focusing on provisions on technical and sanitary standards in trade agreements involving a sample of Latin American countries. Other provisions may have just the opposite effect. For instance, requirements in DTAs to meet higher environmental or labour standards could make it easier for larger exporters relative to smaller firms to export in the destination market, reinforcing the competition effect that we stress in this paper.

[11] An alternative approach would be using machine learning (data-driven) methods for selecting the most important provisions in affecting trade. This approach is adopted by Breinlich et al. (2022) who find that provisions related to antidumping, competition policy, technical barriers to trade, and trade facilitation are associated with a strong trade-increasing effect.

[12] In case of breaks in the firms' identifiers over the time period covered, the Exporter Dynamics Database contains two separate data sets per country (before *versus* after the break in the firms' identifier). In case of breaks in the firms' identifiers, we keep the most recent period (i.e. after-break period).

[13] However, our baseline results do not change by including Oil and Mineral sector.

depth. This translates into 701 country-pairs having time variation in the depth of RTAs in the period 2000-2020. This time variation comes mainly from the entry into force of new trade agreements. Since in the empirical strategy we rely on the within country-pair variation in RTAs' depth – see Section 3 – the final estimation sample considers only 31 exporting countries and contains 4,578,173 firm-destination-year specific observations.[14] Also, to reduce any bias from occasional exporters and noisy firm-destination level exports, we keep only those firms that exported to a certain market for at least four years within the time period. Finally, our estimation sample shrinks to 1,534,360 observations because of missing values in the tariff data.

Table 1: Number of countries and country-pairs with variation in RTAs' depth in the period 2000-2020.

# of exporting countries having at least one trade-partner with :	
constant RTA depth	*changing* RTA depth
48	31
# of country-pairs with:	
constant RTA depth	*changing* RTA depth
1004	701

Notes: Authors' calculation on World Bank Content of Deep Trade Agreement data.

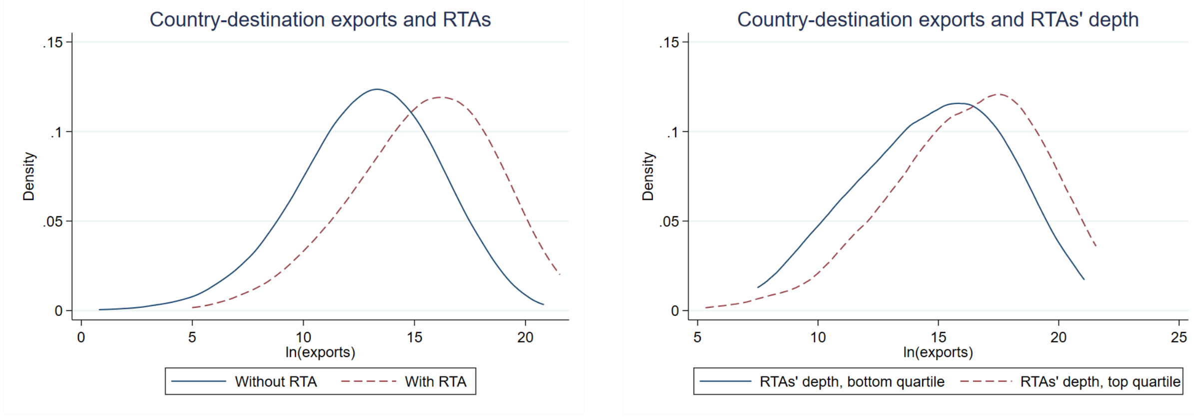
Among the 31 exporting countries having *within* variation in the RTA depth values, 15 countries account for 87% of the overall variation in the RTA depth. These are: Croatia, Georgia, Colombia, Chile, South Africa, Serbia, Slovenia, Peru, Mauritius, Guatemala, Madagascar, Nicaragua, Tanzania, Ecuador, and Malawi. In Figure 1 we show the empirical distribution of country-pair exports (panel a and b), for the sub-sample of the 15 countries that provide much of the variation in RTA depth, in the last available year.[15] In panel (a) of Figure 1 we show the empirical distribution of bilateral exports for country-pairs with and without RTA in force (respectively dash and continuous line in the figure), while in panel (b) of Figure 1 we show the empirical distribution of bilateral exports for pairs having deep *versus* shallow RTA in force (i.e. RTA with number of provisions respectively below the 25^{th} and above the 75^{th} percentile of distribution). The presence of an RTA and its depth matter for the aggregate country-pair exports.

The same stylized facts emerge from firm-level regression results reported in Table 2. While the presence of a RTA (no matter its depth) has positive and significant effect on the exports of

[14] It must be noted that the coverage of the EDD varies by country. While for some countries we have data for the entire period 2000-2020, for other countries we have a shorter time period. In Table A1 we report the period covered by the set of countries covered in our estimations (i.e. those having within variation in RTAs' depth).

[15] In Figure A2 we provide the same evidence for the full set of 31 countries with time variation in RTA depth.

Figure 1: Export values and RTAs, major treated countries.



(a) With vs. without RTA

(b) Shallow vs deep RTA

firms (see RTA_{ijt} coefficients in columns 1-2), the depth of RTAs matters.[16] If the RTA has a number of provisions above the median, it has an even stronger effect on firms' exports. Namely, while the presence of a RTA increases exports by 6-7%, a deep RTA boosts firms' exports by 11%. Reassuringly, these results are in line with a recent paper by Nagengast and Yotov (2024) showing that the signature of a deep RTA (i.e. with more than median number of provisions) boosts *aggregate* export by 0.18 log points.[17]

Finally, in Table A2 we report the in-sample descriptive statistics for the variables included in our econometric exercise. It clearly emerges the large variation in the depth of RTAs across country-pairs. Regardless of the type of provisions considered to define the RTAs' depth (i.e. count of provisions), the standard deviation is almost equal to the average value. We exploit the variability in the depth of RTAs in the econometric exercise reported in the next section.

3 Identification strategy

This section discusses the empirical strategy adopted to test the effect of RTAs' depth on the exports of firms. Since we are primarily interested in testing how trade agreements affect bilateral trade flows we start by estimating the following firm-destination-year specific equation:

$$X_{fijt} = exp\left[\theta_{ft} + \ \theta_{jt} + \theta_{ij} + \beta_1 DTA_{ijt} + \beta_2 ln\left(1 + \tau_{fijt}\right)\right] \times \varepsilon_{fijt} \tag{1}$$

where the subscripts f, i, j, and t stand respectively for firm, origin country (i.e. the country where the exporting firm is located), destination country and year. The explanatory variable of interest – DTA_{ijt} – is the depth of the RTA (if any) that country i has with destination j

[16]Regressions in Table 2 include firm-year, destination-year and origin-destination fixed effects.
[17]See TWFE estimation results in Table 10 in Nagengast and Yotov (2024).

Table 2: Firms' exports and deep vs shallow RTAs.

	Exp	Exp
	(1)	(2)
RTA_{ijt}	0.068***	0.055***
	(0.019)	(0.020)
RTA × Dummy=1 if DTA_{ijt} > Mdn		0.043*
		(0.023)
$\text{Ln}(1+\tau_{ijt})$	-0.578***	-0.579***
	(0.073)	(0.073)
Firm-Year FE	Yes	Yes
Destination-Year FE	Yes	Yes
Origin-Destination FE	Yes	Yes
Observations	1,534,360	1,534,360

Notes: PPML estimates. Standard errors in parenthesis are clustered at firm-destination and origin-destination-year level. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

at time t. As a first coarse measure of RTA depth, we use the count of any type of provisions included in the RTA. However, our preferred measure of RTA depth is the count of legally enforceable provisions. This measure is then refined and we use the count of WTO+, WTO-X, "core", NTM-related and Policy oriented legally enforceable provisions in RTAs, as discussed in Section 2. The firm-year fixed effects (θ_{ft}) control for any unobserved time-variant firm-specific characteristics, such as productivity shocks, size and workforce composition (i.e. quality of the management, etc.). Since each firm is unambiguously located in a country i, the firm-year fixed effects subsume origin-year fixed effects and capture any origin country-year specific shock affecting the export performances of all firms in country i (i.e. degree of competition and the multilateral resistance term on the exporter side). Any country-specific technological or productivity shock, as well as the distribution of firms' productivity in a given exporting country, are captured by firm-year fixed effects. Country-pair (time-invariant) factor, such as geographical distance and any other gravity-type covariate is captured by country-pair fixed effects θ_{ij}. In equation (1) we also control for the multilateral resistance term of the importer country by including destination-year fixed effects θ_{jt}, capturing also any demand shock that affects the import demand at destination. Hence, β_1 can be interpreted as the elasticity of a firm's export (or market share) in a destination to RTA's depth.[18] We adopt a PPML estimator to address the heteroskedasticity problem in structural gravity model for trade (Santos-Silva

[18] Indeed, after log-linearizing and rearranging the market share of firm f in a destination d, i.e. $mktsh = X_{fijt}/\sum_i \sum_f X_{fijt}$, the export sales of a given firm can be expressed as:

$$\ln(X_{fijt}) = \ln(mktsh_{fijt}) + \ln\left(\sum_i \sum_f X_{fijt}\right)$$

Since the last term of this equation, i.e. the total imports of j, is captured by fixed effects, our results can be interpreted in terms of firm's market share at destination. The same point is made by Crowley et al. (2024).

and Tenreyro, 2006), and cluster standard errors at firm-destination (to account for the serial correlation in firm-destination exports) and origin-destination-year level (i.e. the source variation of our main variable of interest).[19]

Given the set of fixed effects included in equation (1) our variable of interest DTA_{ijt} is identified on the *change* in the depth of RTAs (i.e. newly signed RTAs or amendment of pre-existing ones) between country i and j, controlling for any firm- and country-specific factor that may affect the export performances of firms. Specifically, we compare a given firm's exports (from a given origin country) towards destinations with *versus* without changes in RTA's depth (conditional on any firm- and destination-specific shock). The omitted variable concern is therefore very reduced here. Moreover, the concern that a specific firm in country i may affect the signature and the content of a trade agreement between country i and j is in general remote, and the reverse causality argument is unlikely to bias our baseline estimations. In Table E1 we perform a pre-trend test showing the absence of correlation between firms' export growth and the change in RTAs' depth, supporting qualitatively the absence of a reverse causality issue. Nevertheless, in section 5 we propose an Instrumental Variable (IV) approach aimed to reduce further any residual endogeneity concern. This could come, for instance, from large exporters lobbying for deep trade agreements. A plausible exogeneity test à la Conley et al. (2012), discussed in section 5.1, supports the robustness of our 2SLS results to deviations from the perfect validity of the exclusion restriction hypotheses and reinforces the causal interpretation of our results. The parallel trend assumption in the pre-treatment period is tested in section 6 where we adopt an event-study approach to test the effect of deep RTAs on firms' export.

The main empirical challenge here is the high-collinearity between the mere presence of an active RTA (abstracting from its content, RTA_{ijt}) and the depth of the agreement (the count of provisions therein, DTA_{ijt}). This problem is exacerbated when country-pair fixed effects (θij) are included in the estimation and absorb any cross country-pairs variability in the presence of RTAs and depth. For this reason, in our baseline estimations we do not control for the mere presence of an RTA; this is *de facto* subsumed by the DTA_{ijt} variable (i.e. DTA_{ijt} is larger than zero only when there is a RTA in force). However, since the main and ever-present objective of any RTA is to reduce bilateral tariffs, we control for the presence of a RTA by including in all the estimations the weighted average applied tariffs faced by a given firm f, into a given destination j across exported products, $ln\,(1+\tau_{fijt})$.[20] Hence, the trade effect of RTAs' depth (coefficient

[19] As a benchmark, in Appendix Table B1 we show OLS estimations.

[20] We use the share of exports of firm f on a given product p in the initial year $t0$ as a weight ω_{fipt_0}, to calculate the average applied tariffs faced by each firm at destination across exported products:

$$\tau_{fijt} = \sum_p \omega_{fipt_0} \times \tau_{ijpt} \tag{2}$$

β_1 in eq. 1) must be interpreted as on top of the standard market access effect captured by the applied tariff variable.

