

NBER WORKING PAPER SERIES

LEARNING TO USE TRADE AGREEMENTS

Kala Krishna
Carlos Salamanca
Yuta Suzuki
Christian Volpe Martincus

Working Paper 29319
<http://www.nber.org/papers/w29319>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2021, Revised November 2022

We thank Julia Cajal Grossi, Mathieu Parenti, Feodora Teti, Halis Yildiz and Jonas Kasteng, Nils Norell, Patrik Gustavsson Tingvall at the National Board of Trade, Sweden, for valuable discussions. We also thank participants of the NBER ITI Program Meeting Fall 2021, the Geneva Trade and Development Workshop, the Australasian Trade Workshop, the CEA 2021, the EEA-ESEM Virtual 2021, and the seminar at the University of Niigata Prefecture for helpful comments. We are also grateful to Rafael Cornejo for providing valuable and insightful information on preference utilization in our sample countries and Antoni Estevadeordal and Jeremy Harris for sharing their ROO restrictiveness indicators. Nino Doghonadze provided great research assistance. The views and interpretations in this paper are strictly those of the authors and should not be attributed to the Inter-American Development Bank, its executive directors, its member countries, or the National Bureau of Economic Research. Other usual disclaimers also apply.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2021 by Kala Krishna, Carlos Salamanca, Yuta Suzuki, and Christian Volpe Martincus. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Learning to Use Trade Agreements
Kala Krishna, Carlos Salamanca, Yuta Suzuki, and Christian Volpe Martincus
NBER Working Paper No. 29319
October 2021, Revised November 2022
JEL No. F02,F13,F14,F68,N76

ABSTRACT

Preferential trade areas (PTAs) allow firms to pay zero or preferential tariffs as long as Rules of Origin (ROOs) are met. Meeting them is costly for exporters not only in terms of production costs but also in terms of fixed costs, such as documentation costs. We ask if these fixed costs change with the experience of exporters in obtaining preferential tariffs. We explore this using a unique exporter-importer matched transaction-level customs data set on a group of Latin American countries. We estimate a model-based equation and show that these fixed costs depend on the history of preference utilization.

Kala Krishna
Department of Economics
523 Kern Graduate Building
The Pennsylvania State University
University Park, PA 16802
and NBER
kmk4@psu.edu

Carlos Salamanca
Pennsylvania State University
carlos.salamanca91@gmail.com

Yuta Suzuki
The Pennsylvania State University
hit.suzuki.y@gmail.com

Christian Volpe Martincus
IADB
1300 New York Ave.
Washington, DC 20577
christianv@iadb.org

1 Introduction

Recent decades have seen a proliferation of free trade agreements (FTAs).¹ Rules of origin (ROOs) are used to distinguish between products which are eligible for preferential treatment (those originating in the member countries) and those from third countries. Meeting these rules of origin is costly as evidenced by preference utilization rates that are often far below unity.² Certainly, marginal costs could rise as firms change their production processes so as to meet ROOs. In addition, fixed costs would rise as supply chains are altered and/or documentation costs are incurred to show that the shipment complies with ROOs.³

Do these fixed costs change with the experience of the exporting firm in obtaining preferential tariffs?⁴ What type of experience, if any, matters? In this paper, we use the universe of export transactions from Argentina and Peru to Colombia over a long time period to answer this question. We make the case that the probability of using preferences decreases in the fixed costs of using preferences. Fixed costs of using preferences, in turn, should fall with experience if there is learning. In this case, greater experience of the exporter should increase the probability of using preferences.⁵ We accordingly infer the shape of fixed costs of meeting ROOs on the basis of how an exporting firm's history in using preferences affects

¹From 1990 to 2019 they rose from 76 to 443 in number. See Figure 1A in [Dinh et al. \(2019\)](#).

²See for example [UNCTAD \(2018\)](#), which shows that utilization rates of EU exporters to countries with FTAs with the EU are only 67% between 2009 to 2013. For an overview of ROOs, see [Cornejo and Harris \(2007\)](#).

³For instance, firms may need to keep records that would not otherwise have been kept and learn to fill out the documentation required to show ROOs have been met. [Cadot et al. \(2006b\)](#) estimate that the administrative costs of the Pan-European preference scheme are around 6.8% of the value of trade compared to 1.9% for NAFTA. For details on the forms taken by these ROOs and the procedures involved see [Dinh et al. \(2019\)](#).

⁴We use experience by the exporter, not the importer as this is what is relevant as the burden of providing the documentation is on the exporter. As a result we do not expect the importer's experience to matter.

⁵We do not differentiate between fixed costs and sunk costs of using preferences. If sunk costs are incurred only when preferences are used for the first time, we should observe an increase in the probability of using preferences only after the first experience. If the probability of using preferences keeps increasing as more experience is gained, as our estimates actually reveal, we argue that fixed costs exist and decrease with experience.

the likelihood of the firm using preferences in the current transaction.

Furthermore, we ask whether these costs are impacted differently by the form of the experience. Is experience in the same product and with the same importer (i.e., importing firm) more valuable than other kinds of experience? Does experience with specific importers and products spillover to other importers and products? For example, costs of meeting or documenting ROOs for a particular product could fall once a particular exporter has successfully overcome the hurdles imposed by ROOs. In this case, the experience of the exporter with one importer and product would increase the probability of using preferences in transactions with other importers of the same product. If some of this experience can be useful with other products, it may even spillover to other products with the same importer or even to other products and other importers. In other words, the nature of these costs can cast its shadow on the patterns in preference usage across suppliers and products. Consequently, the patterns in preference usage are informative about the nature of these costs.

The choice of using preferences is a tradeoff between the costs and benefits of using ROOs. One element of these benefits is the lower tariffs from using preferences, i.e., the difference in the MFN tariff and the preferential tariff times the value of the transaction, what we call “savings”. We estimate the probability of using preferences on a particular transaction with a particular importer as a function of the history of preference use while controlling for these savings.

However, savings can be endogenous.⁶ To account for this we construct a new instrument for the size of the transaction which is model-based: namely the daily exchange rate of the importing country. Using this instrument, we show that spillovers across products and /or importers are not always evident, though spillovers within the same product, even with different importers, are always present. In other words, experience in obtaining preferences in a particular product (defined as a HS 10 digit category) do not always seem to significantly improve

⁶There could be reverse causation present: not only might a larger transaction size drive the use of preferences, but the desire to use preferences may also drive transaction size.

the probability of obtaining preferences in another product, though experience in obtaining preference in a product with a given importer makes it much more likely preferences will be obtained for a transaction with the same or another importer in the same product. The results we find make sense: The ROOs are at the product level. Documenting that the supply chain complies with ROOs is costly. But doing so for one buyer extends to other buyers. This is less so across products especially for Peru.

