CESifo WORKING PAPERS

10436
2023

May 2023

Deep Trade Agreements and Heterogeneous Firms Exports

Matteo Neri-Lainé, Gianluca Orefice, Michele Ruta

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10436*

Deep Trade Agreements and Heterogeneous Firms Exports

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, the analysis shows that the depth of trade agreements matters for the export performance of firms. Moving from *shallow* to *deep* trade agreements boosts firms' exports, on average, by 3.6 percent. In line with models of trade with heterogeneous firms and mark-ups, the trade impact of deep trade agreements depends on the firm's characteristics. The impact is stronger for large firms and firms involved in global value chains and is negative for small firms. Robustness tests, an event study approach and an Instrumental Variable strategy confirm the causal interpretation of the results. These heterogeneous impacts on firms' exports imply a selection (*pro-competitive*) effect of deep trade agreements with significant welfare consequences for signatory countries.

JEL-Codes: F130, F140, F150.

Keywords: deep trade agreements, exports, firm heterogeneity, developing countries.

Matteo Neri-Lainé
University Paris-Dauphine PSL
Place du Maréchal de Lattre de Tassigny
Paris / France
matteo.neri-laine@dauphine.psl.eu

*Gianluca Orefice**
University Paris-Dauphine PSL
Place du Maréchal de Lattre de Tassigny
Paris / France
gianluca.orefice@dauphine.psl.eu

Michele Ruta
International Monetary Fund
700 19th Street N.W.
USA – Washington, D.C. 20431
mruta@worldbank.org

*corresponding author

This draft: May 15, 2023
We would like to thank Ana Fernandes, Nuno Limao, Andrea Lassmann, Philip Sauré and seminar participants at the World Bank, Indiana University, University of Mainz 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.

1 Introduction

During the last two decades, 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). 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 than at the beginning of the period, with 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 as "deep RTAs" or simply "Deep Trade Agreements" (DTAs hereafter).

The effects of regional trade agreements on trade have been largely studied, but there remain some important gaps in the literature. In part motivated by the changing depth of RTAs, a number of studies have empirically investigated the trade effects of different types of trade agreements. 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 can help explain different trade outcomes (Orefice and Rocha, 2014; Mattoo et al., 2017). At the same time, theoretical models have emphasized a different type of heterogeneity: a common reduction in trade costs affects different types of firms differently. Larger and more productive firms can more easily take advantage of the change 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 trade agreements can reduce trade costs among members by eliminating tariffs and by reducing other frictions due to regulatory differences and/or policy uncertainty, a natural extension of this literature is that DTAs should affect firms differently based on their characteristics, promoting the export performance of larger firms and leading to a selection effect.

This paper empirically investigates the effect of RTAs' depth 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 deep RTAs on trade come from the intensive and extensive margins of trade. Importantly, we carefully investigate 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.

To assess the impact of deep RTAs on the export performances of firms, we 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 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. The baseline results show that one additional legally enforceable policy area in RTAs boosts the exports of firms by 0.3%; corresponding to a 3.6% increase in firms' exports when moving from *shallow* to *deep* RTAs.[1] This effect is larger for policy areas in RTAs that are not regulated by the World Trade Organization (WTO-extra) agreements, such as investment or competition. One additional legally enforceable WTO-extra provision increases the firm's exports by 0.7%. The instrumental variable approach and the plausible exogeneity test both suggest the causal interpretation of our results.

This average effect of deep RTAs on the export performance of firms may hide substantial heterogeneity across firms with different characteristics. In New New Trade Theory models à la Melitz (2003), firms are heterogeneous and a reduction in fixed and/or variable export cost is expected to favour large and highly productive firms' exports at the expense of low-productivity firms, leading to a selection effect. This selection effect is magnified in the presence of heterogeneous demand elasticity: if large and more productive firms face smaller demand elasticity and higher mark-ups, the pro-competitive effect induced by deep RTAs exacerbates the reduction in foreign sales of less productive firms (in line with Atkeson and Burstein 2008 and Crowley et al. 2022). Our second set of estimations tests the effect of deep RTAs on heterogeneous exporters.[2]

In line with the New New Trade Theory findings, we uncover strong heterogeneous effects

[1] 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).

[2] 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 5.2.

of deep RTAs on firms with different characteristics. While large and high-productivity firms benefit from deep RTAs, small and less productive firms suffer the increased within-origin competition at destination and export less – a *pro-competitive* effect of deep trade agreements. On average, including an 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 0.5%-0.6%.[3] Interestingly, and in line with the argument in Limão (2016),[4] high-productivity firms participating in Global Value Chains (GVC), i.e. firms that export and import to/from the same country, benefit the most from the signature of deep RTAs. For these firms, an additional legally enforceable provision in RTAs stimulates exports by 0.8%. The *pro-competitive* effect of deep trade agreements shows that using firm-level data is key to understand the welfare implications of RTAs. Specifically, by favouring exports of large and high-productivity firms and reducing exports of small and less productive firms, deep RTAs promote a reallocation of resources from the latter to the first, which entails adjustment costs but also leads to an overall increase in the average productivity of firms in the exporting country.

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 behaviour to impute the extensive *versus* intensive margin effect of deep RTAs. We show that the extensive margin channel has only a slightly larger contribution (60%) than the intensive margin channel (40%). This evidence, in light of the theoretical predictions in Chaney (2008), suggests that, indeed, deep RTAs represent a reduction in both the fixed and the variable export cost components. The second contribution of this paper is therefore providing a novel firm-level evidence on the channel through which deep RTAs boost aggregate exports.

Finally, we study the dynamic effects of deep trade agreements on firms' exports. The positive effect of *bilateral* deep RTAs may vanish over time because of the worldwide increase in RTAs and their depth and the inclusion in RTAs of non-discriminatory provisions that *de*

[3]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 favourable 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.

[4]In discussing the economic consequences of deep PTAs, 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 [...].*"

facto can reduce trade costs for exporters in third countries (Lee et al., 2019; Dai et al., 2014). 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 two years after the shock. The effect of the change in depth of RTAs vanishes afterwards. The event study approach also shows the validity of the parallel trend assumption and reinforces our baseline results' causal interpretation.

The rest of the paper is organized as follows. Section 2 discusses the possible theoretical channels underlying the firm-specific effect of deep trade agreements. Section 3 presents the data used in the analysis and proposes some descriptive evidence. Section 4 discusses the empirical strategy. Section 5 shows the baseline results and the heterogeneity tests. In section 6 we address the potential endogneneity problem. Section 7 tests the dynamic effect of deep RTAs. Concluding remarks follow.

2 Theoretical framework

Deep trade agreements are wider in scope and more complex than shallow trade agreements as they go beyond a standard improvement in market access via reductions in bilateral import tariffs. Specifically, by filling regulatory gaps among member countries, DTAs may reduce fixed export costs and uncertainty in bilateral trade relationships. These features make any attempt to set a general theoretical framework extremely complicated (and likely unsatisfactory). However, some general predictions can be drawn. The reduction in fixed or variable export costs and the reduction in uncertainty associated with DTAs unambiguously increase market access at destination, thus make it easier for firms of a given origin to export to 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 DTAs 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. But firms differ also along other relevant dimensions. First, firms that participate in global value chains (i.e. GVC firms) can be more affected by DTAs 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 destination market, so that the reduction of trade costs and uncertainty associated to DTAs affects them both on the import and export side – see Limão (2016). A second way firms differ is in the demand elasticity they face. Large and high-productivity firms face lower demand elasticity (Spearot, 2013) and have larger markups than small and less productive firms (Atkeson and Burstein, 2008). This implies that, after the entry into force of a DTA, large and high-productivity firms at origin may reduce their export price at destination to a larger extent to gain market share, exerting a tougher "within origin" competition on small and less productive firms. Such a strong heterogeneous impact of DTAs on firms with different characteristics is even starker in a framework in which the degree of substitutability across varieties of a given origin country is stronger than that between origin countries – see Crowley et al. (2022). In these settings, DTAs may have very different consequences on the export performance of large and small firms at destination, with large firms unambiguously benefiting the most from the trade agreement.