Heteroegeneous effect of DTAs In order to test the heterogeneous effect of the RTAs' depth on firms with different characteristics, we extend eq. (1) by interacting the DTA_{ijt} variable with a firm characteristic indicator $I(k_f > \bar{k})$ as follows:

$$\begin{aligned} X_{fijt} &= exp\Big[\theta_{ft} + \theta_{jt} + \theta_{ij} + \beta_1 DTA_{ijt} + \beta_2 ln\,(1+\tau_{fijt}) \\ &+ \beta_3\left(DTA_{ijt} \times I(k_f > \bar{k})\right)\Big] \times \varepsilon_{fijt}. \end{aligned} \tag{3}$$

The indicator $I(k_f > \bar{k})$ is equal to one if a given firm's characteristics k_f is above a threshold $\bar{k}$. Two firm-specific characteristics are used to define the indicator $I(k_f > \bar{k})$. First, as a proxy for the firm's size, we use the total exports of the firm in the initial year t_0 (across destinations and products),[21] and define a dummy variable equal to one if the size of the firm is above a threshold $\bar{k}$. Percentiles 75^{th} and 90^{th} of the initial firms' exports distribution are used as threshold $\bar{k}$. We use the total exports of the firm in the initial year t_0 (i.e. the first year in which the firm is observed in the data) to avoid the endogenous change in firm's export in the post-RTA period. Second, we capture the GVC status of the firm by a dummy equal to one if the exporting firm is also an importer, indicating that the firm is likely to use imported inputs in production for exports. Finally, we refine the GVC nature of the firm by using a dummy equal to one if the firm exports and imports to/from the same country j – *GVC bilateral*. This last firm characteristic is meant to capture the importance of deep RTAs for firms in developing countries having *bilateral* (import-export) relations with destination j.

Sector specific regression. Finally, to better understand whether certain sectors are more sensitive to RTAs' depth, and to the specific sub-groups of provisions, we resort to the sector dimension (HS 2-digit) and estimate the following firm-destination-sector-year level equation:

$$X_{fijht} = exp\left[\theta_{fht} + \theta_{jht} + \theta_{ij} + \sum_s \beta_s\,(DTA_{ijt} \times D_s) + \beta_2 ln\,(1+\tau_{fijt})\right] \times \varepsilon_{fijht} \tag{4}$$

where X_{fijht} now represents total exports of firm f in HS 2-digit product h to destination j at year t. θ_{fht}, θ_{jht} and θ_{ij} are firm-HS2-year, destination-HS2-year and country-pair fixed effects.

where $\omega_{fipt_0} = exports_{fipt_0}/\sum_p exports_{fipt_0}$, and τ_{ijpt} is the applied tariff faced by firms in country i when exporting product p in market j at time t.

[21] High-productive firms export more. So the total exports of the firm (across all products and destinations) is a plausible proxy for its productivity. See Fontagné et al. (2015).

D_s are dummy variables for each section s of the HS classification (HS 1-digit). Hence β_s indicates the effect of RTAs' depth on firms' exports in each HS section s.[22] This approach has two advantages. First, it allows to uncover the possible heterogeneous effect of DTAs and their content across different sectors. Second, we can control for any product-specific shock at origin θ_{fht} (where the firm is located), i.e. product-specific multilateral resistance term, distribution of productivity in a given country-product; and for any product-destination specific shock θ_{jht}, i.e. the strength of competition and demand shocks in a given destination country-product. Results on sector-specific regressions are reported in Section 7.

4 Results

This section discusses our main results on the effect of RTAs' depth. We start by showing the results obtained by estimating our baseline specification, equation (1), on the full sample of firms in developing countries facing changes in their RTAs' depth - section 4.1. In the same section, we propose a robustness check using the weighted count of provisions as an alternative measure of RTA depth. In section 4.2, we estimate equation (3) and show the heterogeneous effect of RTAs' depth on firms with different characteristics. Finally, in section 4.3, we disentangle the effect of deep RTAs into extensive *versus* intensive margins of exports.

4.1 Baseline results

Table 3 shows our baseline results. The depth of RTAs has a positive and significant effect on the export performance of firms in developing countries no matter the type of provisions considered to approximate the depth of the RTA (all, legally enforceable, WTO+, WTO-X, core, NTM-related or policy). In particular, one additional legally enforceable provision in the agreement boosts the exports of firms by 0.3%. This means that by moving from a *shallow* agreement (here defined as an RTA including only legally enforceable provisions related to tariff cuts in manufacturing and agriculture sectors) to a *deep* RTA containing legally enforceable provisions at the 75th percentile of the distribution of the RTAs depth implies a 3.6% increase in firm exports.[23] This effect adds on top of the the standard tariff-cut effect of RTAs on trade here captured by applied tariffs τ_{fijt}. This must be considered in reconciling the results of the present paper (focusing on the *number* of provisions included in RTAs) with previous papers approximating the overall effect of RTAs (i.e. tariff cuts and other provisions) by a simple dummy variable.

[22] We use the 15 headings of the HS classification to approximate the macro-sector of the firm s. Given the very large set of fixed effects, and the limited RAM capacity of our server, we use OLS to estimate equation 4.
[23] RTAs at the 75th percentile of legally enforceable provisions contain 14 provisions.

If we consider the count of WTO+ provisions, moving from shallow to deep RTAs implies a 3.6% increase of firm exports (i.e. one additional legally enforceable WTO+ provision boosts the exports of firms by 0.4%).[24] The effect *per-provision* is larger for legally enforceable WTO-X provisions: one additional WTO-X provision boosts the export of firms by 0.7% (moving from shallow to deep RTAs in WTO-X provision implies a 2.1% increase in firm exports).[25] The trade effect is even larger for NTM-related and Policy oriented provisions: moving from *shallow* RTAs that do not contain NTM-provisions to *deep* RTAs containing TBT and SPS provision implies a 4.8% increase in firm exports. The effect is 8.7% for provisions on services, investment and competition (i.e. policy-oriented provisions as defined in Dhingra et al. 2018).

As expected, the applied tariffs at destination have a negative and significant effect on the exports of firms. The point estimates on applied tariffs are lower than commonly obtained in the previous literature. This is due to the aggregation bias. Indeed, we aggregate the exports and the tariffs of firms across products, and the consequent aggregation bias produces tariff elasticity that are smaller in magnitude (Redding and Weinstein, 2019).[26]

Using the simple count of provisions to approximate the RTAs' depth implicitly gives the same importance to any type of provision. One may want to assign relatively higher value of depth to those RTAs including rare provisions. So, as a robustness check, in Table 4 we show results by using weighted count to approximate the depth of RTAs (the weight is equal to one minus each provision's frequency in the matrix of the RTAs mapped by the World Bank Deep Trade Agreements database). The resulting index weights relatively more rare provisions in RTAs. Results, reported in Table 4, support the robustness of our baseline results. Interestingly, the estimations coefficients in Table 4 point to a stronger impact of RTAs' depth on exports when approximated by a weighted sum. This suggests that the inclusion of rare provisions in RTAs is a good signal of the extent of trade cost reductions associated with deep RTAs between member countries.[27]

In Appendix Table B4 we replace origin-destination fixed effects with more demanding firm-destination fixed effects controlling for any firm-destination specific factor affecting the export

[24]WTO+ provisions contain standard tariff cut provisions on agriculture and manufacturing sectors, so we consider RTAs with two WTO+ provisions as shallow. The 75^{th} percentile in the count of WTO+ provisions is equal to 11.

[25]WTO-X provisions do not contain standard tariff cut provision, so we consider RTAs with zero WTO-X provision as shallow. The 75^{th} percentile in the count of WTO-X provision is equal to 3.

[26]The negative and significant coefficients on applied tariffs reassure us of the accuracy of our estimations. This accuracy is also supported by point estimates on standard gravity controls reported in appendix Tables B2 and B3. In these appendix tables, we remove country-pair fixed effects and include standard gravity model controls (such as distance, colony, common language and border) to have a benchmark with previous literature on gravity controls' coefficients (Head and Mayer, 2014). These variables have the expected sign and magnitude in line with previous studies.

[27]We do not use weighted sum indices for NTM and Policy oriented provisions as the weighted sum of only two/three provisions is meaningless.

Table 3: The trade effect of RTAs depth. Baseline specification. PPML estimator.

	Exp	Exp	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DTA_{ijt}	0.003^{***}						
	(0.001)						
DTA_{ijt} leg.		0.003^{***}					
		(0.001)					
DTA_{ijt} WTO+			0.004^{**}				
			(0.002)				
DTA_{ijt} WTO-X				0.007^{**}			
				(0.003)			
DTA_{ijt} Core					0.004^{***}		
					(0.001)		
DTA_{ijt} NTM						0.024^{**}	
						(0.010)	
DTA_{ijt} Ser-Inv-Com							0.029^{***}
							(0.011)
$Ln(1+\tau_{ijt})$	-0.578^{***}	-0.578^{***}	-0.578^{***}	-0.579^{***}	-0.578^{***}	-0.578^{***}	-0.578^{***}
	(0.073)	(0.073)	(0.073)	(0.073)	(0.073)	(0.073)	(0.073)
Firm-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Origin-Destination FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,534,360	1,534,360	1,534,360	1,534,360	1,534,360	1,534,360	1,534,360

Notes: PPML estimates. Standard errors in parenthesis are clustered at firm-destination and origin-destination-year level. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

Table 4: The trade effect of RTAs depth. Robustness check using the weighted sum of provisions as a proxy for the RTA depth.

	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)
Weigh. DTA_{ijt}	0.007^{***}				
	(0.001)				
Weigh. DTA_{ijt} leg.		0.005^{**}			
		(0.002)			
Weigh. DTA_{ijt} WTO+			0.008^{*}		
			(0.005)		
Weigh. DTA_{ijt} WTO-X				0.008^{**}	
				(0.003)	
Weigh. DTA_{ijt} core					0.007^{**}
					(0.003)
$Ln(1+\tau_{ijt})$	-0.578^{***}	-0.579^{***}	-0.579^{***}	-0.579^{***}	-0.578^{***}
	(0.073)	(0.073)	(0.073)	(0.073)	(0.073)
Firm-Year FE	Yes	Yes	Yes	Yes	Yes
Dest.-Year FE	Yes	Yes	Yes	Yes	Yes
Orig.-Dest. FE	Yes	Yes	Yes	Yes	Yes
Observations	1,534,360	1,534,360	1,534,360	1,534,360	1,534,360

Notes: PPML estimates. Standard errors in parenthesis are clustered at firm-destination and origin-destination-year level. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

of firm f in destination d. Results are robust in both sign and magnitude. Finally, we estimate eq. (1) on the sub-sample of country-pairs that change RTA status during the period 2000-2020. Results, reported in Appendix Table B5 support the robustness of our results. The non-linear effect of DTAs is shown in Table B6 by including the square term of DTAs measures in eq. (1).

The baseline and the robustness check regressions discussed so far suggest that the depth of RTAs has a statistically significant positive effect on the exports of firms in developing countries. However, such a positive effect may hide strong heterogeneity across firms of different types, a dynamic effect after the change in depth of an RTA, and a different effect depending on the sector. We dig more into these heterogeneous effects of deep RTAs in the next sections.

4.2 Firm heterogeneity

The reduction in fixed and variable export costs and the reduction in uncertainty associated with the depth of RTAs increase market access at destination and make it easier for firms of a given origin to export to that destination. In a world with homogeneous firms, this simply implies larger export values for firms at destination. But firms are heterogeneous (Bernard et al. 2007) and the expected trade effect of RTAs' depth is not as simple. In a standard model of trade with heterogeneous firms à la Melitz (2003), where all firms face the same demand elasticity and have the same mark-up, improved market access at destination makes large and more productive firms export relatively more than less productive firms. A second way firms differ is through their participation in global value chains. In particular, firms that participate in global value chains (GVC firms) can be more affected by deep RTAs either because they tend to be high-productivity firms (Goldberg et al., 2010; Feng et al., 2016) and/or because they import and export from the RTA's partner country, so that the reduction of trade costs and uncertainty associated to a deep RTA affects them both on the import and export side – see Limão (2016).

This section explores the heterogeneous effects of RTAs' depth across different types of firms based on size and GVC participation. Namely, we interact the DTA_{ijt} variable with the four firm-specific indicators $I(k_f > \bar{k})$ discussed in section 3. Results reported in Table 5 show an interesting and robust pattern: large and GVC firms benefit from deep RTAs.[28] For small firms, the depth of RTAs has a negative effect on exports and market shares – see columns (1) and (2) of Table 5.[29] These results suggest that deep trade agreements imply the re-allocation

[28] The same conclusion holds by using a non-parametric binned model where the DTA variable is interacted by three firm size bins based on whether the total exports of the firm (across destination and years) are below the 25^{th} percentile (small firms), above the 75^{th} percentile (big firms), or in between (medium size firms). See appendix table C3.