Our estimates suggest that learning is larger for Argentina than for Peru. We argue that this due to the fact that the FTA between Argentina and Colombia is newer, so that there is a larger learning potential for its firms. We use a natural experiment to verify this intuition. In 2005 new products became covered for Argentina under the expanded FTA. One would expect that there would be more learning evident for these products which is exactly what we find. Moreover, once we distinguish between newly covered products and the previously covered ones, we see that the estimates for learning in the latter for Argentina look much like the learning estimates for Peru.

Understanding the nature of these fixed costs of ROOs is important from the theoretical, empirical, and policy perspectives. On the theory side, we provide a simple model on which we base our estimation approach. On the empirical side, we look for evidence of spillovers from experience across products and trading partners. On the policy front we argue that given that learning is important, policies which encourage preference use early on can have a huge impact in terms of the efficacy of a free trade area.

There is a large literature on ROOs, both on the theory and empirical side. The first set of papers highlight what ROOs are and how they operate, where the second set of papers estimate their effects on trade. [Krishna and Krueger \(1995\)](#) is an early paper that shows theoretically how ROOs can provide hidden protection to input suppliers within the FTA. Also see [Krishna \(2006\)](#) for a slightly dated survey of the literature. [Cadot et al. \(2006a\)](#) focuses on ROOs and has a number of case studies as well as innovations in terms of measuring the restrictiveness of

ROOs. [Anson et al. \(2005\)](#) argue that ROOs limit the use of preferential market access considerably. They estimate that in NAFTA, compliance costs are on average 6% in ad-valorem terms while administrative costs amount to 47% of the preference margin. [Pelkmans-Balaoing and Manchin \(2007\)](#) report that for the ASEAN FTA, preferential tariffs increase intra-regional imports only when preference margins are high (over 25 percentage points).⁷ [Demidova, Kee and Krishna \(2012\)](#) show that the patterns in the use of preferences among Bangladeshi exporters are consistent with firms facing both fixed and marginal costs of meeting ROOs. [Cherkashin et al. \(2015\)](#) set up and estimate a heterogeneous firm model with a view to evaluating the role of ROOs in Bangladeshi exports in Apparel. They find large predicted effects on Bangladeshi exports of reducing fixed and/or marginal costs of meeting ROOs in apparel.

These papers present evidence consistent with high fixed and/or marginal costs of using preferences. To our knowledge, ours is the first paper to use a model-based approach and address concerns about endogeneity of transaction size to show that these fixed costs seem to be differentially affected by different kinds of experience.⁸

We proceed as follows. Section 2 contextualizes rules of origin and explain their associated costs. In Section 3 we first describe the institutional background for the Latin American countries in the data set. We then describe the data used, present some summary statistics, and shows the data patterns that motivate our estimation strategy. Section 4 lays out a simple model of the choice of using preferences that guides our estimating equation. In Section 5, after accounting for possible endogeneity as well as possible measurement error of the saving variable, we show that learning is indeed increasing with experience and with the kind of experience. Section 6 concludes.

⁷They define preference margin as MFN tariff minus the preferential tariff divided by the MFN tariff. We use the term just for the difference.

⁸Concurrent, interesting work by [Benguria \(2022\)](#) and [Kasteng, Norell and Tingvall \(2022\)](#) are not model based, do not differentiate between types of experience, and do not fully account for endogeneity issues.

2 Costs of Rules of Origin

In this paper, we model the effects of ROOs as an increase in both marginal and fixed costs of meeting them. It makes sense that ROOs will raise marginal costs of production: forcing a firm to produce or source in a way it would ordinarily not do so as to get preferences must raise marginal costs.⁹ Documenting that ROOs have been met as well as the costs of changing input suppliers are examples of fixed costs of meeting ROOs. In Vietnam, for example, getting origin for the ASEAN Free Trade Area requires a form (form D) be filled out and the products be inspected. [Kirk \(2007\)](#) (page 12, box 1) reproduced here, outlines the steps needed.

“In Vietnam, the Export-Import Managing Department of the Ministry of Trade is the issuing institution for Form D. An application is submitted to an inspection company authorized by the Ministry of Science to conduct a cost screening to ensure local content of 40 percent or more. VINACONTROL remains the largest inspection firm, but the number of authorized companies has increased over the past few years. This provides for competition. Screening generally takes between one-half to a full day. The applications required for each shipment are submitted to a branch office of the Export-Import Managing Department (9 Branches nationwide) and are accompanied by a certifying letter from the inspection company, a commercial invoice, a customs declaration form, a bill of lading, and a copy of the exporter’s commercial license. Form D is issued within 2 hours.”

A quote from an automobile producer in Thailand gives more details about the

⁹For example, Bangladeshi apparel exports to the EU under the EBA (Everything But Arms) require that the cloth used come from Bangladesh which is more expensive than similar imported cloth as it helps confer origin, see [Cherkashin et al. \(2015\)](#). More recently, [Conconi et al. \(2018\)](#) look at NAFTA and show that the change in sourcing decisions so as to use NAFTA preferences led to increases in the marginal cost in production. Also see [Anson et al. \(2005\)](#) and [Head, Mayer and Melitz \(2021\)](#).

costs involved:¹⁰

“The preparation of documents for the initial cost screening takes two months and the screening procedures themselves about one month. There are 1,000 to 2,000 parts in a completed vehicle, and we must collect documentation (invoices, Form Ds, etc.) certifying local procurement from each supplier”

How hard is obtaining the certificate of origin in Latin America? For example, in Argentina, the certificate of origin must be issued by the designated responsible authorities (or delegated entities) according to a pre-established template. Specifically, it must include the name and the signature of the authorized official and the stamp of the certifying entity, a description of the good that perfectly matches those of the relevant tariff line code and the commercial bill, be complete, and be neither damaged nor amended.¹¹ The procedure for verifying such documentation are clearly laid out.¹² As is evident from this document, suspicious documentation may be investigated and investigation can be very expensive for exporters.

3 Institutional Background and Data Patterns

Since we focus on exports to Colombia from Argentina and Peru, we provide a short history of the relevant trade agreements between these countries. This is important because the behavior in terms of preference use comes from both changes

¹⁰Kirk (2007), page 13.