The impact of DTAs on small and less productive exporters also depends on: (i) whether the new agreement has an impact on the market structure of the origin country, and on (ii) whether the DTA modifies the aggregate export demand (i.e. the price index at destination). If the number of exporters is unaffected by the entry into force of a deep trade agreement, small incumbent exporters may benefit through an increase in exports even if these gains are milder relative to large exporters. Conversely, if the DTA has an impact on the market structure in the origin country by increasing the number of exporters, small exporters may also experience a reduction in their exports (and market share) at destination, and eventually may be induced to exit from the export market due to enhanced competition (*pro-competitive* effect of DTAs). In the same vein, when a given origin is populated by few and highly granular exporters, the signature of a DTA may have an effect on the price index and hence on the aggregate goods' expenditure at destination – see Handley and Limão (2017). In this specific case, if high-productive firms react to the signature of a DTA by reducing prices more than low-productive firms, these last will face a stronger reduction in the aggregate expenditure at destination and may be negatively impacted by the signature of a PTA.[5]

All in all, while DTAs are expected to have an unambiguously positive effect on the export value of large and high-productivity firms, the effect on small and less productive firms is an empirical question. The within-origin competition effect of DTAs can be so strong to offset the reduction of fixed and variable costs and induce the exit of small firms from the destination market. This effect has relevant welfare implications. If small and less productive firms exit from a specific destination market and large and high-productivity firms survive and thrive, resources are reallocated from low- to high-productivity firms. This selection effect implies adjustment

[5]This effect has been often neglected in the literature by assuming a given firm sufficiently small relative to the total number of firms in the country implying *exogenous* price index and aggregate demand parameter – a notable exception is Handley and Limão (2017). In this paper we consider a set of poor and developing countries in which exporting firms are granular and the assumption of *exogenous* price index is less plausible.

costs, but also an increase in the average productivity of firms. The heterogeneous trade effects of DTAs on low- *versus* high-productivity firms, and the consequent selection effect, is at the core of the empirical investigation in the rest of this paper.

3 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, and (ii) the Content of Deep Trade Agreements (Hofmann et al., 2017) on the content of 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.

We build our final dataset in several steps. First, we reorganize 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 provisions included in the trade 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 (independently of 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) the number of WTO-extra provisions (WTO-X) – i.e. provisions not covered by the current mandate of the WTO, and (iv) number of core provisions – i.e. provisions directly related to trade enhancing factors.[6] 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.[7] The simple

[6]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.

[7]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, 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.

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).[8]

One drawback of the DTA database is that it does not cover RTAs that are no longer in force.[9] So, when controlling for the presence of an RTA between two countries (see section 4), we complement the DTA database with the CEPII data covering active and inactive RTAs. The *RTA adjusted* dummy is equal to one if a RTA is observed (from CEPII and/or DTA dataset).

The Exporter Dynamics Database (EDD) takes the form of multiple databases: one per country.[10] 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 complete firm-level dataset. 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. It must be noted that the coverage of the EDD varies by country: for some countries we have data for the entire 2000-2020 period, while for some other countries we have a more limited time period (cf. table A1 in appendix). Although original firm-level export data are (also) product HS6-digit specific, our main variable of interest is not, so we aggregate export data at the level of firm-destination-year (sum across HS 6-digit products for each firm-destination-year combination).

As described in Table 1, the majority of exporters included in the EDD have at least one active RTA in force (i.e. 48 out of 55). However, only a sub-sample of countries (and country-pairs) have changed the depth of their RTAs in the period 2000-2020 (i.e. a newly signed RTA or amendment of pre-existing RTAs). Namely, 31 exporting countries signed new RTAs or amended a pre-existing one, giving *within* variation to our measures of RTAs depth. This translates into 701 country-pairs having time variation in the depth of RTAs in the period 2000-2020. This

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.

[8]In Figure A1 we show how frequently each provision is included in Regional Trade Agreements.

[9]The World Bank provides information on the content of RTAs active in the year of the creation of the DTA database. RTAs that were active in the past, but inactive at the time of the creation of the database are not covered.

[10]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).

time variation comes mainly from the entry into force of new trade agreements.[11] Since in the empirical strategy we rely on the within country-pair variation in RTAs' depth, the final estimation sample considers only 31 exporting countries and contains 4,659,362 observations with non-missing information on export value and content of RTAs. The estimation sample shrinks to 2,924,126 observations because of missing values in the tariff 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) and firm-destination specific exports (panel c and d) for the sub-sample of the 15 countries that provide much of the variation in RTA depth.[12] 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). It clearly emerges that the presence of an RTA and its depth matters for the aggregate country-pair exports. In panel (c) of Figure 1 we show the empirical distribution of firm-destination export sales depending on whether the exporting country has (or not) an active RTA with the destination market. In panel (d) of Figure 1 we consider only RTA partners (i.e. pairs with RTA in force), and distinguish between destinations at the top- and bottom-quartile of the RTA depth. There is only small evidence that RTA depth matters for the export of firms. The positive effect of RTAs and their depth is less stark at the firm level than at aggregate country-pair level, suggesting a strong heterogeneous effect of DTAs on firm-level export. While some firms largely benefit from deep RTAs (boosting the aggregate exports of the country), other firms are hindered by the presence of deep RTAs. The empirical distribution of firms' exports is thus only marginally affected by the presence of deep trade agreements. In what follows we carefully test such a heterogeneous effect of deep RTAs.

Finally, in Table 2 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.

[11] In our final sample, 72 RTAs involve more than two members, while 2 are bilateral for a total of 74 agreements contributing to the within-variation of our RTA depth variable.
[12] In Figure A2 we provide the same evidence for the full set of 31 countries with time variation in RTA depth.

4 Identification strategy

This section discusses the empirical strategy adopted to test the effect of deep RTAs on the exports of firms. Our baseline specification is as follows:

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

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 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 and "core" legally enforceable provisions in RTAs. 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. multilateral resistance term on the exporter side). Any country-specific technological or productivity shock, as well as the distribution of firms' productivity in the exporting country, are implicitly captured by firm-year fixed effects. In equation (1) we also control for the multilateral resistance term in the importer side by including destination-year fixed effects θ_{jt} capturing also any demand shock that affects the import demand at destination. Finally, any country-pair (time-invariant) factor, such as geographical distance and any other gravity-type covariate is captured by country-pair fixed effects θ_{ij}.

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 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 A2 we perform a pre-trend test correlating the (average) firms' export growth before a change in RTA depth and the extent of the change in the depth of RTAs. Table A2 shows the absence of correlation between firms' export growth and the change in RTAs' depth supporting qualitatively the absence of reverse causality issue. Nevertheless, in section 6 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 6.1, 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 7.

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 (DTA_{ijt}). This problem is exacerbated when country-pair fixed effects (θij) are included in the estimation and absorb any cross-country-pair variability in the presence of RTAs and depth. For this reason, in our baseline estimations we do not control for the presence of an RTA; this is *de facto* subsumed by the DTA_{ijt} variable when larger than zero. However, since the main and ever-present objective of any RTA is to reduce bilateral tariffs, we control for the presence of an 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, $ln\,(1+\tau_{fijt})$. We use the product share of the firm's exports in the initial year as a weight in averaging the applied tariffs faced by each firm at destination across exported products. As an alternative test, we disregard the collinearity problem and propose a robustness check explicitly controlling for the RTA_{ijt} dummy (namely the *RTA adjusted* dummy discussed in section 3).

In order to test the heterogeneous effect of deep RTAs 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:

$$\begin{aligned} X_{fijt} &= exp\Bigg[\theta_{ft} + \ \theta_{jt} + \theta_{ij} + \beta_1 DTA_{ijt} + \beta_2 ln\,(1+\tau_{fijt}) \\ &+ \ \beta_3 \left(DTA_{ijt} \times I(k_f > \bar{k})\right)\Bigg] \times \varepsilon_{fijt}. \end{aligned} \tag{2}$$

The indicator $I(k_f > \bar{k})$ is equal to one if a given firm's characteristics k_f is above a threshold $\bar{k}$. Four 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 (across destinations and years),[13] and define dummy variables equal to one if the size of the firm is in turn above the 75^{th} and 90^{th} percentile

[13] 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).

of the distribution. This is an intuitive but coarse proxy of firm size. It can be endogenously affected by the presence of deep trade agreements among firm's destinations. To address this problem, 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). Firm's total export at time t_0 is used to define our second indicator variable $I(k_f > \bar{k})$. Percentiles 75^{th} and 90^{th} of the initial firms' exports distribution are used as threshold $\bar{k}$. Third, 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.