[29] Remember that given the structure of fixed effects included in eq. 3 our results can be interpreted in terms of both firm's export or market share.

of destination-specific sales and market shares from small to big exporters. Specifically, for large high-productivity firms, one additional legally enforceable provision implies a 0.4% - 0.6% increase in their exports (see columns 1 - 2). This means that, moving from *shallow* to *deep* RTAs implies a 4.8% - 7.2% increase in exports. For GVC firms exporting/importing to/from the same country j, moving from *shallow* to *deep* RTAs implies a 8.4% increase in exports (see columns 4). Deep RTAs reduce variable and fixed trade costs between members, and more productive firms are more likely to take advantage of this policy change. In addition to this size/productivity channel, GVC firms benefit as they see a reduction in trade costs on imports of intermediate products used in production for exports. The heterogeneous effect of deep RTAs is confirmed also by: (i) using the weighted sum of provision as a proxy for the depth of RTAs (see Table C1), and (ii) excluding highly-concentrated sectors (see Table C2).[30]

Table 5: The heterogeneous trade effect of RTAs' depth by firm characteristics.

	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)
DTA_{ijt} leg.	-0.004* (0.002)	-0.005*** (0.002)	-0.000 (0.001)	0.003** (0.001)
DTA_{ijt} leg. × ($k_f > 75^{th}$ t0)	0.008*** (0.002)			
DTA_{ijt} leg. × ($k_f > 90^{th}$ t0)		0.011*** (0.002)		
DTA_{ijt} leg. × GVC			0.007*** (0.002)	
DTA_{ijt} leg. × GVC bil.				0.004** (0.002)
$\text{Ln}(1+\tau_{ijt})$	-0.577*** (0.073)	-0.573*** (0.073)	-0.581*** (0.073)	-0.571*** (0.073)
Firm-Year FE	Yes	Yes	Yes	Yes
Dest.-Year FE	Yes	Yes	Yes	Yes
Orig.-Dest. FE	Yes	Yes	Yes	Yes
Observations	1,534,360	1,534,360	1,534,360	1,534,360

Notes: PPML estimates. Standard errors in parenthesis are clustered at firm-destination and origin-destination-year level. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

The negative effect on small firms is in line with the competition effect from larger foreign market access highlighted by trade models with variable mark-ups and heterogeneous elasticity of demand (Atkeson and Burstein, 2008; Mrázová and Neary, 2017; Mrázová and Neary, 2020; Crowley et al., 2024). This class of models predicts a stronger within-country competition effect from larger foreign market access (deep RTAs), that may even hurt small and less-productive

[30] We first assign each firm to a given (main) sector (i.e. the HS 1-digit sector with the largest export sales). Then, we calculate the HHI index of firms' export shares for each sector and country of origin. In Table C2 we exclude sector-country combinations having HHI above 90^{th} percentile of the distribution.

firms, and imply an increase in the concentration of foreign market shares among large and high-productivity firms. Namely, in a very general framework à la Mrázová and Neary (2020), the signature of a *deep* RTA increases both the size of the market (i.e. size effect) and the degree of competition among exporters (i.e. competition effect). In this context, the net effect of the signature of a deep RTA on firm's foreign sales depends on how the elasticity of demand of a specific firm compares with average elasticity of demand of other firms. For large (small) firms, having lower (higher) demand elasticity than the average, the signature of a DTA is expected to increase (decrease) foreign sales – see appendix section D for a detailed discussion of this theoretical interpretation. So, deep RTAs may increase the export sales of firms with certain characteristics (large, more productive and GVC firms) at the expense of small and less productive firms.

Moreover, if the signature of a deep RTA also increases the number of firms in origin country i exporting to j, small exporters may experience a reduction in their market share at destination (see Crowley et al. 2024), and eventually be induced to exit from the export market due to enhanced competition. Such a competition effect of deeper RTAs is shown in Table 6 where we estimate the effect of RTAs' depth on: (i) the number of firms in i that keep exporting at destination j at time t (i.e. surviving exporters), and (ii) the average export *per* firm. The depth of RTAs reduces the number of firms surviving in the export market (columns 1-2) but increase the average export (value) *per* firm (columns 3-4).[31] This confirms the existence of a competition effect of deep RTAs in developing countries, and points to welfare effects associated with adjustment costs and improved productivity deriving from the re-allocation of resources from less- to more-productive firms in each exporting country.

4.3 The extensive and intensive margin channels

The export sales effect of deep RTAs may come from the intensive and/or extensive margin of trade. Indeed, the reduction in trade costs associated with deep trade agreements may lead incumbent firms to boost their exports in destinations with deep RTAs (i.e. *intensive* margin), or may allow new firms to start exporting in such destination (i.e. *extensive* margin). In order to disentangle the overall export sales effect into the extensive and intensive margin channel, we calculate the change in *aggregate* country-pair specific exports from: (i) incumbent firms – i.e. firms always exporting to a given destination (intensive margin); and (ii) entry-exit firms

[31] A more compelling way of testing the competition effect of DTAs would be using the participation margin of firm level exports (i.e. dummy equal to one if the firm exports into a given destination-year). However, the EDD is a customs based dataset and does not include real "zeros". Any artificial inclusion of zeros (i.e. squaring the dataset with zeros) would be arbitrary.

Table 6: The competition effect of RTAs' depth.

	# Survivors	# Survivors	Avg Exp	Avg Exp
	(1)	(2)	(3)	(4)
DTA_{ijt} leg.	-0.006***	-0.005***	0.004**	0.006***
	(0.001)	(0.001)	(0.002)	(0.002)
$Ln(1+\tau_{ijt})$		0.519***		-0.153
		(0.092)		(0.097)
Origin-Year FE	Yes	Yes	Yes	Yes
Destination-Year FE	Yes	Yes	Yes	Yes
Origin-Destination FE	Yes	Yes	Yes	Yes
Observations	37,962	25,251	49,363	30,068

Notes: PPML estimations. The dependent variable in columns (1)-(2) is the number of exporters that survive at destination. The dependent variable in columns (3)-(4) is the average export sales per exporting firm (i.e. total export sales over total number of exporters). PPML estimates in columns (1)-(4). Origin-destination-year cluster standard errors in parentheses. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

to/from a specific market (extensive margin). Then, we adopt the decomposition approach as in Berman et al. (2012). Namely, we regress the aggregate exports by origin-destination-year for respectively incumbent and entry-exit firms on RTAs' depth, and country-year fixed effects. The coefficients on RTA depth are respectively $\hat{\beta}_{incumbent}$ and $\hat{\beta}_{entry}$. We then calculate the share of total exports from respectively incumbent and entry-exit firms ($V_{incumbent}/V_{tot}$ and V_{entry}/V_{tot}), and multiply such shares for the respective elasticity $\hat{\beta}_{incumbent}$ and $\hat{\beta}_{entry}$.

The results of this calculation are reported in Table 7. As expected, the elasticity to deep RTAs for both incumbent and entry firms are positive and significant, with a larger effect for incumbent firms. But since the share of total exports is larger for entry-exit firms than for incumbent firms, we obtain a slightly lower contribution from the intensive margin channel to the aggregate effect. Namely, the intensive (extensive) channel accounts for 46% (54%) of the total impact of RTAs' depth. The fact that RTAs' depth has a non-negligible effect on both the intensive and extensive margin of trade, suggests that both the fixed and the variable export cost are affected by deep RTAs. This is consistent with the view that several deep provisions in RTAs, such as the ones on technical standards, sanitary measures or services, allow to reduce fixed entry costs associated to divergent regulations and policy uncertainty.

5 Endogeneity

The inclusion of the sets of fixed effects discussed above strongly reduces any omitted variable concern. Namely, any firm specific productivity shock, as well as any import demand shock and country-pair-specific transaction cost are captured by fixed effect. In this setting, the only endo-

Table 7: Intensive *vs* extensive margin contribution to export response to RTAs depth.

	$\hat{\beta}$	V_i/V	Aggregate Response	Aggregate Response (% of total)
Intensive	0.044***	0.393	0.0173	46
Extensive	0.033***	0.607	0.0200	53
Total			0.0373	

Notes: $\hat{\beta}$ is the estimated coefficient for RTA depth on a gravity type regression (PPML) having the total country-pair-year specific exports for incumbent and entry-exit exporters. V_i/V is the share of total aggregate exports by respectively incumbent and entry-exit exporters. The aggregate response is calculated as $\hat{\beta} \times V_i/V$.

geneity concern may come from the presence of unobserved factors affecting contemporaneously a *change* in the RTA depth and a *change* in the exports of firm f in destination j. To further reduce any endogeneity concern, we propose an Instrumental Variable (IV) approach based on the idea that each country (i or j) sets the depth of *new* RTAs based on the depth of existing ones (in the vein of the domino effect of RTA formation by Baldwin and Jaimovich 2012). Our instrumental variable is therefore the following:

$$IV_{ijt} = \left[\frac{1}{K-1} \sum_{k \neq j} DTA_{ikt} \right] \times \left[\frac{1}{Z-1} \sum_{z \neq i} DTA_{zjt} \right] \tag{5}$$

where the two terms in brackets represent: (i) the average depth of RTAs signed by country i with trade partners $k \neq j$ within the j's macro-region, and (ii) the average depth of RTAs signed by country j with trade partners $z \neq i$ within the i's macro region.[32] We use the leave-one-out means to construct the average depth in brackets to address the finite sample bias coming from using own-observation information. The exclusion restriction is based on: (i) the absence of a direct effect of firms' exports toward destination j on the average depth of RTAs signed by country i and j with third countries, and (ii) on the absence of a direct effect of average RTA depth signed with third countries on firms' exports toward j. While condition (i) is likely to hold, condition (ii) deserves careful discussion. Indeed, the presence of a trade diversion effect of deep RTAs threatens the validity of our IV: if country i signs a deep RTA with third country k ($k \neq j$) and this diverts firm's exports from j to k, the validity of the IV is challenged. Notice, however, that the average reduction in trade costs of country i and country j with third countries is captured here by respectively firm-year and destination-year fixed effects. We are therefore confident of the exclusion restriction validity in our empirical setting. Still, in the next section, we present a plausible exogeneity test aimed at supporting the validity of our IV results even in the presence of weak deviation from the perfect validity of the exclusion restriction assumption.

[32] Macro-regions are: East Asia and Pacific, Europe and Central Asia, Latin America and Caribbean, Middle East and North Africa, North America, South Asia, Sub-Saharan Africa.

Our baseline results are confirmed by the 2SLS estimations addressing any residual endogeneity concerns - see Table 8 column 1. The (instrumented) measure of RTA depth has a positive and significant effect on the exports of the average firm. Also, our results on the heterogeneous impact of DTAs on firms with different characteristics are confirmed by 2SLS estimations - see Table 8 columns 2-5. The interaction terms are instrumented by simply interacting the firm indicators $I(k_f > \bar{k})$ with the IV discussed above. The bottom part of Table 8 shows the first stage results of the 2SLS approach. Our IVs (one for the RTA depth and the other for its interaction with the firm-type dummy) are good predictors of the endogenous variables (i.e. significant first-stage coefficients). Also, the joint F-stat statistics well above 10 support the absence of a weak instrument problem.[33]

Table 8: The trade effect of RTAs depth. 2SLS approach.

	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)
DTA$_{ijt}$ leg.	0.009* (0.005)	-0.017*** (0.006)	-0.006 (0.006)	0.006 (0.006)	0.004 (0.005)
DTA$_{ijt}$ leg. × (k$_f$> 75th t0)		0.033*** (0.006)			
DTA$_{ijt}$ leg. × (k$_f$> 90th t0)			0.025*** (0.005)		
DTA$_{ijt}$ leg. × GVC				0.006 (0.006)	
DTA$_{ijt}$ leg. × GVC bil.					0.023*** (0.005)
Ln(1+τ_{ijt})	-0.509*** (0.060)	-0.508*** (0.060)	-0.503*** (0.061)	-0.511*** (0.059)	-0.520*** (0.060)
Firm-Year FE	Yes	Yes	Yes	Yes	Yes
Dest.-Year FE	Yes	Yes	Yes	Yes	Yes
Orig.-Dest. FE	Yes	Yes	Yes	Yes	Yes
Observations	1,470,487	1,470,487	1,470,487	1,470,487	1,470,487
IV: DTA$_{ijt}$	0.058***	0.059***	0.059***	0.059***	0.058***
IV: DTA$_{ijt}$ × $I(k_f > \hat{k})$		0.098***	0.099***	0.102***	0.072***
Joint F-stat	65	32	32	32	32

Notes: 2SLS estimates. Standard errors in parenthesis are clustered at origin-year and destination-year level (i.e. the two actual sources of variability of our IV). ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

[33] A limitation of *linear* 2SLS estimates in Table 8 is that they cannot be directly compared with our baseline *non-linear* PPML estimations. To address this issue, we adopt a two-stage OLS/PPML control function approach (Lin and Wooldridge, 2019). In the first stage, we use simple OLS to estimate the effect of our IV on RTA depth (i.e. standard first stage in 2SLS). Then, in the second stage, we use PPML estimator, keep our (endogenous) measure of RTA depth as main explanatory variable, and include the estimated residual of the first stage as a control – i.e. control function approach. Results available upon request support the robustness of our IV strategy.