¹¹The commercial bill must be issued by the exporting firm in the origin country of the goods and a sworn declaration signed by the producer when this is also the exporter and by both, producer and exporter, when they are not the same. Firms must report a large amount of information including: (1) name of the producer (and exporter when they are not the same) and the firm’s legal representative; (2) address as registered with the tax agency; (3) description and tariff line code of the good to be exported; (4) FOB value; (5) information on the value and the tariff line code of each input according to whether it originates in (i) the exporting member country; (ii) other member countries; and (iii) the non-member countries; and (6) a description of the production process.

¹²See for example, article 20 of the document available at SICE website: http://www.sice.oas.org/trade/mrcsrac/Anexos/AnexoIV_s.asp

in these trade policies and learning, both individual and social, by exporters. The details of these treaties can be found in Online Appendix [D](#).

3.1 Institutional Background

There is a long history of preferential trade agreements in Latin America starting in the late 1950s that involved Argentina, Colombia, and Peru. These agreements were deepened over time. By 2000, when our data begins, the state of affairs was as follows.

Prior to 2000, the PTA with Argentina (Economic Complementary Agreement of Partial Scope 11 — AAP.CE 11 for its name in Spanish) was quite shallow and tariffs were reduced only for a limited number of products. In 2000, (under the AAP.CE 48) Colombia granted fixed preferences on around 1,250 products from Argentina (i.e., less than one quarter of the total number of tariff lines), with the preference margin rates (defined as the MFN tariff less the preferential one divided by the MFN tariff) averaging 40 percent. In 2005, (under AAP.CE 59) further cuts to tariffs were made and tariffs were reduced further on a group of products so that tariffs reached an average of 10 percent in 2005.

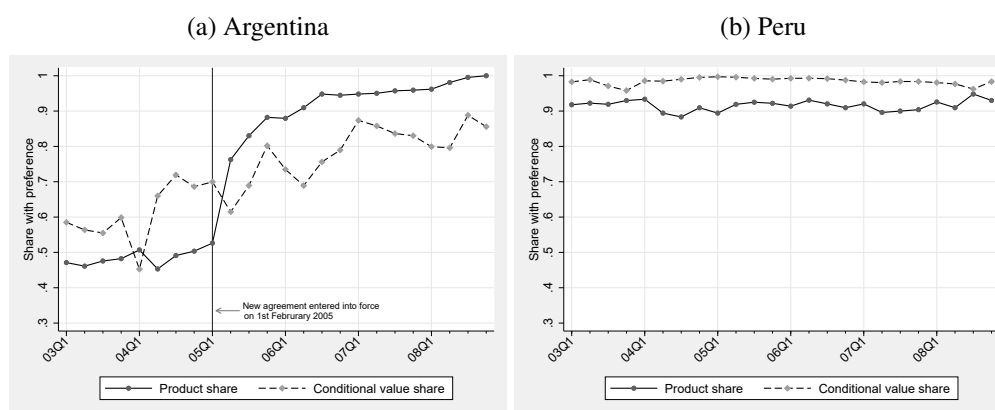
Thus, the trade agreements between Argentina and Colombia were shallow to begin with, but were deepened in 2005. The average preference margin (defined as the value share weighted average of the difference in the MFN and preferential tariff) was 4.7% for Argentina with a bump up after 2005. As a result, the share of products with preferences as well as the value share of transactions using preferences, conditional on the product having preferences, rose after 2005 as depicted in Figure 1.¹³ There is a sharp increase in the share of products with preferences in 2005Q2 and a continuing increase until 2006Q3 as products were phased in. Notice that the conditional value share initially falls, which is consistent with exporters being slow to begin using preferences. However, it soon starts to rise, which would be expected if firms were learning to use preferences.

¹³The figures cover the period 2003-2008 as 1998 to 2002 was a period of considerable turmoil in Argentina due to macro-economic instability which adds noise to this figure.

In contrast, the PTAs with Peru are long-standing and deep. Peru and Colombia had a long-standing history of preferential trade as they were both members of the Andean Pact since the late 1960s and its successor in 1996, the Andean Community. Though the aim was to create a customs union, i.e., bring tariffs down to zero and set a common external tariff for members, this goal was not achieved. Nevertheless, the median tariff imposed by Colombia on exports from Peru decreased from 46 percent in 1985, to 10 percent in 1995, and was close to zero by 2000. The average preference margin is 14.4% for Peru and does not change much over time. The share of products covered by preferences is above 90% and value share conditional on having preferences is even higher. These shares are also very stable over time.

These differences in the nature, depth and duration of the FTAs for Argentina and Peru with Colombia are why we analyze them separately. Note that both the higher average utilization of preferences pointed out above, as well as evidence of less learning for Peru relative to Argentina documented below makes sense in terms of these differences.

Figure 1: Preference Utilization Over Time



Note: Product share is the simple average of the share of products covered by preferences. Conditional value share is the simple average of the value share of transactions using preferences, conditional on the product having preferences.

3.2 Data

The data used here is part of a set of administrative data at the Inter American Development Bank (IDB). It consists of three main databases. First, we have highly disaggregated import data for Colombia from the National Tax and Customs Agency (Dirección de Impuestos y Aduanas Nacionales -DIAN). These data are reported at the transaction-level and cover all transactions entering Colombia over the period 2000-2011. Specifically, each record includes the importing firm's tax ID and name, the origin country, the product code (10-digit HS), the name of the foreign seller, the import value in US dollars, and the tariff actually paid.¹⁴ These data allows us to know time-specific, product-level MFN tariffs (inferred from tariffs on countries without preferential trade agreements with Colombia) and preferential margins applied when preferences are used (inferred from the difference in MFN tariffs and tariff paid).¹⁵

The second and third datasets consist of highly disaggregated export data for Argentina and Peru over 2000-2008 and 2000-2011, respectively, from their respective tax and customs agencies (Administración Federal de Ingresos Públicos—AFIP and Superintendencia Nacional de Administración Tributaria-SUNAT). In the export data, each record includes the exporting firm's tax ID and name, the destination country, the product code (10-digit HS), the export value in US dollars.¹⁶

Using the name of the selling firms reported both in the import database of Colombia and the export databases of Argentina and Peru, we are able to match buyers and sellers for each Colombian import transaction over our sample period and accurately track each exporter's history of preference use and the various kinds of experience by product and importer.¹⁷

¹⁴All transactions are denominated in dollars.

¹⁵We identify that preferences are used whenever the tariff paid is below the MFN one.