We adopt a PPML estimator to address the heteroskedasticity problem in structural gravity model for trade (Santos-Silva and Tenreyro, 2006), and cluster standard errors by origin-destination-year (i.e. the source variation of our main variable of interest). As a benchmark, in Appendix Table A3 we show OLS estimations.[14]

5 Results

This section discusses our main results on the effect of deep RTAs. 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 5.1. In the same section, we propose a robustness check using the weighted count of provisions as an alternative measure of RTA depth. In section 5.2, we estimate equation (2) and show the heterogeneous effect of deep RTAs on firms with different characteristics. Finally, in section 5.3, we disentangle the effect of deep RTAs into extensive *versus* intensive margins of exports.

5.1 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 or core). 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 75^{th} percentile

[14]In the log-linear OLS estimations we set to zero the log of zero firms' exports.

of the distribution of the RTAs depth implies a 3.6% increase in firm exports.[15] If we consider the count of WTO+ provisions, moving from shallow to deep RTAs implies a 2.7% increase of firm exports (i.e. one additional legally enforceable WTO+ provision boosts the exports of firms by 0.3%).[16] 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).[17] 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 disregard the product dimension of both export and tariffs, and the consequent aggregation bias produces tariff elasticity that are smaller in magnitude (Redding and Weinstein, 2019).[18]

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 RTAs containing rare provisions. 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 deep RTAs 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.

The effect of RTA depth is robust to the inclusion of dummies for the presence of an RTA between country i and j at time t. See results reported in Table A6. To take into account the presence of non-mapped active RTAs (i.e. those RTAs not included in the World Bank database), in Table A6 we take the list of RTAs from CEPII and assign a value of depth respectively equal to one (see columns 1-2), or equal to the closest (in time) country-pair's RTA depth to non-mapped active RTAs (see columns 3-4). As expected, the presence of an empty (active) RTA –

[15]RTAs at the 75^{th} percentile of legally enforceable provisions contain 14 provisions.

[16]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.

[17]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.

[18]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 A4 and A5. 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.

i.e. an RTA with zero depth – has a null effect on the export of firms (see coefficient on RTA *adjusted* in table A6 columns 1-3). The presence of a RTA has positive effects on export only if the agreement has some depth (i.e. positive number of provisions) - see columns (1),(2) and (3). While reassuring, results in Table A6 must be taken *cum grano salis* because of the high collinearity between the RTA dummy and the measure of depth in regional trade agreements. As a last robustness check, we estimate equation (1) on the sub-sample of country-pairs that change RTA status during the period 2000-2020. Results, reported in Table A7 support the robustness of our baseline results.

The baseline and the robustness check regressions discussed so far suggest that deep RTAs have 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, and dynamic effect after the change in depth of an RTA. We dig more into the heterogeneous effects of deep RTAs in the next section.

5.2 Firm heterogeneity

In line with the theoretical discussion in section 2, deep RTAs may have a different effect on low- *versus* high-productivity firms and/or on firms with different involvement in GVCs. This section explores the heterogeneous effects of deep RTA across different types of firms based on size and GVC participation. Namely, we interact the DTA_{ijt} variable with four firm-specific indicators: (i) large firm dummy equal to one if the total exports of the firm (across destination and years) are in turn above the 75^{th} and 90^{th} percentile of the distribution, (ii) large firm dummy if the exports of the firm at t_0 is in turn above the 75^{th} and 90^{th} percentile of the distribution, (iii) a dummy equal to one if the exporting firm imports some products (i.e. GVC firm indicator), and (iv) a dummy variable equal to one if the firm exports and imports to/from the same destination j (*GVC bilateral*).

Results reported in Table 5 show an interesting and robust pattern: large and GVC firms benefit from deep RTA.[19] For small firms, and firms that do not participate in GVCs, the depth of RTAs has a negative effect on exports. Specifically, for large high-productivity firms, one additional legally enforceable provision implies a 0.3% - 0.6% increase in exports (see columns 1 - 4). This means that, moving from *shallow* to *deep* RTAs implies a 3.6% - 7.2% increase in exports. For GVC firms exporting/importing to/from the same country j, moving from *shallow* to *deep* RTAs implies a 9.6% increase in exports (see columns 6). These results uncover

[19] 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 A8.

an interesting pro-competitive effect of deep trade agreements. 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 using the weighted sum of provision as a proxy for the depth of RTAs (see Table A9).

The negative effect on small firms is in line with trade models with heterogeneous elasticity of demand (Atkeson and Burstein, 2008; Crowley et al., 2022) that predict a strong within-country competition effect hurting less productive firms, and implying an increase in the concentration of foreign market shares among high-productivity firms. So, deeper RTAs favour the export sales of firms with certain characteristics (large, more productive and GVC firms) at the expense of small and less productive firms. Such a selection effect of deeper RTAs is shown in Table 6 where we estimate the effect of deep RTAs 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.[20] Deep RTAs reduce the number of firms surviving in the export market (columns 1-2) but increase the average export (value) *per* firm (columns 3-4).[21] This confirms the selection effect of DTAs 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.

5.3 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 channel), or may allow new firms to start exporting in such destination (i.e. *extensive* margin channel). 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 toward a given destination (intensive margin channel); and (ii) entry-exit firms to/from a specific market (extensive margin channel). 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

[20] PPML estimator is used in Table 6.

[21] A more compelling way of testing the pro-competitive 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.

$\hat{\beta}_{incumbent}$ 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 lower contribution from the intensive margin channel to the aggregate effect. Namely, the intensive (extensive) channel accounts for 40% (60%) of the total impact of deep RTAs. The not-too-different contribution of the intensive and extensive margin channel suggests that deep RTAs represent at the same time a reduction in both the fixed and the variable export cost component. 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.

In a trade model with Pareto distributed firm productivity as in Chaney (2008), the aggregate exports of country i to country j, and hence the extensive and the intensive margin of trade, depend on the firm productivity dispersion parameter. When low- and medium-productivity firms represent a large fraction of firms in the country (i.e. skewed productivity distribution), a change in the productivity threshold induced by higher market access is expected to have large aggregate export consequences because many firms enter the export market (large extensive margin effect). Conversely, when the distribution of firm productivity is more dispersed and large and high-productive firms represent a larger fraction of firms in the country, a change in the productivity threshold has only a marginal effect on the extensive margin, and the aggregate exports of the country increase only through the intensive margin channel. Thus, no matter the specific type of export costs impacted by deep RTAs, the *aggregate* effect of deep RTAs depends on the productivity distribution of firms in the exporting country.

In Table 8, we interact the RTA depth variable with two proxies of the country-specific Pareto distribution parameter (falling with an increase in the share of high-productive firms). One obtained by following the QQ approach in Head et al. (2014),[22] the other by following Gabaix and Ibragimov (2011).[23] To facilitate the interpretation of the results, we use a dummy variable for countries having above-the-median Pareto shape parameter. In line with the intuition, deep

[22] 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 $\sigma = 5$.

[23] To reduce any endogeneity concern, we use the beginning of the sample data to calculate Pareto share parameters.

RTAs have a strong positive effect on the extensive margin (i.e. total exports of entry/exit firms) of countries having very skewed productivity distribution (i.e. high Pareto shape parameter) – see columns (1)-(2). Interestingly, in countries with a larger share of highly productive firms (i.e. Pareto shape parameter below the median) the extensive margin of export is negatively affected. This is likely due to the selection effect of deep RTAs discussed above. When the exporting country is populated by a large share of high-productivity firms (low Pareto shape parameter), the within-origin pro-competitive effect of DTAs makes less-productive firms exiting the market. Results for the intensive margin channel are shown in columns (3)-(4). Deep RTAs have a strong positive effect on the intensive margin of countries having a large share of highly productive firms (i.e. Pareto parameter below the median) and a tiny/weak negative effect on countries having very skewed firm productivity distribution (i.e. large Pareto shape parameter).[24] Results on total aggregate exports are reported in columns (5)-(6) and reflect results on the extensive and intensive margin channels discussed above.