5.1 IV validity

As discussed above, the validity of our instrumental variable is based on the absence of a direct effect of the depth of RTAs signed with third-country on fij-specific exports. In this section, we test the robustness of our baseline results to a deviation from the perfect validity of the exclusion restriction (Conley et al., 2012). A degree of deviation from the exclusion restriction can be obtained by regressing the exports of firms X_{fijt} on our main DTA_{ijt} and IV variable. The coefficient associated to the IV represents an approximation of the direct effect of the IV on the outcome variable (i.e. degree of deviation from the exclusion restriction) – van Kippersluis and Rietveld (2018). Reassuringly, we obtain a small and not statistically significant direct effect of IV on firm exports (i.e. not-significant direct effect of the IV on the outcome variable). See parameter ν in Table E2. Then, we abstract from the statistically insignificance of parameter ν, plug such a degree of deviation from exclusion restriction in the plausible exogeneity test à la Conley et al. (2012), and obtain lower- and upper-bound coefficients that do not cross the zero – see Table E2. Thus, we can safely argue that the depth of RTAs has an unambiguous positive causal effect on the export of firms in developing countries even in presence of small deviations from the perfect exclusion restriction assumption of the IV.[34]

6 The dynamic effect of deep RTAs: An event study approach

We relied so far on a fixed-effects approach delivering the average effect of deep RTAs on the exports of firms (i.e. pre- *versus* post-change in RTA depth). However, deep RTAs may stimulate firms' exports dynamically. In this section, we adopt an event-study approach to visualize the dynamic effect of deep RTAs. Namely, we follow Fajgelbaum et al. (2020), and compare the treated varieties (i.e. firm-destination combinations that face a change in RTA depth) to *never* treated varieties, using the following specification:[35]

[34]For computational reasons (i.e. maximum number of covariates allowed by `plausexog` STATA command) we had to reduce the number dummies (fixed effects) included in the estimations. First, we replaced destination-year fixed effects by destination-period fixed effects, with periods containing 6-year each. Second, we restricted the number of destinations by: (i) using top-50 destinations for all the 31 exporting countries of our sample (panel a in Table E2), (ii) keep only destinations representing at least the 0.5% of total exports of all the 31 exporting countries of our sample.

[35]In the vein of De Chaisemartin and d'Haultfoeuille (2020), to avoid the negative weight bias arising from comparing treated to not-yet-treated pairs (i.e. country pairs that become treated later in the time period), in estimating eq.(6) we compare treated origin-destination pairs with *never* treated pairs. As a further robustness check we strictly follow De Chaisemartin and d'Haultfoeuille (2020) and perform robust two-way fixed effects estimations of the trade effect of a change in the number of legally enforceable provisions in RTAs (dummy equal to one if positive change in the number of legally enforceable provisions in RTAs). Reassuringly, the two-way fixed effects estimations à la De Chaisemartin and d'Haultfoeuille (2020) produce qualitative identical results. See appendix section B for more details on the two-way fixed effects estimations.

$$X_{fijt} = exp\Bigg[\theta_f + \theta_{jt} + \theta_{ij} + \sum_{z=-2}^{3} \beta_{0z}\mathbf{I}\left(event_{ijt} = z\right) + \sum_{z=-2}^{3} \beta_{1z}\mathbf{I}\left(event_{ijt} = z\right) \times treated_{fj}\Bigg] \times \varepsilon_{fijt} \quad (6)$$

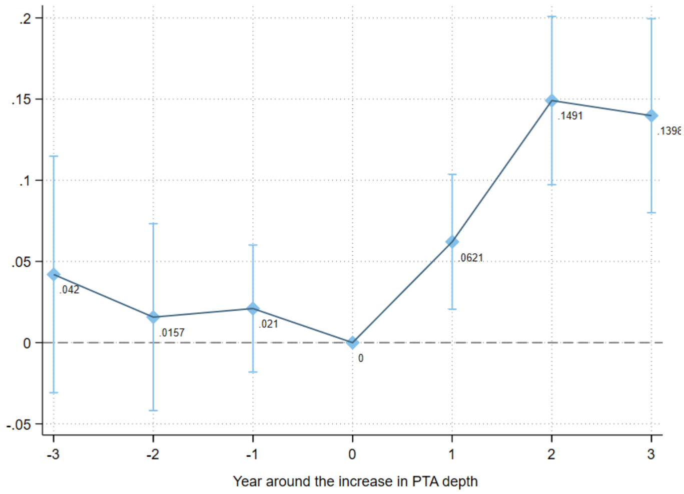
.

The event-study specification includes firm (θ_f), destination-year (θ_{jt}) and origin-destination (θ_{ij}) fixed effects. The indicator variable $\mathbf{I}\left(event_{ijt} = z\right)$ captures the event time coefficient before ($z = -3, -2, -1$) and after ($z = 1, 2, 3$) the change in the RTA depth ($z = 0$). The treatment variable, $treated_{fj}$, is in turn : (i) a dummy variable for varieties (i.e. firm-destination) that experience a change in the RTA depth during the period, and (ii) the extent of such a *change* in RTA's depth. The presence of firm fixed effects implies that β_{1z} coefficients are identified using the variation between treated and non-treated varieties at each point in time z. For treated varieties, the event date is the year of the first change in the RTA depth between country i and j. For non-treated varieties, we assign the event date to be the earliest year in which the destination market j experiences a change in RTA's depth with at least one of its trade partners. As in the baseline specification, we adopt a PPML estimator and cluster standard errors by firm-destination and origin-destination-year.

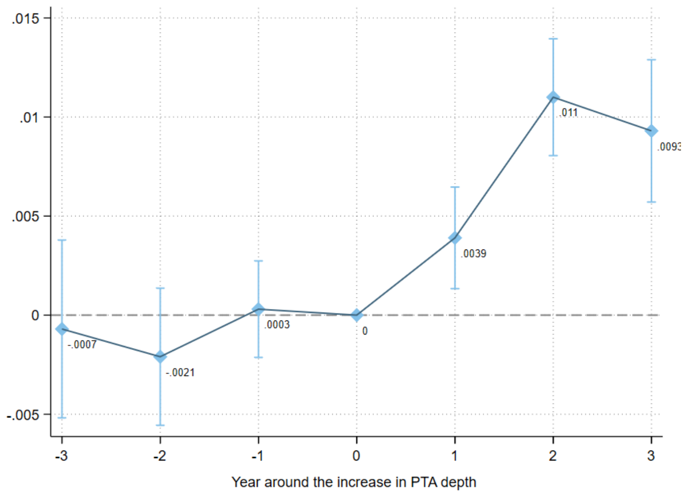
Figure 2 reports the impact of RTA depth on treated varieties for respectively all (Panel a and b) and large firms (Panel c). In Panel (a) the treatment variable is a dummy 0/1 indicating the change in the RTAs' depth. In Panel (b) and (c) the treatment variable is the extent of such a change in RTAs' depth. Panel (a) shows the positive effect of a *change* in the RTAs' depth on the exports of firms in developing countries. Reassuringly, the magnitude of such an effect is in line with Nagengast and Yotov (2024). Panel (b) and (c) show the positive and significant effect of the *extent* of the change in the number of provisions (i.e. RTA's depth) on firms' exports. In line with previous evidence suggesting that tariff elasticity is larger in the long-run (Boehm et al., 2023), Panel (a) and (b) show that the positive effect of RTAs depth increases over time from $t + 1$ to $t + 3$. Interestingly, and in line with our baseline results, the average dynamic effect of DTAs is driven by large firms – see Panel (c) of Figure 2.

Importantly, the event study approach also addresses concerns on the anticipation of changes in RTA depth (and/or RTA signature). During the pre-treatment period treated and non-treated varieties show a parallel trend. The absence of pre-trend reassures us about the causal interpretation of our baseline results.

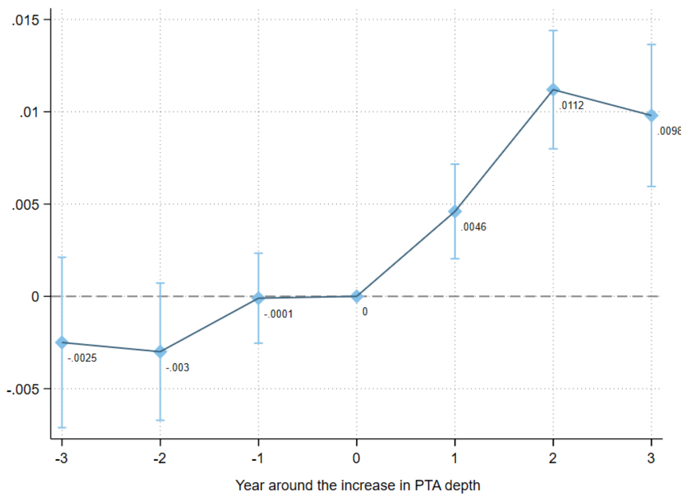
Figure 2: Deep Trade Agreements and firm-level exports. Event study approach.



(a) All firms. Treatment: event dummy.



(b) All firms. Treatment: num. provisions.



(c) Large firms. Treatment: num. provisions

Note: Figure plots event time dummies for treated firms relative to untreated firms. Regression includes firm, destination-year and origin-destination fixed effects. Standard errors are clustered by firm-destination and origin-destination-year. In Panel (a) the treatment variable is a dummy 0/1 indicating the change in the RTAs' depth. In Panel (b) and (c) the treatment variable is the extent of such a change in RTAs' depth. Big firms in Panel (c) are those with initial total export above the 75^{th} percentile. Error bars show 90% confidence intervals.

7 Sector heterogeneity

Deep RTAs, and the specific type of provisions included therein, may have different effects across sectors. When low-productivity firms represent a large fraction of firms in the country-sector (i.e. very skewed distribution of firms productivity), a change in the productivity threshold induced by higher market access (deep RTA) causes many small firms to enter the export market (large extensive margin effect). Conversely, when the distribution of firm productivity is more dispersed and large and high-productivity firms represent a larger fraction of firms in the country-sector, a change in the productivity threshold has a marginal effect on the extensive margin, but strong positive effect on the intensive margin of export (i.e. large firms exporting more). Hence, the aggregate effect of deep RTAs on a given *sector* depends on the productivity distribution of firms in the exporting country-sector.

In this section, we resort to the sector dimension of the EDD data, estimate equation (4), and show the effect of RTA depth (i.e. count of legally enforceable provisions) on firms' exports for each of the 14 sector-heading of the HS classification (i.e. HS 1-digit sections).[36] Results are shown in Figure 3.[37] Deep RTAs have positive and significant effect on almost all sectors, with strongest effect in "Transportation", "Footwear/Headgear", "Vegetable Products" and "Plastic/Rubber" sectors. We also uncover a strong negative effect in "Animal & Animal Products" and "Raw Hides, Skins & Leather" sectors. In light of the heterogeneous effect of DTAs shown in the previous sections, the average negative effect on "Animal & Animal Products" and "Raw Hides, Skins & Leather" sectors can be rationalised by the presence of many small firms suffering the increase competition from few (very) large firms, i.e. highly skewed distribution of firm productivity (size).

To further test this possible explanation for the negative effect of DTAs in some sectors, in Table 9 we interact the RTA depth variable with two proxies of the country-sector-specific Pareto distribution parameter (falling with an increase in the share of high-productive firms). The first proxy is obtained by following the QQ approach in Head et al. (2014),[38] the other by following Gabaix and Ibragimov (2011). We also approximate the skewness of country-sector firms' export sales by the coefficient of variation.[39] To facilitate the interpretation of the results,

[36] As discussed in section 2 we exclude Oil/Mining sectors.