¹⁶The quantity (weight) in kilograms is available though we do not use this data.

¹⁷The merging of these different datasets is challenging. Details of the data cleaning exercise as well as an explanation of the standardization and matching procedures can be found in Appendix A.

Tables 1 and 2 give the summary statistics for the data for Argentina and Peru separately over the entire sample period, based on almost 30,000 transactions for Argentina and approximately 110,000 transactions for Peru.¹⁸ The first panel of each table gives the utilization rate by product. The median is 1.00 for exporters from Peru while for exporters from Argentina have a lower utilization at the median of .75. The median value of a transaction was also higher for Peruvian exporters at US dollars 7,866 versus 4,511 for Argentinian ones.

The second panel presents information on a number of variables at the exporter level. At this level as well, the utilization rate is lower in Argentina than in Peru. In Argentina, 20 percent of exporters do not ever use preferences, while in Peru, this is only so for the bottom 10 percent. More than 60 percent of Argentinian exporters sell a single product in Colombia, whereas more than 50 percent of their Peruvian counterparts do so. More than 70 percent of exporters from Argentina have a single importer and over 80 percent have a single importer per product. The numbers are similar for Peru. This is important because with exporter-importer fixed effects only data on firms with more than one transaction per partner will be useful in identifying the shape of the cost of preferences. The median number of transactions per exporter is 2 and transactions per product are 1 for both countries. At the 90th percentile of the relevant distributions, they reach 20 and 10 in the case of Argentina and 43 and 17 in the case of Peru, respectively. The value of transactions per exporter is higher for Peru at each decile, and this is so especially at the top deciles. The place where they differ more is in terms of the age of exporters: in 2008 Peru has a higher age at the top deciles than does Argentina reflecting the long standing nature of the FTA with Peru.

The third panel of both tables provide summary statistics for the same variables, but for the importer side. Importers are larger than exporters in most dimensions, but similar patterns hold for importers. Importers are much older than exporters in both countries: the median age is 5 in Argentina and 4 in Peru.

A pattern clearly evident in the data is that preference use becomes more likely

¹⁸Note that the numbers are integers as would be expected, except where we look at per product.

Table 1: Summary Statistics: Argentina

		Percentiles								
Variable		10	20	30	40	50	60	70	80	90
per Exporter	Utilization rate by product	0.25	0.41	0.50	0.64	0.75	0.90	1.00	1.00	1.00
	Value of transaction	446	870	1558	2743	4511	7255	11400	18973	35550
	Utilization rate	0.00	0.00	0.33	0.50	0.73	0.88	1.00	1.00	1.00
	#Products	1	1	1	1	1	1	2	2	4
	#Importers	1	1	1	1	1	1	1	2	3
	#Importers per product	1	1	1	1	1	1	1	1	2
	#Transactions	1	1	1	1	2	3	5	9	21
	#Transactions per product	1	1	1	1	1	2	3	4	10
	Value of transaction	627	1233	2121	3400	5040	7798	11466	17698	31755
	Age distribution in 2008	0	0	0	1	1	2	3	4	6
	Utilization rate	0.00	0.00	0.00	0.50	0.75	1.00	1.00	1.00	1.00
	#Products	1	1	1	1	1	1	2	3	5
per Importer	#Exporters	1	1	1	1	1	1	1	2	3
	#Exporters per product	1	1	1	1	1	1	1	1	2
	#Transactions	1	1	1	1	2	3	5	9	21
	#Transactions per product	1	1	1	1	1	2	2	4	7
	Value of transaction	625	1200	2000	3116	4600	6900	10315	16367	26880
	Age distribution in 2008	0	0	2	3	5	7	7	8	8
	#Transactions	1	1	1	1	2	2	3	6	12
	#Transactions per product	1	1	1	1	1	2	2	3	7
per Exporter-Importer	Utilization rate	0.00	0.00	0.00	0.50	0.75	1.00	1.00	1.00	1.00
	#Products	1	1	1	1	1	1	2	3	5
	#Exporters	1	1	1	1	1	1	1	2	3

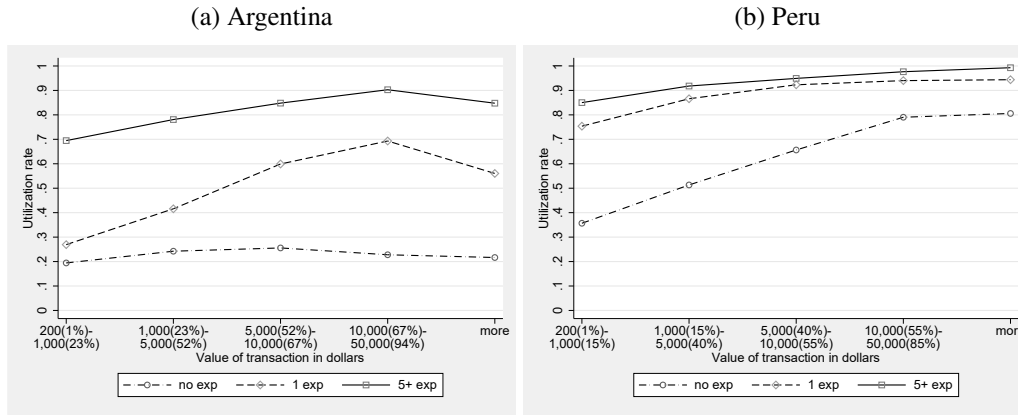
Note: Data for Argentina is from 2000Q4 to 2008Q4.

Table 2: Summary Statistics: Peru

		Percentiles								
Variable		10	20	30	40	50	60	70	80	90
per Exporter	Utilization rate by product	0.35	0.67	0.86	0.96	1.00	1.00	1.00	1.00	1.00
	Value of transaction	660	1465	2771	4715	7866	12751	21367	36875	65760
	Utilization rate	0.00	0.67	0.95	1.00	1.00	1.00	1.00	1.00	1.00
	#Products	1	1	1	1	1	2	2	4	6
	#Importers	1	1	1	1	1	1	1	2	4
	#Importers per product	1	1	1	1	1	1	1	1	2
	#Transactions	1	1	1	1	2	3	6	13	43
	#Transactions per product	1	1	1	1	1	2	3	6	17
	Value of transaction	700	1348	2400	3662	5580	8745	13791	22977	42028
	Age distribution in 2008	0	0	0	0	0	2	3	6	7
	Utilization rate	0.00	0.44	0.95	1.00	1.00	1.00	1.00	1.00	1.00
	#Products	1	1	1	1	1	2	3	4	7
per Importer	#Exporters	1	1	1	1	1	1	2	2	4
	#Exporters per product	1	1	1	1	1	1	1	1	2
	#Transactions	1	1	1	2	3	5	8	18	47
	#Transactions per product	1	1	1	1	2	2	4	7	16
	Value of transactions	797	1685	2941	4725	7501	11518	18056	28070	48270
	Age distribution in 2008	0	0	0	2	4	6	7	8	8
	#Transactions	1	1	1	2	2	3	5	9	24
	#Transactions per product	1	1	1	1	1	2	3	5	12
per Exporter-Importer	Utilization rate	0.00	0.00	0.00	0.50	0.75	1.00	1.00	1.00	1.00
	#Products	1	1	1	1	1	1	2	3	5
	#Exporters	1	1	1	1	1	1	1	2	3

Note: Data for Peru is from 2000Q4 to 2011Q4.