6 Endogeneity

The inclusion of the set 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 endogeneity 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:

$$IV_{ijt} = \left[\frac{1}{K-1}\sum_{k\neq j} DTA_{ikt}\right] \times \left[\frac{1}{Z-1}\sum_{z\neq i} DTA_{zjt}\right] \tag{3}$$

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.[25] 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

[24] In the regressions reported in table 8 we include origin-year, destination-year and origin-destination fixed effects.
[25] 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.

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 results even in the presence of weak deviation for the perfect validity of the exclusion restriction assumption.

Our baseline results are confirmed by the 2SLS estimations addressing any residual endogeneity concerns - see Table 9 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 9 columns 2-7. 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 9 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.

A limitation of *linear* 2SLS estimates in Table 9 is that they cannot be directly compared with our baseline *non-linear* PPML estimations. Also, 2SLS estimations require the log-linearization of the gravity equation and hence suffer the heteroscedasticity problem (Santos-Silva and Tenreyro, 2006). To address these issues, in line with Lin and Wooldridge (2019), we propose a two-stage OLS/PPML. In the first stage, we use the OLS estimator to obtain the predicted RTA depth based on the IV (i.e. standard first stage in 2SLS). Then, the predicted RTA depth is used as the main explanatory variable in a PPML estimator with bootstrapped clustered standard errors (second stage). Results reported in Table A10 support the robustness of our IV strategy.

6.1 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). We obtain a small and not statistically significant direct effect of IV on firm exports - see parameter ν in Table 10. We, therefore, 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. 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.[26]

7 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 and for a limited amount of time. In this section, we adopt an event-study approach to visualize the dynamic effect of deep RTAs. Namely, we follow Fajgelbaum et al. (2019), and compare the targeted varieties (i.e. firm-destination combinations that face a change in RTA depth) to non-targeted varieties, using the following specification:

$$\begin{aligned} X_{fijt} &= exp\Bigg[\theta_f + \theta_{jt} + \theta_{ij} + \sum_{z=-2}^{3} \beta_{0z}\mathbf{I}\left(event_{ijt} = z\right) \\ &+ \sum_{z=-2}^{3} \beta_{1z}\mathbf{I}\left(event_{ijt} = z\right) \times target_{fj}\Bigg] \times \varepsilon_{fijt} \end{aligned} \tag{4}$$

.

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 ($z = -2, -1$) and after ($z = 1, 2, 3$) the change in the RTA depth ($z = 0$).[27] The target

[26]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 10), (ii) keep only destinations representing at least the 0.5% of total exports of all the 31 exporting countries of our sample.

[27]We adopt an asymmetric time period before *versus* after a change in RTA depth because in our data the pre-treatment period is limited.

variable, $target_{fj}$, is a dummy for varieties (i.e. firm-destination) that experience a change in the RTA depth during the period. The presence of firm fixed effects implies that β_{1z} coefficients are identified using the variation between targeted and non-targeted destinations at each point in time z. For targeted varieties, the event date is the year of the first change in the RTA depth between country i and j. For non-targeted 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 origin-destination-year.

Figure 2 reports the impact of RTA depth on targeted varieties. On impact, we find a positive but imprecisely estimated effect of RTA depth on the export value of firms. The effect becomes statistically significant one and two years after the change in RTAs' content. The positive effect vanishes after two years from the change in RTA depth. This is coherent with RTAs including non-discriminatory provisions (Lee et al., 2019) that *de facto* reduce the trade costs also for third (non-RTA) countries, and hence imply the weakening of any preferential bilateral relationship among RTA signatory countries in the medium- and long-run. Interestingly, the event study approach also addresses concerns on the anticipation of changes in RTA depth (and/or RTA signature). During the pre-treatment period targeted and non-targeted varieties show a parallel trend. The absence of pre-trend reassures us about the causal interpretation of our baseline results.

7.1 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. Moreover, RTAs' depth is a continuous treatment that can change over time and may have a dynamic effect (i.e. the initial variation in RTA depth may affect future variations in depth and the outcome variable), implying a second possible source of bias.[28] Therefore, we follow De Chaisemartin and d'Haultfoeuille (2020) and perform a robust two-way fixed effects estimation of the trade effect of 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 (2020). The intuition behind this estimation is that to avoid negative weights that may bias standard two-way fixed effects estimators, one

[28] Notice that this second source of bias is already taken into account in the event study exercise where we focus exclusively on the first variation in RTAs depth.

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. [29]

We present the results of the estimation with $l \in [0, 3]$ in Table 11.[30] In line with the event study approach reported in Figure 2, the depth of RTAs has a weak positive effect on the export of firms in the year of the first variation in RTA depth (i.e. t). The effect of RTA depth becomes strongly significant the year after $t+1$, increases in magnitude at time $t+2$ and vanishes at $t+3$. While the pattern is the same as in the standard event study approach reported in the previous section, the point estimates in Table 11 differ with respect to those reported in Figure 2. However, such a difference is small if standard errors are considered; suggesting the presence of a (very) small negative bias in the standard two-way fixed effects event study approach.

Conclusion

This paper studies the effect of deep trade agreements on the export performance of firms in 31 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. Large and GVC firms are more positively affected by deep RTAs. For firms belonging to the top-quartile of size distribution, moving from *shallow* to *deep* RTAs implies 4.8% increase in their exports. For GVC firms importing and exporting from/to the same country, moving from *shallow* to *deep* RTAs implies a 9.6% increase in their exports. Conversely, small and less productive firms are negatively affected by deep RTAs as they suffer the higher degree of competition induced by deep RTA at destination. These results are robust to a number of extensions and robustness checks that confirm the causal impact of deep RTAs on firms' exports. These findings have relevant welfare and policy implications for developing countries. While the selection 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 process in developing countries can be significant.

[29] See De Chaisemartin and D'Haultfoeuille (2020) for more details on the robust two-way fixed effects estimator.
[30] In Table A11 we show the same estimations including firm's tariffs as controls and results hold.

Bibliography

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

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

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

Baier, S. L., Y. V. Yotov, and T. Zylkin (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 D. Jaimovich (2012). Are Free Trade Agreements contagious? *Journal of International Economics 88*(1), 1–16.

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

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

Chaney, T. (2008). Distorted gravity: the intensive and extensive margins of international trade. *American Economic Review 98*(4), 1707–1721.

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

Crowley, M., L. Han, and T. Prayer (2022). The pro-competitive effects of trade agreements. Working Paper 2240, University of Cambridge, Cambridge.

Dai, M., Y. V. Yotov, and T. Zylkin (2014). On the trade-diversion effects of free trade agreements. *Economics Letters 122*(2), 321–325.

De Chaisemartin, C. and X. D'Haultfoeuille (2020). Difference-in-differences estimators of intertemporal treatment effects. *arXiv preprint arXiv:2007.04267*.

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

Fajgelbaum, P. D., P. K. Goldberg, P. J. Kennedy, and A. K. Khandelwal (2019, 11). The Return to Protectionism. *The Quarterly Journal of Economics 135*(1), 1–55.

Fajgelbaum, P. D., P. K. Goldberg, P. J. Kennedy, and A. K. Khandelwal (2020, February).

The return to protectionism. *The Quarterly Journal of Economics 135*(1), 1–55.

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

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

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

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

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

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

Goldberg, P. K., A. K. Khandelwal, N. Pavcnik, and P. Topalova (2010, 11). Imported Intermediate Inputs and Domestic Product Growth: Evidence from India*. *The Quarterly Journal of Economics 125*(4), 1727–1767.

Handley, K. and N. Limão (2017, September). Policy uncertainty, trade, and welfare: Theory and evidence for china and the united states. *American Economic Review 107*(9), 2731–83.

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

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

Horn, H., P. C. Mavroidis, and A. Sapir (2010, November). Beyond the WTO? An Anatomy of

EU and US Preferential Trade Agreements. *The World Economy 33*(11), 1565–1588.

Lee, W., A. Mulabdic, and M. Ruta (2019, November). Third-Country Effects of Regional Trade Agreements: A Firm-Level Analysis. Policy Research Working Paper Series 9064, The World Bank.

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

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

Mattoo, A., A. Mulabdic, and M. Ruta (2017, September). Trade creation and trade diversion in deep agreements. Policy Research Working Paper Series 8206, The World Bank.

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

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

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

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

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

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

Spearot, A. C. (2013). Variable demand elasticities and tariff liberalization. *Journal of International Economics 89*(1), 26–41.