[37] In the Online Appendix Figure F1 we show sector-specific results using the other proxies of RTAs' depth (i.e. respectively the count of WTO-plus, WTO-X, Core, NTM and Policy related provisions).

[38] As in Head et al. (2014), we retrieve the Pareto-shape parameter of firm size distribution by using the QQ estimator. We regress the empirical quantiles of the sorted log exports on the theoretical quantiles (i.e. $-ln(1-((k-0.3)/(n+0.4))$, where k is the firm's ascending order of exports and n the rank of the firm having the highest export value). The coefficient of such regression, $1/\tilde{\theta}$, gives us the inverse of the Pareto shape parameter, which we recover as $\theta = (\sigma - 1)\tilde{\theta}$. We use the elasticity of substitution $\sigma = 5$.

[39] To reduce any endogeneity concern, we use the beginning of the sample data to calculate Pareto shape parameters and the coefficient of variation.

we use a dummy variable for country-sector combinations having above-the-median Pareto shape parameter and below-the-median coefficient of variation, both indicating highly-skewed country-sectors distribution of sales. As in section 4.3, we calculate the extensive and intensive margin of exports as the total sales of respectively entry-exit and continuous exporting firms in a given origin-sector-destination cell.[40]

In line with the intuition, deep RTAs have a strong positive effect on the intensive margin of country-sector pairs with more disperse distribution of productivity, i.e. in country-sectors with large share of high-productivity firms. See Panel (a) of Table 9. Conversely, the effect of DTAs is negative in country-sector pairs having a very skewed productivity distribution (i.e. high Pareto shape parameter) – see columns (1)-(2), or low coefficient of variation – see column (3). In highly concentrated sectors, the many low-productivity firms suffer the increased competition from the few large firms; making the average aggregate effect of DTAs negative. This explains why in some sectors, such as "Animal & Animal Products" and "Raw Hides, Skins & Leather", deep trade agreements can have on average a negative effect on exports. As for the extensive margin of trade, in country-sector with a larger share of highly productive firms (i.e. Pareto shape parameter below the median) the extensive margin of exports is negatively affected. This is in line with the competition effect of deep RTAs discussed in section 4.2. When the exporting country-sector is populated by a large share of high-productivity firms (low Pareto shape parameter), the stronger within origin-sector competition induced by DTAs makes less productive firms exit the market (i.e. negative extensive margin effect).

Conclusion

This paper investigates the effects of Regional Trade Agreements' depth (i.e. number of provisions included therein) on firms' exports in developing countries. We show a moderate but statistically significant effect of RTA depth on the exports of the *average* firm. Namely, one additional legally enforceable provision boosts the export of firms by 0.3%. This implies that moving from shallow to deep RTAs leads to a 3.6% increase in firms' exports. This average effect is however strongly *heterogeneous* across firms with different characteristics. For large and for GVC firms, moving from shallow to deep RTAs implies a 6% and a 7.2% increase in exports, respectively. Conversely, small firms are negatively affected by deep RTAs as they suffer the higher degree of competition induced by deep RTAs. These results are robust to a number of

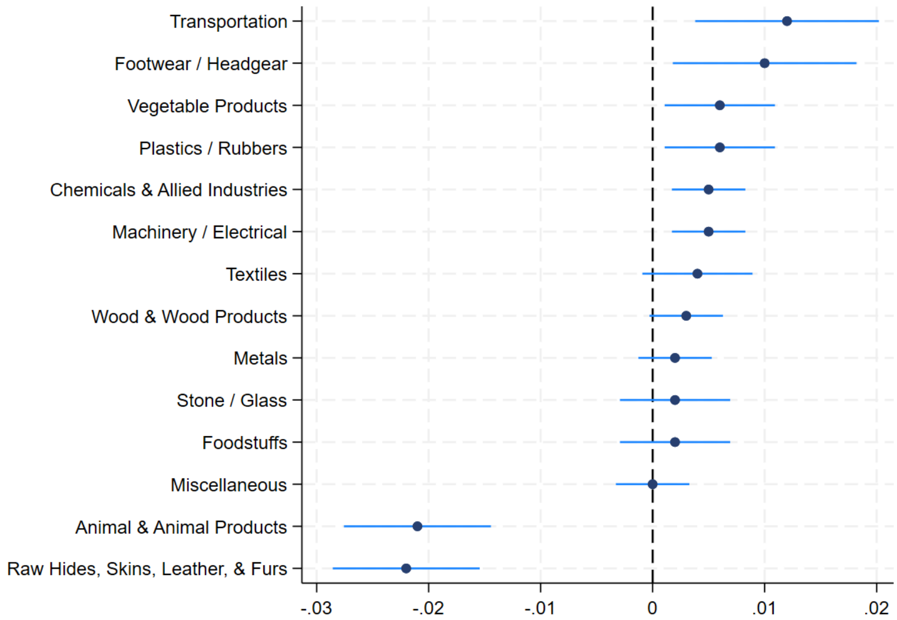
[40] In the regressions reported in Table 9 we include origin-sector-year, destination-sector-year and origin-destination fixed effects.

Table 9: The trade effect of RTAs depth. The role of firm size distribution.

	Panel a: Intensive margin		
	(1)	(2)	(3)
DTA_{ijt} leg.	0.010***	0.008***	0.004**
	(0.002)	(0.002)	(0.002)
DTA_{ijt} leg. × (Pareto shape^a > Mdn)	-0.019***		
	(0.003)		
DTA_{ijt} leg. × (Pareto shape^b > Mdn)		-0.014***	
		(0.003)	
DTA_{ijt} leg. × (Coef. Var. < Mdn)			-0.011***
			(0.002)
Observations	112,099	112,099	112,099
	Panel b: Extensive margin		
	(1)	(2)	(3)
DTA_{ijt} leg.	-0.004***	-0.004***	-0.003**
	(0.001)	(0.001)	(0.001)
DTA_{ijt} leg. × (Pareto shape^a > Mdn)	0.002		
	(0.002)		
DTA_{ijt} leg. × (Pareto shape^b > Mdn)		0.004***	
		(0.002)	
DTA_{ijt} leg. × (Coef. Var. < Mdn)			0.001
			(0.001)
Observations	158,397	158,397	158,397
Dest.- Sec.- Year FE	Yes	Yes	Yes
Orig. - Sec. - Year FE	Yes	Yes	Yes
Origin-Destination FE	Yes	Yes	Yes

Notes: PPML estimates, origin-destination-year cluster standard errors in parentheses. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1. (a) Pareto shape parameter estimated following Head et al. (2014). (b) Pareto share parameter estimated following Gabaix and Ibragimov (2011)

Figure 3: The impact of RTAs depth on firms' export by sector.



Source: Authors' calculations. *Notes*: This figure shows the effect of RTAs' depth (i.e. total count of legally enforceable provisions) of the export of firms. 90% confidence intervals are shown for our parameter of interest.

extensions and robustness checks that confirm the causal impact of deep RTAs on firms' exports.

The findings in this study have relevant welfare and policy implications for developing countries. While the competition effect, through the reallocation of resources toward more productive firms, is expected to improve welfare in countries joining deep trade agreements, the negative export performance of small firms signals that the adjustment costs in developing countries can be significant. Policy makers should therefore complement the implementation of deep trade agreements with the introduction of domestic policies that facilitate adjustment and support those that are hurt from the reallocation process.

Bibliography

Atkeson, A. and Burstein, A. (2008). Pricing-to-market, trade costs, and international relative prices. *American Economic Review,*, 98(5):1998–2031.

Baier, S. L. and Bergstrand, J. H. (2007). Do free trade agreements actually increase members' international trade? *Journal of international Economics*, 71(1):72–95.

Baier, S. L., Bergstrand, J. H., and Clance, M. W. (2018). Heterogeneous effects of economic integration agreements. *Journal of Development Economics*, 135:587–608.

Baier, S. L., Bergstrand, J. H., and Feng, M. (2014). Economic integration agreements and the margins of international trade. *Journal of International Economics*, 93(2):339–350.

Baier, S. L., Yotov, Y. V., and Zylkin, T. (2019). On the widely differing effects of free trade agreements: Lessons from twenty years of trade integration. *Journal of International Economics*, 116:206–226.

Baldwin, R. and Jaimovich, D. (2012). Are Free Trade Agreements contagious? *Journal of International Economics*, 88(1):1–16.

Berman, N., Martin, P., and Mayer, T. (2012). How do different exporters react to exchange rate changes? *Quarterly Journal of Economics*, 127(1):437–492.

Bernard, A. B., Jensen, J. B., Redding, S. J., and Schott, P. K. (2007). Firms in International Trade. *Journal of Economic Perspectives*, 21(3):105–130.

Bernard, A. B., Jensen, J. B., Redding, S. J., and Schott, P. K. (2018). Global firms. *Journal of Economic Literature*, 56(2):565–619.

Bernard, A. B., Jensen, J. B., and Schott, P. K. (2009). Importers, Exporters and Multinationals: A Portrait of Firms in the U.S. that Trade Goods. In *Producer Dynamics: New Evidence from Micro Data*, NBER Chapters, pages 513–552. National Bureau of Economic Research, Inc.

Blanga-Gubbay, M., Conconi, P., and Parenti, M. (2024). Lobbying for Globalization. *The Economic Journal*, forthcoming.

Blanga-Gubbay, M., Conconi, P., Song, K., and Parenti, M. (2021). Lobbying on deep trade agreements: How large firms buy favourable provisions. In *The Economics of Deep Trade Agreements*, pages 37–48. World Bank, Washington DC.

Boehm, C. E., Levchenko, A. A., and Pandalai-Nayar, N. (2023). The long and short (run) of trade elasticities. *American Economic Review*, 113(4):861–905.

Breinlich, H., Corradi, V., Rocha, N., Ruta, M., Zylkin, T., and Santos Silva, J. (2022). Machine

Learning in International Trade Research - Evaluating the Impact of Trade Agreements. CEPR Discussion Papers 17325, C.E.P.R. Discussion Papers.

Conley, T. G., Hansen, C. B., and Rossi, P. E. (2012). Plausibly Exogenous. *The Review of Economics and Statistics*, 94(1):260–272.

Crowley, M. A., Han, L., and Prayer, T. (2024). The pro-competitive effects of trade agreements. *Journal of International Economics*, 150:103936.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.

De Chaisemartin, C. and d'Haultfoeuille, X. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, pages 1–45.

Dhingra, S., Freeman, R., and Mavroeidi, E. (2018). Beyond Tariff Reductions: What Extra Boost From Trade Agreement Provisions? CEP Discussion Paper 1532, LSE.

Dhingra, S. and Morrow, J. (2019). Monopolistic competition and optimum product diversity under firm heterogeneity. *Journal of Political Economy*, 127(1):196–232.

Edmond, C., Midrigan, V., and Xu, D. Y. (2015). Competition, markups, and the gains from international trade. *American Economic Review*, 105(10):3183–3221.

Fajgelbaum, P. D., Goldberg, P. K., Kennedy, P. J., and Khandelwal, A. K. (2020). The return to protectionism. *The Quarterly Journal of Economics*, 135(1):1–55.

Feng, L., Li, Z., and Swenson, D. L. (2016). The connection between imported intermediate inputs and exports: Evidence from chinese firms. *Journal of International Economics*, 101:86–101.

Fernades, A., Lefebvre, K., and Rocha, N. (2021). Heterogeneous Impacts of SPS and TBT Regulations : Firm-Level Evidence from Deep Trade Agreements. Policy Research Working Paper Series 9700, The World Bank.

Fernandes, A., Rocha, N., and Ruta, M. (2021). *The Economics of Deep Trade Agreements*. CEPR Press.

Fernandes, A. M., Freund, C., and Pierola, M. D. (2016). Exporter behavior, country size and stage of development: Evidence from the exporter dynamics database. *Journal of Development Economics*, 119:121–137.

Fontagné, L., Orefice, G., Piermartini, R., and Rocha, N. (2015). Product standards and margins of trade: Firm-level evidence. *Journal of International Economics*, 97(1):29–44.

Freund, C. and Ornelas, E. (2010). Regional trade agreements. *Annual Review of Economics*, 2(1):139–166.

Freund, C. and Pierola, M. D. (2015). Export superstars. *The Review of Economics and Statistics*, 97(5):1023–1032.

Gabaix, X. and Ibragimov, R. (2011). Rank- 1/2: a simple way to improve the ols estimation of tail exponents. *Journal of Business & Economic Statistics*, 29(1):24–39.