Figure 2: Preference Utilization by Value of the Transaction and Experience



Note: Preference utilization rate is the ratio of the number of transactions using preferences over the total number of transaction at every value bin of the transactions. *no exp* shows the utilization rates for firms with no experience in using preferences prior to the time of transaction, *1 exp* shows the utilization rates for firms with one experience, and *5+ exp* shows the utilization rates for firms with 5 or more experience.

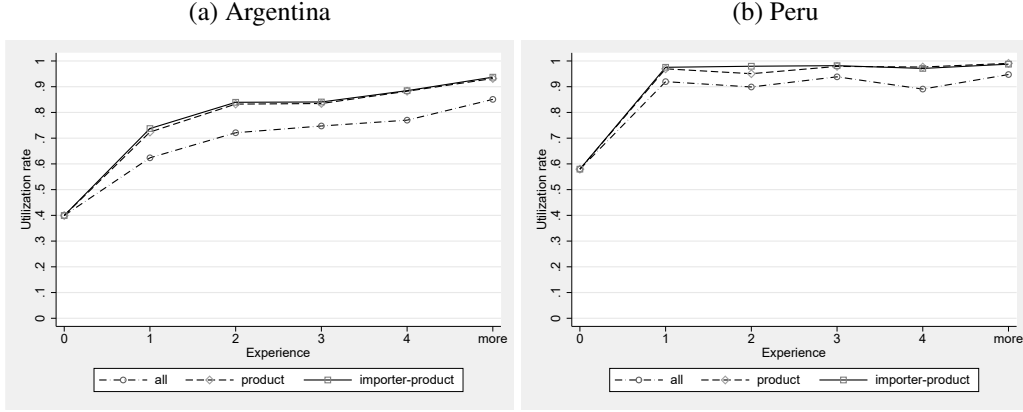
with experience and when the transaction value rises. This is depicted in Figure 2 for Argentina and Peru separately. The curves show preference utilization as a function of transaction size for exporters with different levels of experience. There is a distinct upward slope. It is clear that greater experience raises the likelihood of using preferences for all transaction value levels, conditional on there being preferences that could be used.¹⁹

Figure 3 looks at the preference use as a function of experience, but distinguishes between the kinds of experience. Note that preference use when total experience of the exporter to Colombia (across all products and importers) is considered is substantially lower than when experience in the same product, conditional on the exporter having no experience in using preferences in other products, is considered.²⁰ The third curve depicts preference use as a function of experience of the exporter with a given importer and product, conditional on having no ex-

¹⁹The transaction levels and the percentiles that those levels correspond to are depicted on the x-axis

²⁰Note that here experience of the exporter could come from many importers in the given product.

Figure 3: Preference Utilization by Type of Experience



Note: This figure shows the preference utilization rates over transactions at every experience bin. *all* refers to all kinds of experience in using preferences prior to the transaction, *product* refers to all experience in the product, while *importer-product* refers to all experience with the importer in the product.

perience of preference use in other importer product combinations. This curve is close to the one for a given product. This suggests that experience is likely to be product specific rather than importer product specific. However, to the extent that firms included in the importer-product definition of experience differ from those in the product definition, this may give a biased view of the importance of these two kinds of experience. These kinds of issues are exactly why we use a regression approach below.

4 Motivating the Empirical Specification

It seems reasonable, as we assume, to have the exporter choose whether to invoke preferences or not as it is the exporter who will need to prove that the products meet origin requirements. Hence, for each transaction in a product, we can think of a supplier who decides whether to use preferences or not depending on whether the costs of doing so, both fixed and marginal, exceed the benefits or not. On the one hand, the exporter gains from lower tariffs and this gain is larger the larger

is the preference margin. On the other hand, its costs may rise in order to meet Rules of Origin (ROOs). The increase in marginal costs comes about because the firm has to modify some aspect of its production to meet ROOs when it chooses to invoke preferences. It may also incur higher fixed costs such as the costs of changing its supply chain (i.e., finding new suppliers) and the administrative costs of associated with proving compliance of the ROOs. In practice, an exporter needs to have a certificate of origin to obtain preferences. This requires the exporter to provide documentation that shows their product qualifies for preferences as it meets the relevant rules of origin.²¹

These fixed costs can depend on experience in using preferences. Knowledge about input suppliers and providing documentation in past transactions may reduce fixed costs and make easier for an exporter use preferences. There may also be spillovers to other importers or other products. That is, experience in one product and importer may help the exporter to use preferences with other products and/or importers. Our central insight is that if there is no learning, then the probability of using preferences should be unaffected by any history the exporter has in using preferences. Marginal costs, in contrast, are likely not to depend on experience in using preferences, but are likely to vary considerably by product as the rules of origin are defined at a very detailed product level.

Formally, suppose that exporter e exports a particular product p and faces a constant elasticity of demand from importer i for transaction t .²² The demand curve is:

$$q_{eipt}(a_{eipt}) = \left((1 + \tau_{pt}^{\text{mfn}})^{1-a_{eipt}} (1 + \tau_{pt}^{\text{pref}})^{a_{eipt}} p_{eipt}(a_{eipt}) \right)^{-\eta} \psi_{eipt}, \quad (1)$$

²¹Implementation varies across settings. In some settings, like NAFTA, self-certification is enough. Approved exporter authorization may be required for the exporter to be allowed to issue origin declarations and incorrectness of the issued origin declaration can lead to withdrawal of the authorization and further consequences applicable under the domestic law. Exporters are also required to keep the needed paperwork to document origin for a certain period of time. Hence, even with self-certification, there are considerable documentation costs involved. In other instances, a governmental authority may be charged with this certification and there may be additional bureaucratic costs involved.