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

Tables and Figures

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

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

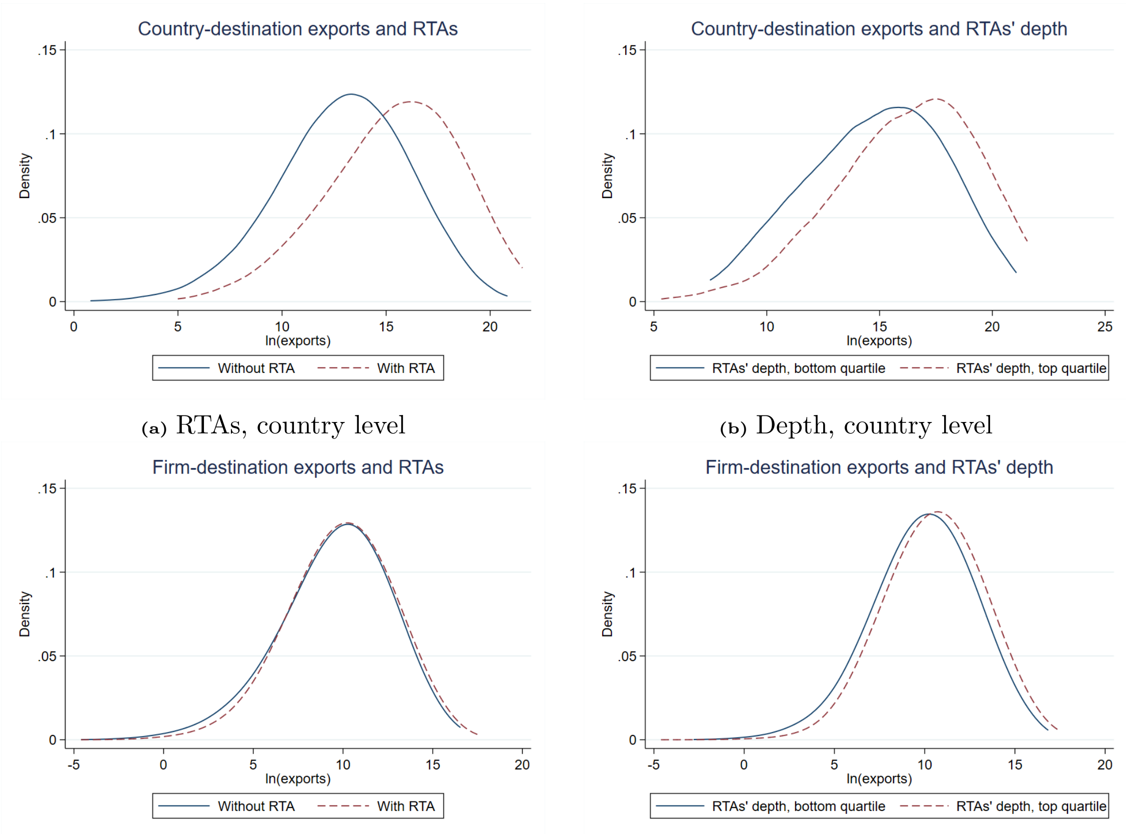
Notes: Authors' calculation on World Bank Content of Deep Trade Agreement data.

Table 2: In-sample descriptive statistics.

	Mean	Std Dev	Min	Max
Export	279032	1.1e+06	0	1.4e+06
RTA depth	14.4	13.8	0	48
RTA depth legally enf.	8.3	7.5	0	43
RTA depth WTO+	6.2	5.2	0	14
RTA depth WTOX	2.1	3.1	0	29
RTA depth core	7.4	6.3	0	18
ln(1+τ)	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.

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



(a) RTAs, country level

(b) Depth, country level

(c) RTAs, firm level

(d) Depth, firm level

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

Table 3: The trade effect of deep RTA. Baseline specification. PPML estimator.

	Exp	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)
RTA_{ijt}	0.047***					
	(0.015)					
DTA_{ijt}		0.003***				
		(0.001)				
DTA_{ijt} leg.			0.003***			
			(0.001)			
DTA_{ijt} WTO+				0.003*		
				(0.001)		
DTA_{ijt} WTO-X					0.007***	
					(0.002)	
DTA_{ijt} Core						0.002**
						(0.001)
$Ln(1+\tau_{ijt})$	-0.671***	-0.671***	-0.671***	-0.671***	-0.672***	-0.671***
	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)
Firm-Year FE	✔	✔	✔	✔	✔	✔
Destination-Year FE	✔	✔	✔	✔	✔	✔
Origin-Destination FE	✔	✔	✔	✔	✔	✔
Observations	2,388,213	2,388,213	2,388,213	2,388,213	2,388,213	2,388,213

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.

Table 4: The trade effect of deep RTA. 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}	0.006***				
	(0.001)				
Weigh. DTA_{ijt} leg.		0.005***			
		(0.002)			
Weigh. DTA_{ijt} WTO+			0.004		
			(0.004)		
Weigh. DTA_{ijt} WTO-X				0.009***	
				(0.002)	
Weigh. DTA_{ijt} core					0.004
					(0.003)
Ln(1+τ_{ijt})	-0.671***	-0.672***	-0.671***	-0.672***	-0.671***
	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)
Firm-Year FE	✔	✔	✔	✔	✔
Destination-Year FE	✔	✔	✔	✔	✔
Origin-Destination FE	✔	✔	✔	✔	✔
Observations	2,388,213	2,388,213	2,388,213	2,388,213	2,388,213

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.

Table 5: The heterogeneous trade effect of deep RTA by firm characteristics.

	Exp	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)
DTA_{ijt} leg.	-0.042***	-0.022***	-0.006***	-0.005***	-0.011***	-0.001
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
DTA_{ijt} leg. × ($k_f > 75^{th}$)	0.045***					
	(0.002)					
DTA_{ijt} leg. × ($k_f > 90^{th}$)		0.026***				
		(0.002)				
DTA_{ijt} leg. × ($k_f > 75^{th}$ t0)			0.010***			
			(0.001)			
DTA_{ijt} leg. × ($k_f > 90^{th}$ t0)				0.011***		
				(0.001)		
DTA_{ijt} leg. × GVC					0.016***	
					(0.002)	
DTA_{ijt} leg. × GVC bil.						0.008***
						(0.001)
GVC bil.						0.141***
						(0.014)
Ln(1+τ_{ijt})	-0.670***	-0.667***	-0.670***	-0.667***	-0.665***	-0.662***
	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)
Firm-Year FE	✔	✔	✔	✔	✔	✔
Destination-Year FE	✔	✔	✔	✔	✔	✔
Origin-Destination FE	✔	✔	✔	✔	✔	✔
Observations	2,388,213	2,388,213	2,388,213	2,388,213	2,388,213	2,388,213

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.

Table 6: The *pro-competitive* (selection) effect of deep RTA.

	# Survivors	# Survivors	Avg Exp	Avg Exp
	(1)	(2)	(3)	(4)
DTA$_{ijt}$ leg.	-0.006***	-0.005***	0.003**	0.006***
	(0.001)	(0.001)	(0.002)	(0.002)
Ln(1+τ_{ijt})		0.518***		-0.165*
		(0.092)		(0.098)
Origin-Year FE	✔	✔	✔	✔
Destination-Year FE	✔	✔	✔	✔
Origin-Destination FE	✔	✔	✔	✔
Observations	38,045	25,308	49,772	30,212

Notes: 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.

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

	$\hat{\beta}$	V_i/V	Aggregate Response	Aggregate Response (% of total)
Intensive	0.045***	0.336	0.015	40
Extensive	0.034***	0.664	0.023	60
Total			0.038	

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 $\hat{\beta} \times V_i/V$.

Table 8: The trade effect of deep RTA. The role of firm size distribution.