Goldberg, P. K., Khandelwal, A. K., Pavcnik, N., and Topalova, P. (2010). Imported Intermediate Inputs and Domestic Product Growth: Evidence from India*. *The Quarterly Journal of Economics*, 125(4):1727–1767.

Head, K. and Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of International Economics*, volume 4, Handbook of International Economics, chapter 4. Gita Gopinath, Elhanan Helpman and Kenneth Rogoff editors.

Head, K., Mayer, T., and Thoenig, M. (2014). Welfare and trade without pareto. *American Economic Review*, 104(5):310–16.

Hofmann, C., Osnago, A., and Ruta, M. (2017). Horizontal depth: A new database on the content of preferential trade agreements. Policy Research Working Paper 7981, World Bank Group, Washington, DC.

Horn, H., Mavroidis, P. C., and Sapir, A. (2010). Beyond the WTO? An Anatomy of EU and US Preferential Trade Agreements. *The World Economy*, 33(11):1565–1588.

Limão, N. (2016). Preferential trade agreements. In *Handbook of commercial policy*, volume 1, pages 279–367. Elsevier.

Lin, W. and Wooldridge, J. M. (2019). Testing and correcting for endogeneity in nonlinear unobserved effects models. In *Panel data econometrics*, pages 21–43. Elsevier.

Mattoo, A., Mulabdic, A., and Ruta, M. (2022). Trade creation and trade diversion in deep agreements. *Canadian Journal of Economics*, 55(3):1598–1637.

Mattoo, A., Rocha, N., and Ruta, M. (2020). *Handbook of deep trade agreements*. World Bank Publications.

Mayer, T., Melitz, M. J., and Ottaviano, G. I. P. (2021). Product Mix and Firm Productivity Responses to Trade Competition. *The Review of Economics and Statistics*, 103(5):874–891.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

Melitz, M. J. and Ottaviano, G. I. P. (2008). Market size, trade, and productivity. *Review of Economic Studies*, 75(1):295–316.

Mrázová, M. and Neary, J. P. (2017). Not so demanding: Demand structure and firm behavior. *American Economic Review*, 107(12):3835–74.

Mrázová, M. and Neary, J. P. (2020). Io for exports(s). *International Journal of Industrial Organization*, 70:102561.

Nagengast, A. J. and Yotov, Y. (2024). Staggered difference-in-differences in gravity settings: Revisiting the effects of trade agreements. *American Economic Journal: Applied Economics*, forthcoming.

Orefice, G. and Rocha, N. (2014). Deep Integration and Production Networks: An Empirical Analysis. *The World Economy*, 37(1):106–136.

Parenti, M. and Gonzague, V. (2022). A simple theory of deep trade integration. CEPR Discussion Paper 17199, CEPR.

Redding, S. J. and Weinstein, D. E. (2019). Aggregation and the gravity equation. *AEA Papers and Proceedings*, 109:450–55.

Santos-Silva, J. M. C. and Tenreyro, S. (2006). The Log of Gravity. *The Review of Economics and Statistics*, 88(4):641–658.

van Kippersluis, H. and Rietveld, C. A. (2018). Beyond plausibly exogenous. *Econometrics Journal*, 21(3):316–331.

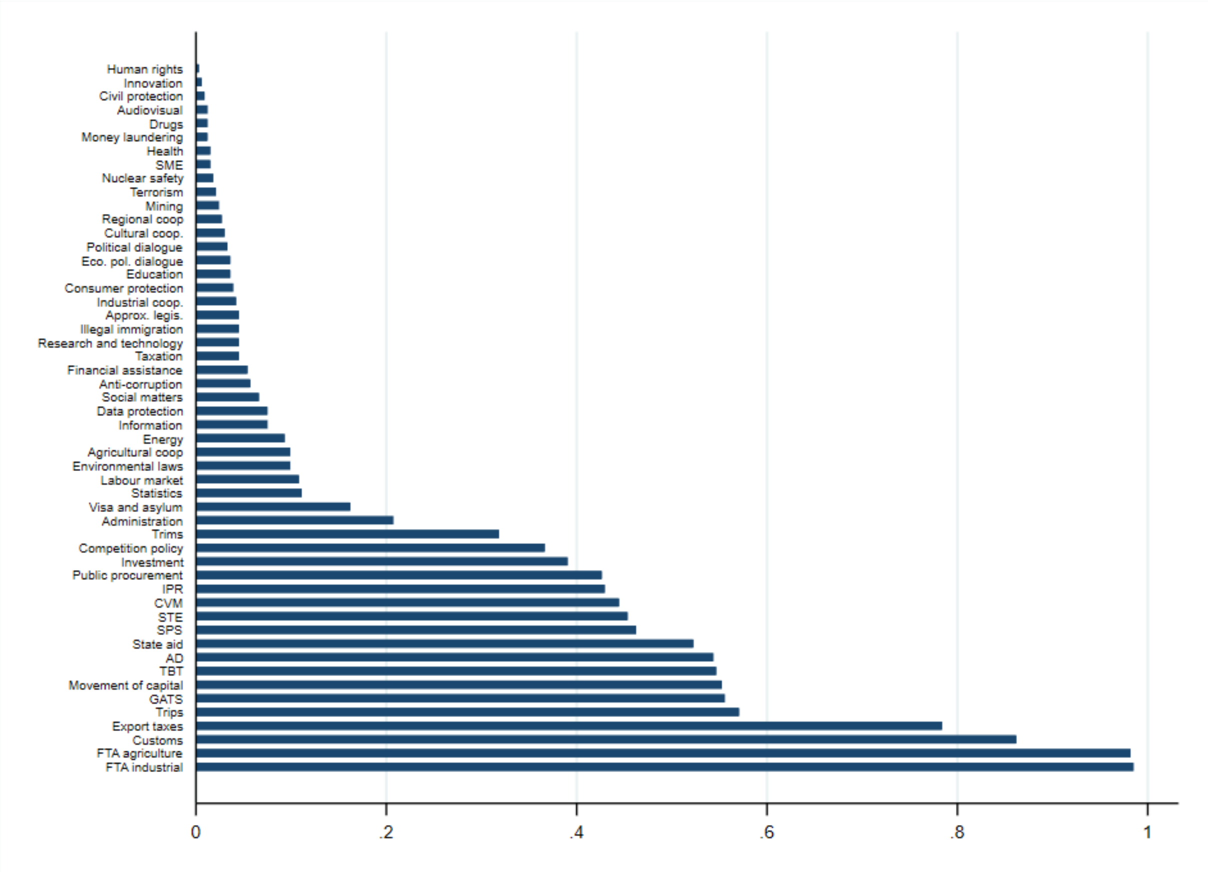
Online Appendix

Deep Trade Agreement and Heterogeneous Firms Exports

Matteo Neri--Lainé Gianluca Orefice Michele Ruta

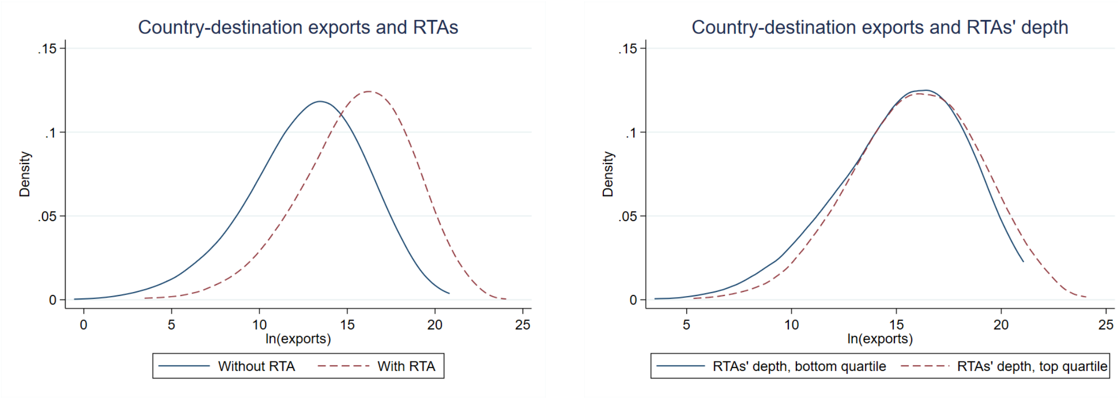
Appendix A. Additional descriptive statistics

Figure A1: Frequency of legally enforceable provisions in RTAs.



Note: Provisions are ranked from the less frequent to the most frequent.

Figure A2: Exports value and RTAs, all available countries.



(a) RTAs, country level

(b) Depth, country level

Note: K-density graphs are realized compiling country's exports for last year of available data.

Table A1: Time periods covered by country specific EDD data.

Country (iso3)	Period	Country (iso3)	Period
ALB	2007-2019	MEX	2011-2016
BWA	2003-2013	MKD	2008-2017
CHL	2000-2020	MUS	2010-2020
CIV	2009-2019	MWI	2005-2020
COL	2000-2020	NIC	2012-2014
DOM	2006-2020	PER	2000-2020
ECU	2002-2019	PRY	2012-2020
EGY	2005-2016	RWA	2005-2016
GEO	2000-2020	SLV	2006-2020
GHA	2010-2019	SRB	2006-2019
GTM	2005-2013	TZA	2003-2017
HRV	2007-2015	UGA	2000-2010
JOR	2003-2012	URY	2001-2020
KEN	2006-2020	YEM	2008-2012
LBN	2008-2012	ZAF	2009-2020
MDG	2007-2012		

Notes: World Bank Export Dynamics Database.

Table A2: In-sample descriptive statistics.

	Mean	Std Dev	Min	Max
Export	375671	1.3e+06	1	1.3e+08
RTA depth	15.0	13.4	0	48
RTA depth legally enf.	8.8	7.4	0	43
RTA depth WTO+	6.6	5.2	0	14
RTA depth WTOX	2.2	3.0	0	29
RTA depth core	7.9	6.3	0	18
RTA depth NTM	0.8	0.9	0	2
RTA depth Policy	0.9	0.9	0	3
$\ln(1+\tau)$	0.04	0.09	0	2.40

Notes: Authors' calculation on Export Dynamic Database, World Bank Content of Deep Trade Agreement data and MacMap (CEPII) dataset.

Appendix B. Robustness checks on baseline results

OLS estimations. In Table B1 we estimate equation (1) using linear OLS estimator. Hence we express our dependent variable in log (by relying on custom firm-level data, we do not have zeros trade flows). While we consider PPML our baseline estimator because of the heteroskedasticy problem in log-linear gravity estimation (Santos-Silva and Tenreyro, 2006), it is somehow reassuring that OLS estimations strongly support baseline results reported in Table 3.

Cross-section identification. In Tables B2 and B3 we estimate equation (1) by replacing country-pair fixed effects by standard gravity-type variables (i.e. distance, contiguity, common language, colony) respectively without and with tariff variable as a control. Two important results emerge from Tables B2 and B3. First, our coefficients of interest capturing the different aspects of Deep RTAs remain positive and highly significant. Second, coefficients on gravity-type variables have the expected sign and magnitude – see Head and Mayer (2014).

Firm-destination FE. In Table B4 we replace origin-destination with firm-destination fixed effects in equation (1). The idea is controlling for any firm-destination specific relationship affecting the pattern of trade of a given firm into a given destination. The source of identification is then within firm-destination over time. Results in Table B4 confirm our baseline estimations in both sign, magnitude and significance of coefficients.

Sub-sample of pairs changing RTA status. In table B5 we estimate equation (1) on the sub-set of country-pairs that change RTA status during the period 2000-2020. Results in Table B5 confirm our baseline estimations in both sign, magnitude and significance of coefficients.

Non-linear effect of DTAs. In table B6 we test the non-linear effect of RTAs' depth by including the square term of DTAs measures in eq. 1.

Two-way (robust) fixed effects estimations. As recently argued by De Chaisemartin and d'Haultfoeuille (2020), two-way fixed effects estimations with heterogeneous treatment across groups and over time may be biased by negative weights. Indeed, comparing the outcomes of *treated* country-pairs with those of *non-treated* pairs that may (or may not) be treated afterwards may cause negative weights in the difference-in-difference estimator. For this reason, in the event study estimations reported in Figure 2 we compare treated with *never* treated pairs. However, as a further check, we follow De Chaisemartin and d'Haultfoeuille (2020) and perform a robust two-way fixed effects estimation of the trade effect of a change in the number of legally enforceable provisions in RTAs (dummy equal to one in case of an increase in the number of legally enforceable provisions in RTAs). Since we are interested in the dynamic effect, we specifically adopt the estimator proposed by De Chaisemartin and d'Haultfoeuille (2024). The intuition behind this estimation is to avoid negative weights that may bias standard two-way fixed effects estimators. One should compare the outcome change from $t-1$ to $t+l$ for only the first-time switchers' (i.e. firm-destination pairs at the first change in RTA depth occurred at t), to the outcome change of firm-destination pairs whose treatment has remained stable until t.[41] We present the results of the estimation with $l \in [0, 3]$ in Table B7. In line with the event study approach reported in Figure 2, the increase in the depth of RTAs has a positive effect on the export of firms already one year after the change in the depth (t+1). The effect of RTA depth becomes strongly significant and larger in magnitude at $t+1$ and $t+2$.