²²The details of the derivations below are to be found in Appendix B

where τ_{pt}^k for $k = \text{mfn}, \text{pref}$ is the tariff imposed, a_{eipt} is the dummy variable for preference use in the transaction, p_{eipt} is the price in dollars, and ψ_{eipt} is the idiosyncratic demand shifter. Note that the demand in Colombia is a function of the price in Colombian Pesos. Therefore, the relevant price in the demand function is the dollar price times the exchange rate of the importing country vis-a-vis the dollar. Below, we have the exporting firm choosing the price in dollars.²³ As a result, the importing country exchange rate vis-a-vis the dollar serves as a time-variant demand shifter which we use later as an instrument in our regression.

The profit from the transaction for the exporter is:

$$\pi_{eipt}(a_{eipt}) = (p_{eipt}(a_{eipt}) - R_p^{a_{eipt}} c_{et}) q_{eipt}(a_{eipt}) - a_{eipt} \varepsilon_{eipt} F_{eipt}, \quad (2)$$

where c_{et} is the marginal cost of production (in dollars), R_p is the increase of the marginal cost to meet the rules of origin, F_{eipt} is the fixed cost of using preferences, and ε_{et} is the idiosyncratic shock in using preferences in the transaction.²⁴ The exporter maximizes his profit in two steps. First, the exporter determines the price of the transaction with and without using preferences. The profit maximizing price of the transaction using preferences or not ($a_{et} = 0$ or 1) is

$$p_{eipt}(a_{eipt}; c_{ept}) = \frac{\eta}{\eta - 1} R_p^{a_{eipt}} c_{ept}. \quad (3)$$

Note that the price does not depend on the importer country's exchange rate vis-a-vis the dollar.

Then the exporter chooses to use preferences only if doing so raises his profits. In Appendix B, we show that the decision rule is (see equation (A8))

$$a_{eipt} = \mathbb{1} \{ \pi_{eipt}(1) - \pi_{eipt}(0) > 0 \} = \mathbb{1} \{ \ln s_{eipt} - \ln F_{eipt} + \chi_{pt} > \epsilon_{eipt} \}, \quad (4)$$

²³Transactions between the three countries are denominated in dollars. [Gopinath and Stein \(2021\)](#) explain why firms maximize profits in dollars.

²⁴Note that, as pointed out in Section 1, we do not differentiate between fixed costs and sunk costs.

where $s_{eipt} \equiv (\tau_{pt}^{mf n} - \tau_{pt}^{pref})r_{it}(0)$ is the tariff savings by invoking preferences, i.e. the preference margin times the value of the transaction if the preferences had not been met χ_{pt} is a term that is product transaction specific and captures the MFN tariff and the preferential tariff, and the increase in the marginal cost of production from meeting ROOs.

Motivated by equation (4) we estimate a linear probability model.²⁵ For exporter e , importer i , product p , and transaction t , we estimate the following equation:

$$a_{eipt} = \alpha_0 \ln s_{eipt}^* + \alpha_1 Large_{eipt} \ln s_{eipt}^* + \alpha_2 (\tau_{pt}^{mf n} - \tau_{pt}^{pref}) + \alpha_3 \ln(er_t^o) + \sum_{n=1}^4 \beta_n Exp_{et}^n + \beta_{more} Exp_{et}^{more} + \gamma Age_{et} + \delta_{ei} + \delta_p + u_{eipt}, \quad (5)$$

where the dependent variable is one if preferences are used in the transaction and zero otherwise. We allow F_{eipt} to be a function of experience so that an increase in the probability of using preferences with experience can be seen as evidence that fixed costs fall with experience. ϵ_{eipt} is the idiosyncratic cost of using preference. s_{eipt}^* proxies for s_{eipt} . In the data, we cannot observe $r_{it}(0)$ when preferences are used in the transaction. We use $r_{it}(a_{eipt})$ as a proxy for the value of the transaction when preferences are used and define $s_{eipt}^* \equiv (\tau_{pt}^{mf n} - \tau_{pt}^{pref})r_{it}(a_{eipt})$. This creates measurement error. In Appendix C, we show that this measurement error does not necessarily invalidate our results. As is evident in (A10), the variation in χ_{pt} comes from the changes in the tariff margin and the increase in the marginal costs from meeting the ROOs. The former is controlled for by the tariff margin, $(\tau_{pt}^{mf n} - \tau_{pt}^{pref})$, and the latter is accounted for by the product fixed effect, which captures the time invariant component of the ROOs, and by the exchange rate of the origin country $\ln(er_t^o)$, which controls for the change in relative costs of

²⁵We prefer to use the linear probability model as we know of no way to estimate an IV with high dimensional fixed effects in a non-linear setting. Estimates of alternative non linear specifications without IVs are presented in Online Appendix E.2. These estimates are in line with the baseline reported here.

meeting the ROOs that vary over time through the changes in the cost of imported inputs.

Our focus is on the coefficients of the exporter's experience.²⁶ Total experience, denoted by $Exp^n(ai, ap)$ is a dummy variable that takes the value 1 if the *exporter* has total experience of exactly n transactions using preferences to date t , across all importers (ai) and all products (ap).²⁷ This total experience is the sum of four kinds of experience, i.e., across the same importer and the same product (si, sp), the same importer and other products (si, op), the other importers and the same product (oi, sp), and other importers and other products (oi, op). Formally,

$$Exp^n_{et}(ai, ap) = 1 \left\{ \underbrace{Exp_{et}(si, sp)}_{\substack{\text{Same Importer (si),} \\ \text{Same Product (sp)}}} + \underbrace{Exp_{et}(si, op)}_{\substack{\text{Same Importer (si),} \\ \text{Other Products (op)}}} + \underbrace{Exp_{et}(oi, sp)}_{\substack{\text{Other Importers (oi),} \\ \text{Same Product (sp)}}} + \underbrace{Exp_{et}(oi, op)}_{\substack{\text{Other Importers (oi),} \\ \text{Other Products (op)}}} = n \right\}. \quad (6)$$

By allowing for different coefficients on this four-way classification of experience, we can gain a better understanding of the shape of the fixed costs of obtaining preferences, something that has not been done to date. More precisely, we can explore how fixed costs of meeting ROOs vary by the various kinds of experience a firm might have in obtaining preferences and specifically the extent to which experience with using preferences creates positive spillovers across products and/or importers – the possibility that an exporter that learns how to use preferences for a product or with a specific importer might be more likely to use preferences with other products or importers. If, for example, using preferences in another product

²⁶The exporter's experience is relevant because the exporter needs to provide documentation that shows that goods comply with the relevant ROOs. To validate this hypothesis, we assessed whether the importer's experience matters once the exporter's experience was accounted for and found that it did not. These results are available in Online Appendix E.1.