	Extensive	Extensive	Intensive	Intensive	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)
DTA$_{ijt}$ leg.	-0.008***	-0.006***	0.009***	0.008***	-0.005***	-0.004**
	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
DTA$_{ijt}$ leg. $\times$ Pareto shape^{a}	0.010***		-0.008***		0.007***	
	(0.002)		(0.002)		(0.002)	
DTA$_{ijt}$ leg. $\times$ Pareto shape^{b}		0.004*		-0.005***		0.003*
		(0.002)		(0.002)		(0.002)
Ln(1+τ_{ijt})	-0.021	-0.020	0.185	0.186	-0.005	-0.005
	(0.115)	(0.115)	(0.130)	(0.130)	(0.105)	(0.105)
Destination-Year FE	✔	✔	✔	✔	✔	✔
Origin-Year FE	✔	✔	✔	✔	✔	✔
Origin-Destination FE	✔	✔	✔	✔	✔	✔
Observations	30,182	30,182	12,537	12,537	30,212	30,212

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)

Table 9: The trade effect of deep RTA. 2SLS approach.

	Exp	Exp	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DTA_{ijt} leg.	0.007*** (0.003)	-0.042*** (0.007)	-0.024*** (0.005)	-0.019*** (0.004)	-0.008** (0.004)	-0.004 (0.004)	0.001 (0.003)
DTA_{ijt} leg. × ($\text{k}_f > 75^{th}$)		0.054*** (0.007)					
DTA_{ijt} leg. × ($\text{k}_f > 90^{th}$)			0.041*** (0.006)				
DTA_{ijt} leg. × ($\text{k}_f > 75^{th}$ t0)				0.037*** (0.005)			
DTA_{ijt} leg. × ($\text{k}_f > 90^{th}$ t0)					0.030*** (0.004)		
DTA_{ijt} leg. × GVC						0.016*** (0.004)	
DTA_{ijt} leg. × GVC bil.							0.015*** (0.002)
GVC bil.							-0.008 (0.023)
$\text{Ln}(1+\tau_{ijt})$	-0.575*** (0.040)	-0.572*** (0.040)	-0.567*** (0.040)	-0.574*** (0.040)	-0.570*** (0.040)	-0.574*** (0.040)	-0.577*** (0.040)
Firm-Year FE	✔	✔	✔	✔	✔	✔	✔
Destination-Year FE	✔	✔	✔	✔	✔	✔	✔
Origin-Destination FE	✔	✔	✔	✔	✔	✔	✔
Observations	2,280,704	2,280,704	2,280,704	2,280,704	2,280,704	2,280,704	2,280,704
IV DTA_{ijt}	0.064***	0.065***	0.065***	0.065***	0.065***	0.066***	0.064***
IV $\text{DTA}_{ijt} \times I(k_f > \bar{k})$		0.108***	0.111***	0.109***	0.110***	0.114***	0.085***
Joint F-stat	318	159	159	159	159	163	161

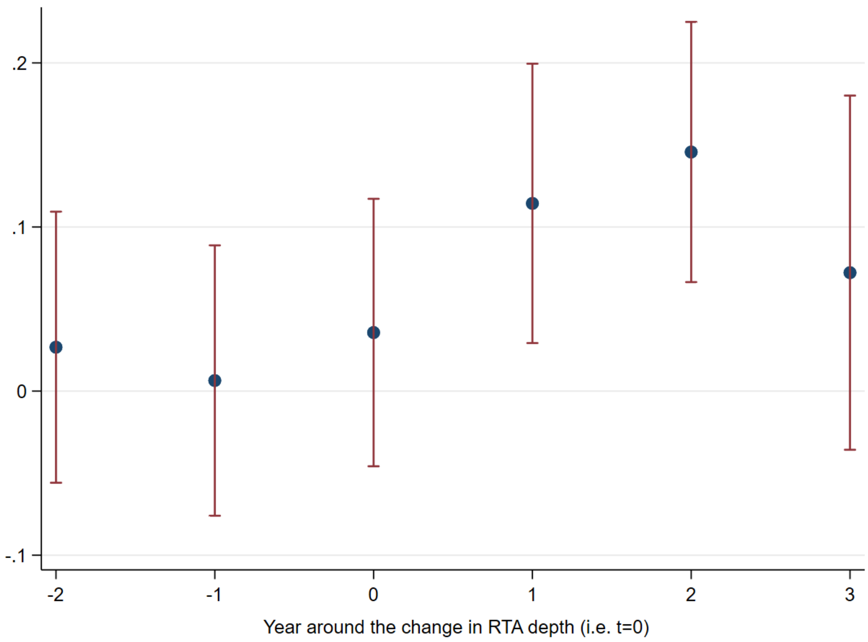
Notes: 2SLS estimates, origin-destination-year cluster standard errors in parentheses. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

Table 10: Deep trade agreements and the export of firms with plausibly exogenous instrument.

	Union of Confidence Interval estimations			
Dep Var	ν	Coeff. in tab 9	Min 90% CI	Max 90% CI
	Panel (a): Top-50 destinations.			
Firm Exports	0.0006 (0.0004)	0.008*** (0.003)	0.028	0.042
	Panel (b): market share above 0.5%.			
Firm Exports	0.0006 (0.0004)	0.008*** (0.003)	0.027	0.041

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).

Figure 2: Firm Exports Event Study.



Note: Figure plots event time dummies for targeted firms relative to untargeted firms. Regression includes firm-year, destination-year and origin-destination fixed effects. Standard errors are clustered by origin-destination-year. Error bars show 90% confidence intervals.

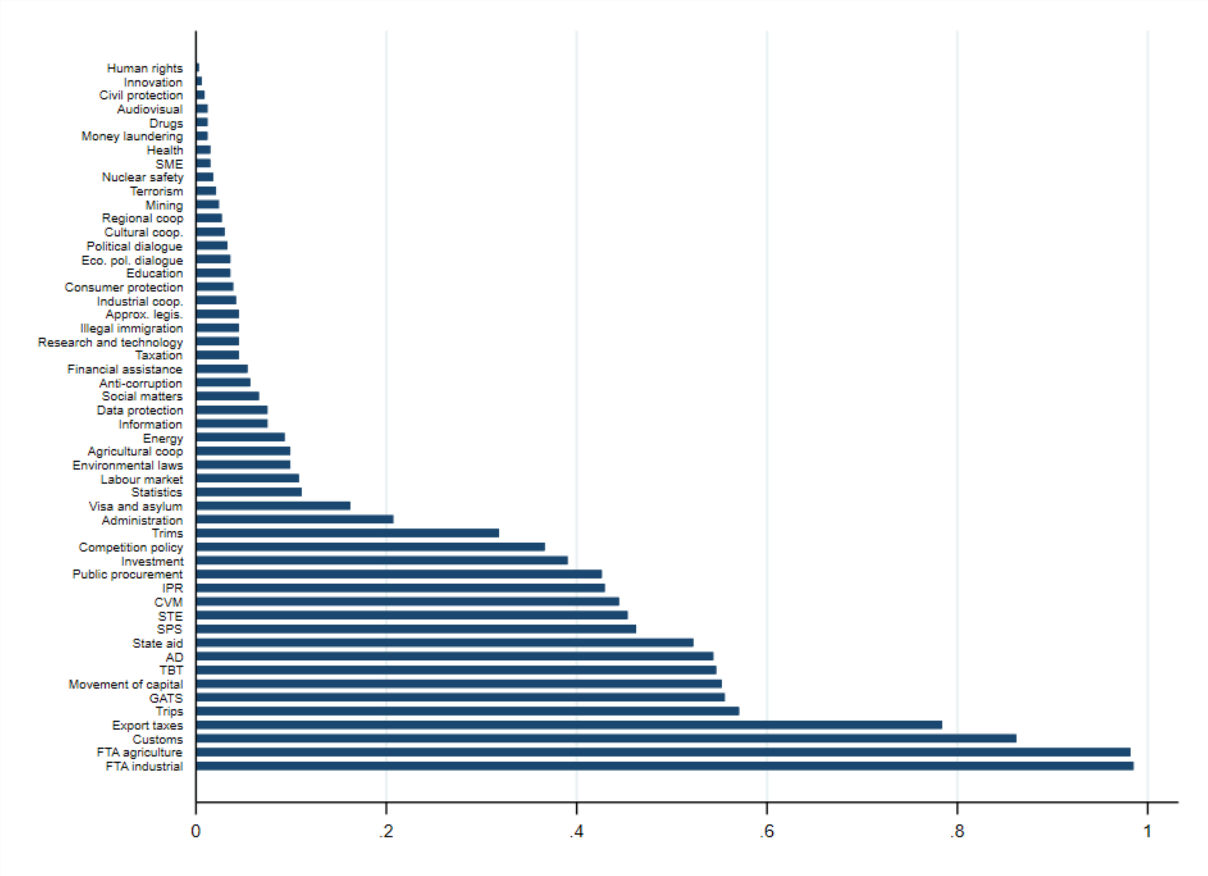
Table 11: Robustness check using the estimator by Chaisemartin D'Haultfoeuille (2020).