Table B1: The trade effect of RTAs depth. Robustness check using OLS estimations.

	Exp	Exp	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DTA_{ijt}	0.003^{***}						
	(0.001)						
DTA_{ijt} leg.		0.005^{***}					
		(0.001)					
DTA_{ijt} WTO+			0.005^{**}				
			(0.002)				
DTA_{ijt} WTO-X				0.014^{***}			
				(0.003)			
DTA_{ijt} Core					0.004^{**}		
					(0.002)		
DTA_{ijt} NTM						0.037^{***}	
						(0.009)	
DTA_{ijt} Ser-Inv-Com							0.037^{***}
							(0.013)
$Ln(1+\tau_{ijt})$	-0.497^{***}	-0.497^{***}	-0.498^{***}	-0.498^{***}	-0.498^{***}	-0.497^{***}	-0.498^{***}
	(0.057)	(0.057)	(0.057)	(0.057)	(0.057)	(0.057)	(0.057)
Firm-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dest.-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Orig.-Dest. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,534,360	1,534,360	1,534,360	1,534,360	1,534,360	1,534,360	1,534,360

Notes: OLS estimates. Standard errors in parenthesis are clustered at firm-destination and origin-destination-year level. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

[41] See De Chaisemartin and d'Haultfoeuille (2024) for more details on the robust two-way fixed effects estimator.

Table B2: The trade effect of RTAs depth. Cross-section identification.

	Exp	Exp	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DTA_{ijt}	0.005***						
	(0.001)						
DTA_{ijt} leg.		0.018***					
		(0.002)					
DTA_{ijt} WTO+			0.027***				
			(0.002)				
DTA_{ijt} WTO-X				0.032***			
				(0.004)			
DTA_{ijt} Core					0.021***		
					(0.002)		
DTA_{ijt} NTM						0.079***	
						(0.012)	
DTA_{ijt} Ser-Inv-Com							0.119***
							(0.013)
Distance (ln)	-0.404***	-0.357***	-0.357***	-0.388***	-0.360***	-0.419***	-0.381***
	(0.018)	(0.019)	(0.018)	(0.019)	(0.019)	(0.017)	(0.018)
Contiguity	0.305***	0.312***	0.314***	0.306***	0.315***	0.281***	0.331***
	(0.032)	(0.029)	(0.029)	(0.031)	(0.029)	(0.029)	(0.031)
Language	0.145***	0.151***	0.145***	0.157***	0.147***	0.131***	0.137***
	(0.025)	(0.023)	(0.023)	(0.025)	(0.023)	(0.024)	(0.023)
Colony	0.229***	0.208***	0.206***	0.199***	0.216***	0.213***	0.225***
	(0.073)	(0.073)	(0.073)	(0.073)	(0.073)	(0.073)	(0.074)
Firm-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dest.-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Orig.-Dest. FE	No	No	No	No	No	No	No
Observations	2,120,273	2,120,273	2,120,273	2,120,273	2,120,273	2,120,273	2,120,273

Notes: PPML estimates. Standard errors in parenthesis are clustered at firm-destination and origin-destination-year level. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

Table B3: The trade effect of RTAs depth. Cross-section identification controlling for applied tariffs.

	Exp	Exp	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DTA_{ijt}	0.003***						
	(0.001)						
DTA_{ijt} leg.		0.013***					
		(0.002)					
DTA_{ijt} WTO+			0.020***				
			(0.002)				
DTA_{ijt} WTO-X				0.022***			
				(0.004)			
DTA_{ijt} Core					0.015***		
					(0.002)		
DTA_{ijt} NTM						0.045***	
						(0.012)	
DTA_{ijt} Ser-Inv-Com							0.085***
							(0.013)
$\text{Ln}(1+\tau_{ijt})$	-0.589***	-0.560***	-0.549***	-0.602***	-0.553***	-0.608***	-0.574***
	(0.071)	(0.071)	(0.071)	(0.071)	(0.071)	(0.072)	(0.071)
Distance (ln)	-0.442***	-0.393***	-0.394***	-0.423***	-0.397***	-0.452***	-0.413***
	(0.016)	(0.018)	(0.017)	(0.017)	(0.017)	(0.016)	(0.017)
Contiguity	0.230***	0.251***	0.254***	0.236***	0.253***	0.217***	0.261***
	(0.030)	(0.029)	(0.029)	(0.030)	(0.029)	(0.029)	(0.030)
Language	0.094***	0.108***	0.103***	0.106***	0.104***	0.084***	0.097***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Colony	0.243***	0.229***	0.227***	0.225***	0.234***	0.236***	0.239***
	(0.077)	(0.076)	(0.076)	(0.077)	(0.077)	(0.077)	(0.077)
Firm-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dest.- Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Orig.- Dest. FE	No	No	No	No	No	No	No
Observations	1,534,387	1,534,387	1,534,387	1,534,387	1,534,387	1,534,387	1,534,387

Notes: PPML estimates. Standard errors in parenthesis are clustered at firm-destination and origin-destination-year level. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

Table B4: The trade effect of RTAs depth. Robustness check using firm-destination fixed effects.

	Exp	Exp	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DTA_{ijt}	0.002***						
	(0.001)						
DTA_{ijt} leg.		0.003***					
		(0.001)					
DTA_{ijt} WTO+			0.003**				
			(0.002)				
DTA_{ijt} WTO-X				0.009***			
				(0.002)			
DTA_{ijt} Core					0.003**		
					(0.001)		
DTA_{ijt} NTM						0.013	
						(0.009)	
DTA_{ijt} Ser-Inv-Com							0.029***
							(0.010)
$\text{Ln}(1+\tau_{ijt})$	-0.204***	-0.205***	-0.205***	-0.205***	-0.205***	-0.205***	-0.205***
	(0.074)	(0.074)	(0.074)	(0.074)	(0.074)	(0.074)	(0.074)
Firm-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dest.-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Orig.-Dest. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,526,082	1,526,082	1,526,082	1,526,082	1,526,082	1,526,082	1,526,082

Notes: PPML estimates. Standard errors in parenthesis are clustered at firm-destination and origin-destination-year level. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

Table B5: The trade effect of RTAs depth. Robustness check using sub-sample of country pairs that switch RTA status during the period.

	Exp	Exp	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DTA_{ijt}	0.003^{***}						
	(0.001)						
DTA_{ijt} leg.		0.003^{**}					
		(0.001)					
DTA_{ijt} WTO+			0.004^{**}				
			(0.002)				
DTA_{ijt} WTO-X				0.010^{***}			
				(0.004)			
DTA_{ijt} Core					0.003^{**}		
					(0.002)		
DTA_{ijt} NTM						0.020	
						(0.013)	
DTA_{ijt} Ser-Inv-Com							0.028^{**}
							(0.014)
$Ln(1+\tau_{ijt})$	-0.792^{***}	-0.792^{***}	-0.792^{***}	-0.792^{***}	-0.792^{***}	-0.792^{***}	-0.792^{***}
	(0.115)	(0.115)	(0.115)	(0.115)	(0.115)	(0.115)	(0.116)
Firm-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dest.-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Orig.-Dest. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,068,395	1,068,395	1,068,395	1,068,395	1,068,395	1,068,395	1,068,395

Notes: PPML estimates. Standard errors in parenthesis are clustered at firm-destination and origin-destination-year level. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

Table B6: The non-linear trade effect of RTAs depth. PPML estimator.

	Exp	Exp	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DTA_{ijt}	0.002						
	(0.002)						
DTA_{ijt} sq.	0.000						
	(0.000)						
DTA_{ijt} leg.		0.006^{***}					
		(0.002)					
DTA_{ijt} leg. sq.		-0.000^{**}					
		(0.000)					
DTA_{ijt} WTO+			0.041^{***}				
			(0.008)				
DTA_{ijt} WTO+ sq.			-0.003^{***}				
			(0.001)				
DTA_{ijt} WTO-X				0.017^{***}			
				(0.005)			
DTA_{ijt} WTO-X sq.				-0.001^{***}			
				(0.000)			
DTA_{ijt} Core					0.028^{***}		
					(0.006)		
DTA_{ijt} Core sq.					-0.002^{***}		
					(0.000)		
DTA_{ijt} NTM						-0.004	
						(0.053)	
DTA_{ijt} NTM sq.						0.014	
						(0.028)	
DTA_{ijt} Ser-Inv-Com							0.068^{**}
							(0.033)
DTA_{ijt} Ser-Inv-Com sq.							-0.018
							(0.015)
$\text{Ln}(1+\tau_{ijt})$	-0.578^{***}	-0.578^{***}	-0.577^{***}	-0.578^{***}	-0.577^{***}	-0.578^{***}	-0.578^{***}
	(0.073)	(0.073)	(0.073)	(0.073)	(0.073)	(0.073)	(0.073)
Firm-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dest.-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Orig.-Dest. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,534,360	1,534,360	1,534,360	1,534,360	1,534,360	1,534,360	1,534,360

Notes: PPML estimates. Standard errors in parenthesis are clustered at firm-destination and origin-destination-year level. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

Table B7: Robustness check using the estimator by Chaisemartin D'Haultfoeuille (2024).

	Bilateral exports (ln)			
	t	*t+1*	*t+2*	*t+3*
DTA_{ijt} (1 if positive change in # of provisions)	0.013	0.038*	0.059**	0.091***
	(0.014)	(0.020)	(0.022)	(0.025)
Observations	792,277	727,223	694,221	619,297
# of switchers	26,596	26,233	24,614	22,138

Notes: dependent variable is the log of exports from the country i to the country j at time t in current million dollars. t is the year of the first depth variation in the country-pair. The reference period is t-1. We control for tariffs. Standard errors clustered at the country-pair level are in parentheses. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

Appendix C. Robustness checks on firm heterogeneity results

Weighted sum of provisions. In Table C1 we test the robustness of our results on the heterogeneous effect of RTAs depth using the weighted sum of provisions as a proxy of their depth. Results in Table C1 strongly support the robustness of our main estimation. While large and high-productive firms benefit from Deep RTAs, small firms are negatively impacted.

Excluding high-concentration sectors. In high-concentrated sectors firms may have very peculiar pricing strategy, almost unrelated to easier (or tougher) market access in presence (absence) of Deep RTAs. So, in Table C2 we test the robustness of our results by excluding high-concentrated sectors. Namely, we calculate the Herfindahl-Hirschman (HH) index of export sales of firms in a given origin-country and sector (we consider the main sector of the firm), and exclude from the sample those firms in country-sector combinations belonging to the top 10 percentile in HH index. Results in Table C2 show the robustness of our baseline results both sign, magnitude and significance of coefficients.

Binned model. In Table C3, we estimate the heterogeneous effect of RTAs depth by firms characteristics using a binned model, where the size bin of firms are based on their total exports in the initial year t_0. Small firms are those below the 25^{th} percentile in total export at t_0, big firms are those above the 75^{th} percentile in total export at t_0, and medium-size firms are those in between.

Table C1: The heterogeneous effect of RTAs depth. Robustness check using the weighted sum of provisions as a proxy for the RTA depth.

	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)
Weigh. DTA_{ijt} leg.	-0.008*	-0.009**	-0.001	0.004*
	(0.004)	(0.003)	(0.003)	(0.002)
Weig. DTA_{ijt} leg. × ($k_f > 75^{th}$ t0)	0.014***			
	(0.004)			
Weig. DTA_{ijt} leg. × ($k_f > 90^{th}$ t0)		0.018***		
		(0.003)		
Weig. DTA_{ijt} leg. x GVC			0.012***	
			(0.004)	
Weig. DTA_{ijt} leg. x GVC bil.				0.004
				(0.003)
$\text{Ln}(1+\tau_{ijt})$	-0.578***	-0.575***	-0.580***	-0.570***
	(0.073)	(0.073)	(0.073)	(0.073)
Firm-Year FE	Yes	Yes	Yes	Yes
Dest.-Year FE	Yes	Yes	Yes	Yes
Orig.-Dest. FE	Yes	Yes	Yes	Yes
Observations	1,534,360	1,534,360	1,534,360	1,534,360

Notes: PPML estimates. Standard errors in parenthesis are clustered at firm-destination and origin-destination-year level. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

Table C2: The heterogeneous effect of RTAs depth. Robustness check excluding high-concentration (HHI) sectors.