²⁷For example, suppose an exporter at time t has invoked preferences three times to date. Then the dummy is equal to one for the third experience but 0 for the rest. Note that we choose to group more than 4 experiences together because we see from Figure 3 the effects of experience flatten out.

and the same importer in the past has no effect on the likelihood of using preferences in the current transaction, then it must be that costs are not impacted by such an experience.

Given our fixed effects, our estimates are driven by variation in meeting preferences for a given importer-exporter pair in a particular product. These fixed effects, which account for a wide range of potential confounding factors, include exporter-importer (ei) fixed effects and product (p) fixed effects. Thus, the former set of fixed effects control for any unobservable heterogeneity across exporter-importer pairs as well as for any systematic variation in the fixed costs of meeting ROOs across exporters and importers –associated, for instance, with their different initial sourcing structures and locations and hence access to the physical location of the government agencies in charge of reviewing the documents needed to obtain preferences. The latter set of fixed effects absorb any product specific forces at work that might affect the choice of using preferences, including differences in the ROOs across products that raise the costs of meeting them.

The variable Age_{et} is the exporter’s age in terms of years he has been in exporting to Colombia. This is meant to capture any effect on fixed costs of meeting ROOs that vary by the age of the firm as an exporter in this market at the time of the transaction.²⁸

We account for the fact that once the size of the transaction becomes large enough, further increases are unlikely to increase the probability of using preferences (see Figure 2). We do so by interacting savings with a dummy for large transactions (top 10 percent of transaction in value at the country level).²⁹ Furthermore, because firms with high levels of experience are likely to use preferences

²⁸Since our data start from 2000, the age variable is potentially truncated from above. One might be concerned that this would create measurement error and bias our results. However, our exporter-importer fixed effect will control for this. The experience variable could have a similar problem. Note that, in Section 5.2 where we look at products newly covered by the deepening of the trade agreement between Argentina and Colombia, this problem does not exist and our results on the pattern of learning are unaffected.

²⁹The results when the top 5% or top 20% are used instead are available in the Online Appendix.

independent of the size of the transaction, we drop observations of transactions when experience of the exporter at the product level exceeds 30.³⁰

A remaining concern might be that savings could be correlated with the residual for a variety of reasons including reverse causality and simultaneity. We do not take a stand on what the source of the bias is exactly since, whatever that source is, if we can find an instrument for savings, we would be able to correct for it. We use the daily dollar-Colombian peso exchange rate as an instrument, because it affects the demand for imports (see equation (1)) and hence the value of the transaction and the savings variable without having direct effects on preference utilization.³¹ In particular, we use this high frequency data on exchange rate in order to maximize the variation and, specifically, 9-week lags to capture the relevant exchange rate at the time of the order.³²

5 Results

In this section, we present both IV and OLS estimates of equation (5) using the four-way classification of experience as in equation (6).³³ We do so separately for Argentina and for Peru. Recall that Peru has a long-standing FTA with Colombia and, as a consequence, deeper integration with this country than Argentina. Because of these differences we would expect that there is less room for learning in Peru than in Argentina. Thus, we look at them separately and find that there is indeed evidence of stronger learning effects for Argentina. In addition, preferences were extended to additional products in Argentina in the period we study.

³⁰Our results are robust to changing this cutoff when we drop observations of transactions where the experience of the exporter exceeds 20, 40, 60, 100, as well as when the full sample is used. These results are available in the Online Appendix.

³¹Appreciation of the Colombian peso relative to the dollar would raise the willingness to pay for imports, which would raise savings.

³²We do not include time fixed effects for the variation to be fully exploited. We experimented with different lags, and the lag of 9-week used worked best as an instrument.

³³A problem which remains is measurement error. As argued in Appendix C, it would not affect the signs of our estimates, though it would affect the scale so that the ratios of the estimates would be unbiased.

This natural experiment lets us further explore whether our hypothesis is correct. The results of this exercise reveal that the greater learning for Argentina is indeed coming from these newly covered products.

Baseline results for each country, both for the IV and OLS regressions are presented in Tables 3 and 4. The first stage estimates indicate that the exchange rate instruments are strongly correlated with the endogenous saving variables after conditioning by relevant covariates and fixed effects in both countries. As expected, an increase in the price of the dollar in terms of the Colombian peso (depreciation of the peso) reduces the savings for both exporting countries, though less so for large transactions. According to the second stage estimates, savings have a positive effect on preference utilization in both countries, which is weaker for large transactions.

Tables 3 and 4 also show how experience of different kinds affects preference utilization. Note that experience in the same product and with the same importer matters. The coefficients for this kind of experience are positive, significant and increasing. The coefficients are roughly two to three times larger for Argentina than for Peru. This was expected given the shorter history of the FTA between Argentina and Colombia and thus the larger learning potential for its firms. Moreover, while there is some evidence of cross product and cross importer spillovers for Argentina, other kinds of experience do not seem to matter much for Peru, at least once we use the IV. This could be because new products were covered in 2005 in Argentina as the FTA coverage expanded.³⁴ Our results in Table 5 below are consistent with the hypothesis that newly covered products have greater learning.

5.1 Is Learning Increasing with Experience?

While Tables 3 and 4 indicate whether or not the coefficients estimated are significantly different from zero or not, we are also interested in whether they are increasing with experience. In this subsection, we focus on experience with the

³⁴See Section 3.1 for more.