	Bilateral exports (ln)			
	t	*t+1*	*t+2*	*t+3*
DTA_{ijt}	0.094**	0.170**	0.270***	0.154
	(0.045)	(0.069)	(0.082)	(0.107)
Observations	1.568.450	1.440.873	1.279.444	1.100.902
# of switchers	174.852	167.748	159.379	141.907

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 first depth variation in the country-pair. Bootstrapped (100) standard errors clustered at the country-pair-time level are in parentheses. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

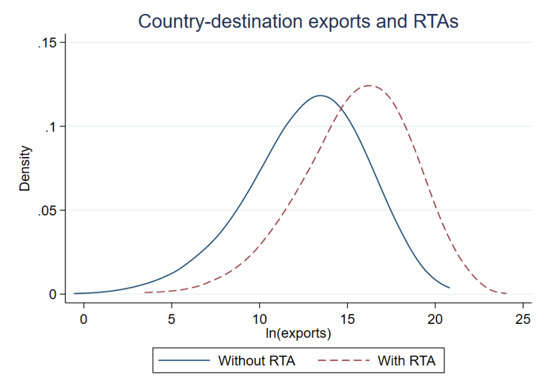
Appendix tables

Figure A1: Frequency of legally enforceable provisions in RTAs.

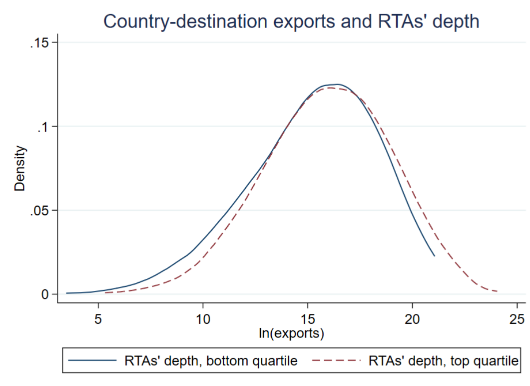


Note: Provisions are ranked from the less frequent to the most frequent.

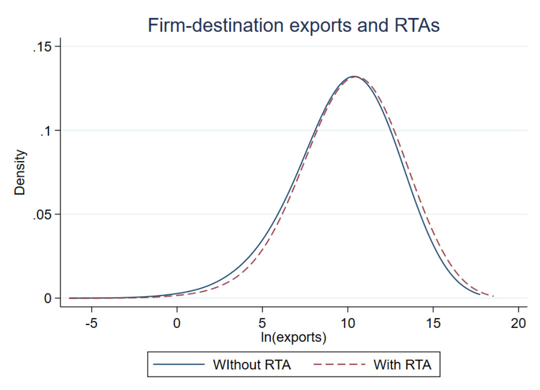
Figure A2: Exports value and RTAs, all available countries.



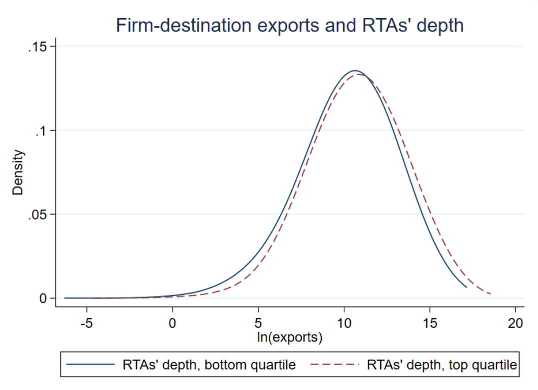
(a) RTAs, country level



(b) Depth, country level



(c) RTAs, firm level



(d) Depth, firm 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
ECU	2002-2019	ROU	2005-2011
EGY	2005-2016	RWA	2005-2016
ETH	2008-2017	SEN	2000-2020
GAB	2009-2015	SLV	2006-2020
GHA	2010-2019	SRB	2006-2019
GIN	2009-2012	STP	2014
GTM	2005-2013	SWZ	2012
HRV	2007-2015	TLS	2006-2012
IRN	2006-2010	TZA	2003-2017
MEX	2011-2016	UGA	2000-2010
JOR	2003-2012	URY	2001-2020
KEN	2006-2020	YEM	2008-2012
KGZ	2006-2012	ZAF	2009-2020
KHM	2016-2019	ALB	2007-2019
KWT	2009-2010	BDI	2010-2016
LBN	2008-2012	BEN	2016-2020
MDG	2007-2012	BFA	2005-2012
MKD	2008-2017	BGD	2005-2016
MLI	2005-2008	BGR	2001-2006
MMR	2011-2013	BOL	2006-2012
MUS	2010-2020	BWA	2003-2013
MWI	2005-2020	CHL	2000-2020
NER	2008-2010	CIV	2009-2019
NIC	2012-2014	COL	2000-2020
NPL	2011-2014	CPV	2010-2020
PAK	2015-2017	DOM	2006-2020
PER	2000-2020	GEO	2000-2020
PRY	2012-2020		

Notes: World Bank Export Dynamics Database.

Table A2: 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	-0.000
	(0.001)	(0.001)	(0.002)
Origin FE	✔		✔
Destination FE		✔	✔
Observations	216,877	216,875	216,874

Notes: OLS estimates, robust standard errors in parentheses. ***, ** and * indicate statistical significance levels for p-val. $<$ 0.01, p-val. $<$ 0.05, and p-val.$<$ 0.1.

Table A3: The effect of Deep Trade Agreements. Robustness check using Linear OLS estimations.

	Exp	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)
RTA_{ijt}	0.054***					
	(0.018)					
DTA_{ijt}		0.002***				
		(0.001)				
DTA_{ijt} leg.			0.005***			
			(0.001)			
DTA_{ijt} WTO+				0.005**		
				(0.002)		
DTA_{ijt} WTO-X					0.014***	
					(0.002)	
DTA_{ijt} Core						0.004**
						(0.002)
$Ln(1+\tau_{ijt})$	-0.560***	-0.560***	-0.560***	-0.561***	-0.561***	-0.561***
	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)
Firm-Year FE	✔	✔	✔	✔	✔	✔
Destination-Year FE	✔	✔	✔	✔	✔	✔
Origin-Destination FE	✔	✔	✔	✔	✔	✔
Observations	2,388,213	2,388,213	2,388,213	2,388,213	2,388,213	2,388,213

Notes: OLS 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.

Table A4: The trade effect of deep RTA. Cross-section identification.

	Exp	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)
RTA_{ijt}	0.266***					
	(0.019)					
DTA_{ijt}		0.006***				
		(0.001)				
DTA_{ijt} leg.			0.022***			
			(0.001)			
DTA_{ijt} WTO+				0.032***		
				(0.002)		
DTA_{ijt} WTO-X					0.038***	
					(0.004)	
DTA_{ijt} Core						0.025***
						(0.002)
Distance (ln)	-0.368***	-0.406***	-0.352***	-0.351***	-0.390***	-0.355***
	(0.017)	(0.016)	(0.017)	(0.016)	(0.017)	(0.016)
Contiguity	0.392***	0.380***	0.381***	0.382***	0.380***	0.391***
	(0.027)	(0.027)	(0.024)	(0.023)	(0.026)	(0.024)
Language	0.168***	0.209***	0.214***	0.207***	0.222***	0.209***
	(0.019)	(0.021)	(0.019)	(0.019)	(0.021)	(0.019)
Colony	0.174***	0.199***	0.177***	0.170***	0.168***	0.184***
	(0.047)	(0.045)	(0.044)	(0.044)	(0.043)	(0.044)
Firm-Year FE	✔	✔	✔	✔	✔	✔
Destination-Year FE	✔	✔	✔	✔	✔	✔
Origin-Destination FE	No	No	No	No	No	No
Observations	3,488,011	3,488,011	3,488,011	3,488,011	3,488,011	3,488,011

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.

Table A5: The trade effect of deep RTA. Cross-section identification controlling for applied tariffs.