	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)
DTA_{ijt} leg.	-0.004**	-0.005***	-0.000	0.003**
	(0.002)	(0.002)	(0.001)	(0.001)
DTA_{ijt} leg. × ($k_f > 75^{th}$ t0)	0.009***			
	(0.002)			
DTA_{ijt} leg. × ($k_f > 90^{th}$ t0)		0.011***		
		(0.002)		
DTA_{ijt} leg. × GVC			0.008***	
			(0.002)	
DTA_{ijt} leg. × GVC bil				0.004**
				(0.002)
$\text{Ln}(1+\tau_{ijt})$	-0.577***	-0.573***	-0.581***	-0.571***
	(0.073)	(0.073)	(0.073)	(0.073)
Firm-Year FE	Yes	Yes	Yes	Yes
Dest.-Year FE	Yes	Yes	Yes	Yes
Orig.-Dest. FE	Yes	Yes	Yes	Yes
Observations	1,531,758	1,531,758	1,531,758	1,531,758

Notes: PPML estimates. Standard errors in parenthesis are clustered at firm-destination and origin-destination-year level. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

Table C3: The heterogeneous trade effect of RTAs depth by firm characteristics. Binned model.

	Exp	Exp
	(1)	(2)
DTA_{ijt} leg. × Big	0.004***	0.004***
	(0.001)	(0.001)
DTA_{ijt} leg. × Medium	-0.007***	-0.004
	(0.002)	(0.002)
DTA_{ijt} leg. × Small	-0.012***	-0.007
	(0.004)	(0.005)
$\text{Ln}(1+\tau_{ijt})$		-0.577***
		(0.073)
Firm-Year FE	Yes	Yes
Destination-Year FE	Yes	Yes
Origin-Destination FE	Yes	Yes
Observations	2,384,411	1,534,360

Notes: PPML estimates. Standard errors in parenthesis are clustered at firm-destination and origin-destination-year level. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

Appendix D. Theoretical framework

Following Mrázová and Neary (2020) we assume a given foreign market of size k having L identical consumers. With x being the amount the firm's sales to an individual consumer, the total foreign sales of the firm are $y = kLx$. We also assume that preferences are additively separable, so that the demand function is defined by the consumer's first order condition $u'(x) = \lambda p$, where λ is the marginal utility of income and plays the role of the level of competition in the market. In this setting, the firm's profits from selling abroad can be written as:

$$\pi\left(c,\lambda,k\right) = \max_{y}\left[p\left(y,\lambda,k\right) - c\right] y \tag{7}$$

where $p\left(y,\lambda,k\right) = \lambda^{-1}u'(y/kL)$ and c is the variable cost. By indicating with $\underline{c}$ and $\overline{c}$ respectively the firm with lowest and highest variable cost, the free entry condition can be written as:

$$\overline{\upsilon}(\lambda,k) = \int_{\underline{c}}^{\overline{c}} \upsilon(c,\lambda,k)g(c)dc = f_e \tag{8}$$

where $\upsilon(c,\lambda,k) = max\left[0, \pi\left(c,\lambda,k\right) - f\right]$, f_e is the sunk entry cost, and f is the fixed cost of production (common to all firms). The effect of the signature of a *deep* RTA on firms exports can be modeled as an increase in market size k, and obtained by totally differentiating foreign sales $r = py$:

$$\frac{d\ \log\ r\left(c,\lambda,k\right)}{d\ \log\ k} = \underbrace{\frac{\partial\ \log\ r\left(c,\lambda,k\right)}{\partial\ \log\ k}}_{\text{size-effect}} + \underbrace{\frac{\partial\ \log\ r\left(c,\lambda,k\right)}{\partial\ \log\ \lambda}\frac{d\ \log\ \lambda}{d\ \log\ k}}_{\text{competition-effect}} \tag{9}$$

where the change in competition $d\ \log\ \lambda/d\ \log\ k$ is determined by the free entry condition (2):

$$\frac{d\log\lambda}{d\log k} = -\frac{\partial\log\overline{\upsilon}/\partial\log k}{\partial\log\overline{\upsilon}/\partial\log\lambda} \tag{10}$$

By noticing that the total differentiation of foreign revenues is $\hat{r} = -\varepsilon\hat{\lambda} + (1-\varepsilon)\hat{p} + \hat{k}$, with $\varepsilon(c,\lambda,k)$ is the firm's demand elasticity, we have:

$$\frac{d\log r(c,\lambda,k)}{d\log k} = 1 + \left(-\varepsilon(c,\lambda,k)\right)\left(\frac{1}{\overline{\varepsilon}(\lambda,k)}\right) = \underbrace{1}_{\text{size-effect}} - \underbrace{\frac{\varepsilon(c,\lambda,k)}{\overline{\varepsilon}(\lambda,k)}}_{\text{competition effect}} \tag{11}$$

where $\overline{\varepsilon}(\lambda,k)$ is the weighted average elasticity of demand across firms, derived from eq.(2) by computing partial effect of λ on r for all firms weighted by their market share. With subconvex demand, the firm-specific elasticity of demand $\varepsilon(c,\lambda,k)$ decreases with sales. So, the signature of deep RTA increases (decreases) the foreign sales and market share of large (small) firms.

Appendix E. IV validity checks

DTAs and past exports growth. In Table E1, we follow Fajgelbaum et al. (2020) and qualitatively tests the absence of reverse causality concerns in our baseline estimations. Namely, in Table E1 we show the absence of correlation between firms' export growth in the period preceding a change in the RTA's depth and the extent of the *change* in RTAs' depth. Coefficients in Table E1 are "true" zero (i.e. small in magnitude and not-significant) and suggest the absence of any firm-specific trend in export growth affecting the subsequent dept in RTAs.

Plausible exogeneity test. In Table E2, we test the robustness of our 2SLS results to deviations from the perfect validity of the exclusion restriction assumption, and run the plausible exogeneity test as proposed by Conley et al. (2012). We first regress our outcome of interest X_{fijt} on the measure of RTAs depth DTA_{ijt} and IV using the same sets of fixed effects as in eq. (1). The coefficient on the IV, parameter ν in Table E2, can be fairly considered a proxy of the extent of deviation from perfect validity of the exclusion restriction – see van Kippersluis and Rietveld (2018). Reassuringly this coefficient is close to zero and not statistically significant, suggesting that the IV does not have a direct effect on the export of firms on top of the (potentially) endogenous variable. After using this parameter in the plausible exogeneity test we obtain point estimate's bounds strictly above zero. This suggests that the positive effect of RTAs depth is robust to deviation from perfect validity of the exclusion restriction assumption.

Table E1: Firms' export growth and change in RTAs' depth.

	Firm's average export growth before RTA depth change		
	(1)	(2)	(3)
Change in RTA depth	0.001	0.001	0.002
	(0.001)	(0.001)	(0.002)
Origin FE	Yes	No	Yes
Destination FE	No	Yes	Yes
Observations	137,896	137,896	137,896

Notes: OLS estimates. Standard errors in parenthesis are clustered at origin-destination level. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

Table E2: Deep trade agreements and the export of firms with plausibly exogenous instrument.

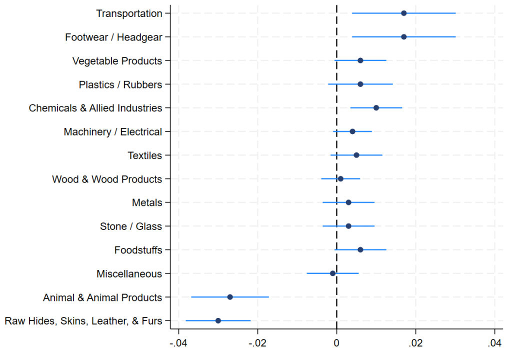
	Union of Confidence Interval estimations			
Dep Var	ν	Coeff. in tab 8	Min 90% CI	Max 90% CI
	Panel (a): Top-50 destinations.			
Firm Exports	0.0019 (0.0013)	0.009* (0.005)	0.006	0.037
	Panel (b): market share above 0.5%.			
Firm Exports	0.0019 (0.0013)	0.009* (0.005)	0.008	0.038

Notes: UCI based on γ coefficients from a regression of firm exports on the IV. Standard errors in parenthesis cluster by origin-destination-year. To meet the maximum number of covariates allowed by plausexog STATA command we had to reduce the number of destinations (i.e. number of origin-destination dummies). In panel (a) we use top-50 destination countries. In panel (b) we keep destinations counting for at least the 0.5% of total exports of all origin countries in the sample. For the same computational reasons, the plausibly exogeneity test has been conducted by replacing destination-year fixed effects by destination-period fixed effects (each period covering a six-year windows).

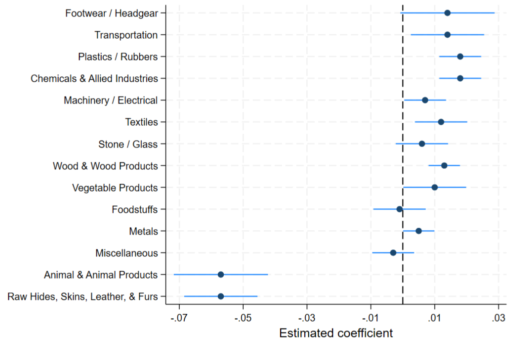
Appendix F. Additional results by sector

Sector-specific results by provision category. In figure F1, we show the effect of alternative measures of RTAs' depth on the exports of firms in each of the 14 sector headings of the HS classification (i.e. HS 1-digit). Namely, we replicate estimation as in eq. (4) for proxy of Deep RTAs based on : (i) WTO-plus provision (panel a), (ii) WTP-X provisions (panel b), (iii) Core provisions (panel c), (iv) NTM related provisions (panel d) and (v) Policy provisions (panel e).

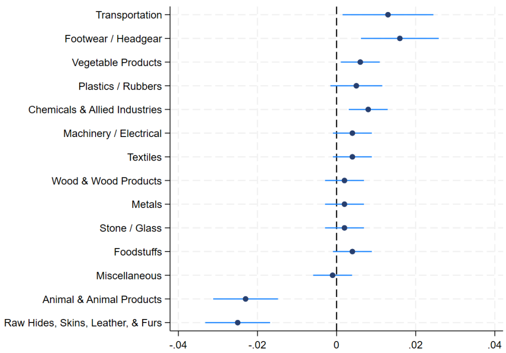
Figure F1: The impact of RTAs depth on firms' export by sector. Provision categories.



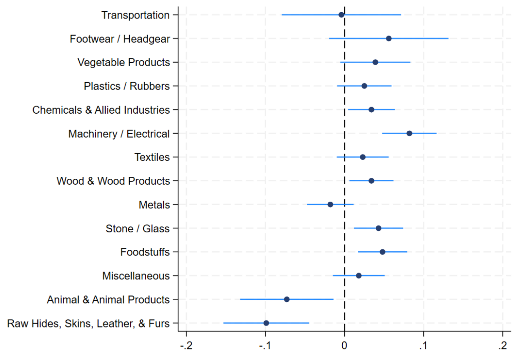
(a) DTA WTO+



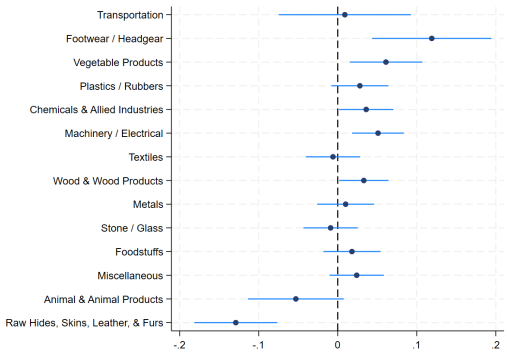
(b) DTA WTO X



(c) DTA Core



(d) DTA NTM



(e) DTA Policy

Source: Authors' calculations. *Notes:* This figure shows the effect of RTAs' depth (i.e. total count of legally enforceable provisions) on the export of firms by provision category. 90% confidence intervals are shown for our parameter of interest.