	Exp	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)
RTA_{ijt}	0.164*** (0.017)					
DTA_{ijt}		0.003*** (0.001)				
DTA_{ijt} leg.			0.015*** (0.001)			
DTA_{ijt} WTO+				0.022*** (0.002)		
DTA_{ijt} WTO-X					0.022*** (0.003)	
DTA_{ijt} Core						0.017*** (0.001)
$\text{Ln}(1+\tau_{ijt})$	-0.613*** (0.048)	-0.661*** (0.048)	-0.625*** (0.048)	-0.612*** (0.048)	-0.673*** (0.049)	-0.617*** (0.049)
Distance (ln)	-0.425*** (0.012)	-0.458*** (0.012)	-0.403*** (0.012)	-0.402*** (0.012)	-0.439*** (0.012)	-0.407*** (0.012)
Contiguity	0.287*** (0.024)	0.268*** (0.024)	0.289*** (0.023)	0.292*** (0.022)	0.274*** (0.024)	0.291*** (0.023)
Language	0.106*** (0.015)	0.124*** (0.015)	0.140*** (0.015)	0.135*** (0.015)	0.136*** (0.015)	0.136*** (0.015)
Colony	0.186*** (0.049)	0.199*** (0.048)	0.187*** (0.047)	0.184*** (0.048)	0.183*** (0.047)	0.192*** (0.047)
Firm-Year FE	✔	✔	✔	✔	✔	✔
Destination-Year FE	✔	✔	✔	✔	✔	✔
Origin-Destination FE	No	No	No	No	No	No
Observations	2,388,510	2,388,510	2,388,510	2,388,510	2,388,510	2,388,510

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.

Table A6: The trade effect of deep RTA. Baseline specification controlling for the presence of a RTA.

	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)
DTA_{ijt}	0.005***		0.005**	
	(0.001)		(0.002)	
DTA_{ijt} leg.		0.004**		0.000
		(0.002)		(0.002)
$Ln(1+\tau_{ijt})$	-0.715***	-0.714***	-0.662***	-0.662***
	(0.118)	(0.118)	(0.119)	(0.119)
RTA adjusted	-0.042	0.022	-0.054	0.062
	(0.039)	(0.034)	(0.059)	(0.038)
Depth non-coded RTA	=1	=1	Closest	Closest
Firm-Year FE	✔	✔	✔	✔
Destination-Year FE	✔	✔	✔	✔
Origin-Destination FE	✔	✔	✔	✔
Observations	2,670,797	2,670,797	2,447,547	2,447,547

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.

Table A7: The trade effect of deep RTA. Robustness check using sub-sample of country pairs that switch RTA status during the period.

	Exp	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)
RTA_{ijt}	0.059***					
	(0.016)					
DTA_{ijt}		0.002***				
		(0.001)				
DTA_{ijt} leg.			0.003***			
			(0.001)			
DTA_{ijt} WTO+				0.003**		
				(0.001)		
DTA_{ijt} WTO-X					0.006***	
					(0.002)	
DTA_{ijt} Core						0.003**
						(0.001)
Firm-Year FE	✔	✔	✔	✔	✔	✔
Destination-Year FE	✔	✔	✔	✔	✔	✔
Origin-Destination FE	✔	✔	✔	✔	✔	✔
Observations	2,718,720	2,718,720	2,718,720	2,718,720	2,718,720	2,718,720

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.

Table A8: The heterogeneous trade effect of deep RTA by firm characteristics. Binned model.

	Exp	Exp
	(1)	(2)
DTA_{ijt} leg. × Big	0.003***	0.003***
	(0.001)	(0.001)
DTA_{ijt} leg. × Medium	-0.045***	-0.040***
	(0.002)	(0.002)
DTA_{ijt} leg. × Small	-0.150***	-0.156***
	(0.007)	(0.010)
$Ln(1+\tau_{ijt})$		-0.670***
		(0.045)
Firm-Year FE	✔	✔
Destination-Year FE	✔	✔
Origin-Destination FE	✔	✔
Observations	3,898,746	2,388,213

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.

Table A9: The heterogeneous effect of deep RTA. Robustness check using the weighted sum of provisions as a proxy for the RTA depth.

	Exp	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)
Weig. DTA_{ijt} leg.	-0.077***	-0.037***	-0.009***	-0.009***	-0.022***	-0.002
	(0.006)	(0.005)	(0.003)	(0.003)	(0.004)	(0.002)
Weig. DTA_{ijt} leg. × ($k_f > 75^{th}$)	0.082***					
	(0.006)					
Weig. DTA_{ijt} leg. × ($k_f > 90^{th}$)		0.044***				
		(0.004)				
Weig. DTA_{ijt} leg. × ($k_f > 75^{th}$ t0)			0.016***			
			(0.003)			
Weig. DTA_{ijt} leg. × ($k_f > 90^{th}$ t0)				0.018***		
				(0.002)		
Weig. DTA_{ijt} leg. × GVC					0.031***	
					(0.003)	
Weig. DTA_{ijt} leg. × GVC bil.						0.013***
						(0.002)
GVC bil.						0.166***
						(0.013)
$Ln(1+\tau_{ijt})$	-0.671***	-0.669***	-0.671***	-0.668***	-0.666***	-0.663***
	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)
Firm-Year FE	✔	✔	✔	✔	✔	✔
Destination-Year FE	✔	✔	✔	✔	✔	✔
Origin-Destination FE	✔	✔	✔	✔	✔	✔
Observations	2,388,213	2,388,213	2,388,213	2,388,213	2,388,213	2,388,213

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.

Table A10: The trade effect of deep RTA. Two-stage OLS/PPML estimations.

	Exp	Exp	Exp	Exp	Exp	Exp	Exp
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DTA_{ijt} leg.	0.010***	-0.075***	-0.040***	-0.011***	-0.006	-0.007	0.007
	(0.005)	(0.010)	(0.007)	(0.005)	(0.005)	(0.006)	(0.006)
DTA_{ijt} leg. $\times$ ($k_f > 75^{th}$)		0.077***					
		(0.008)					
DTA_{ijt} leg. $\times$ ($k_f > 90^{th}$)			0.044***				
			(0.005)				
DTA_{ijt} leg. $\times$ ($k_f > 75^{th}$ t0)				0.021***			
				(0.003)			
DTA_{ijt} leg. $\times$ ($k_f > 90^{th}$ t0)					0.018***		
					(0.002)		
DTA_{ijt} leg. $\times$ GVC						0.022***	
						(0.004)	
DTA_{ijt} leg. $\times$ GVC bil.							0.005**
							(0.002)
GVC bil.							0.209
							(0.014)
$\text{Ln}(1+\tau_{ijt})$	-0.665***	-0.665***	-0.665***	-0.666***	-0.667***	-0.664***	-0.660***
	(0.048)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.051)
Firm-Year FE	✔	✔	✔	✔	✔	✔	✔
Destination-Year FE	✔	✔	✔	✔	✔	✔	✔
Origin-Destination FE	✔	✔	✔	✔	✔	✔	✔
Observations	2,280,704	2,280,704	2,280,704	2,280,704	2,280,704	2,280,704	2,280,704
IV DTA_{ijt}	0.064***	0.065***	0.065***	0.065***	0.065***	0.066***	0.064***
IV $\text{DTA}_{ijt} \times I(k_f > \bar{k})$		0.108***	0.111***	0.109***	0.110***	0.114***	0.085***
Joint F-stat	318	159	159	159	159	163	161

Notes: two stage OLS/PPML estimates, bootstrapped standard errors in parentheses. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.

Table A11: Robustness check using the estimator by Chaisemartin D'Haultfoeuille (2020) controlling for tariffs.

	Bilateral exports (ln)			
	t	*t+1*	*t+2*	*t+3*
DTA_{ijt}	0.090*	0.251**	0.162	-0.064
	(0.053)	(0.109)	(0.141)	(0.161)
Observations	1.003.047	913.310	803.168	684.178
# of switchers	88.713	82.109	77.587	61.992

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 first depth variation in the country-pair. Bootstrapped (100) standard errors clustered at the country-pair-time level are in parentheses. ***, ** and * indicate statistical significance levels for p-val. < 0.01, p-val. < 0.05, and p-val.< 0.1.