Table 3: Linear Probability Model with Fixed Effects: Argentina

IV				OLS			
$\ln Savings$		0.381				0.026	
		(0.165)				(0.005)	
$\ln Savings$		-0.051				0.043	
$\times Large\ transaction$		(0.022)				(0.007)	
Age		0.016				-0.069	
		(0.013)				(0.022)	
$\ln(er_o)$		-0.031				-0.004	
		(0.031)				(0.001)	
$Margin$		-0.059				-0.010	
		(0.026)				(0.005)	
$Exp^1(si, sp)$	0.061	$Exp^1(oi, sp)$	0.012	$Exp^1(si, sp)$	0.097	$Exp^1(oi, sp)$	0.057
	(0.025)		(0.033)		(0.017)		(0.023)
$Exp^2(si, sp)$	0.094	$Exp^2(oi, sp)$	0.038	$Exp^2(si, sp)$	0.146	$Exp^2(oi, sp)$	0.088
	(0.034)		(0.042)		(0.020)		(0.028)
$Exp^3(si, sp)$	0.135	$Exp^3(oi, sp)$	0.051	$Exp^3(si, sp)$	0.173	$Exp^3(oi, sp)$	0.092
	(0.036)		(0.045)		(0.024)		(0.030)
$Exp^4(si, sp)$	0.093	$Exp^4(oi, sp)$	0.017	$Exp^4(si, sp)$	0.161	$Exp^4(oi, sp)$	0.089
	(0.046)		(0.054)		(0.028)		(0.032)
$Exp^{more}(si, sp)$	0.136	$Exp^{more}(oi, sp)$	-0.022	$Exp^{more}(si, sp)$	0.186	$Exp^{more}(oi, sp)$	0.045
	(0.047)		(0.048)		(0.033)		(0.027)
$Exp^1(si, op)$	0.099	$Exp^1(oi, op)$	0.003	$Exp^1(si, op)$	0.048	$Exp^1(oi, op)$	-0.026
	(0.039)		(0.041)		(0.030)		(0.033)
$Exp^2(si, op)$	0.054	$Exp^2(oi, op)$	0.132	$Exp^2(si, op)$	0.043	$Exp^2(oi, op)$	0.043
	(0.034)		(0.068)		(0.031)		(0.030)
$Exp^3(si, op)$	0.076	$Exp^3(oi, op)$	0.062	$Exp^3(si, op)$	0.057	$Exp^3(oi, op)$	0.034
	(0.040)		(0.074)		(0.033)		(0.070)
$Exp^4(si, op)$	0.161	$Exp^4(oi, op)$	0.061	$Exp^4(si, op)$	0.113	$Exp^4(oi, op)$	0.026
	(0.044)		(0.058)		(0.035)		(0.050)
$Exp^{more}(si, op)$	0.133	$Exp^{more}(oi, op)$	0.135	$Exp^{more}(si, op)$	0.117	$Exp^{more}(oi, op)$	0.069
	(0.037)		(0.052)		(0.035)		(0.044)
Observations		19,579				19,579	
Fixed Effects:							
Exporter-Importer		✓				✓	
Product		✓				✓	
First stage: Savings				First stage: Savings $\times Large\ transaction$			
2 month lagged $\ln(er_{CO})$		-0.490					
		(0.136)					
2 month lagged $\ln(er_{CO})$		0.216				1.614	
$\times Large\ transaction$		(0.008)				(0.009)	
F		46.03				1446	

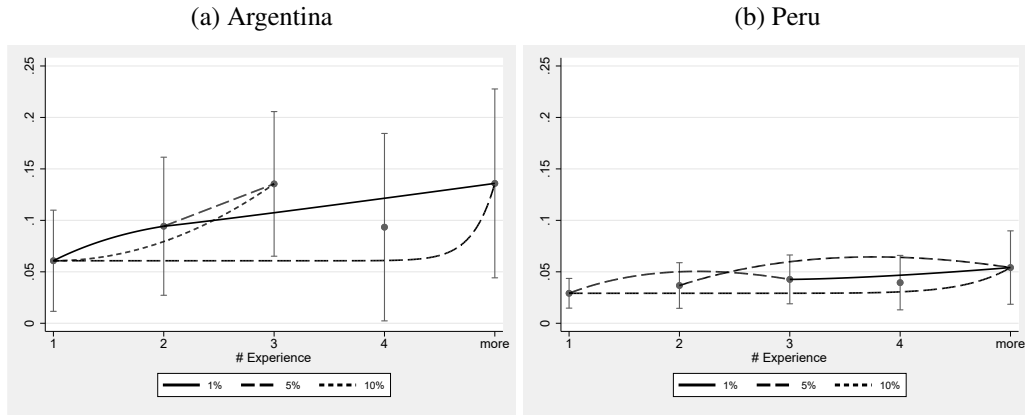
Importer-exporter clustered standard errors in parentheses. sp =same products, op =other goods, si =same importer, oi =other importers.

Table 4: Linear Probability Model with Fixed Effects: Peru

IV				OLS			
$\ln Savings$		0.120 (0.049)				0.010 (0.004)	
$\ln Savings$		-0.011 (0.005)				-0.017 (0.006)	
$\times Large transaction$							
Age		-0.016 (0.006)				0.116 (0.071)	
$\ln(er_o)$		0.211 (0.098)				0.000 (0.000)	
$Margin$		-0.005 (0.003)				0.001 (0.001)	
$Exp^1(si, sp)$	0.029 (0.007)	$Exp^1(oi, sp)$	0.020 (0.009)	$Exp^1(si, sp)$	0.044 (0.008)	$Exp^1(oi, sp)$	0.023 (0.007)
$Exp^2(si, sp)$	0.037 (0.011)	$Exp^2(oi, sp)$	0.010 (0.010)	$Exp^2(si, sp)$	0.056 (0.012)	$Exp^2(oi, sp)$	0.020 (0.007)
$Exp^3(si, sp)$	0.043 (0.012)	$Exp^3(oi, sp)$	0.006 (0.012)	$Exp^3(si, sp)$	0.062 (0.012)	$Exp^3(oi, sp)$	0.022 (0.008)
$Exp^4(si, sp)$	0.040 (0.013)	$Exp^4(oi, sp)$	0.009 (0.013)	$Exp^4(si, sp)$	0.064 (0.012)	$Exp^4(oi, sp)$	0.026 (0.009)
$Exp^{more}(si, sp)$	0.054 (0.018)	$Exp^{more}(oi, sp)$	0.000 (0.014)	$Exp^{more}(si, sp)$	0.084 (0.019)	$Exp^{more}(oi, sp)$	0.019 (0.009)
$Exp^1(si, op)$	-0.012 (0.008)	$Exp^1(oi, op)$	-0.087 (0.022)	$Exp^1(si, op)$	-0.014 (0.007)	$Exp^1(oi, op)$	-0.079 (0.017)
$Exp^2(si, op)$	-0.014 (0.010)	$Exp^2(oi, op)$	-0.012 (0.016)	$Exp^2(si, op)$	-0.020 (0.009)	$Exp^2(oi, op)$	-0.021 (0.016)
$Exp^3(si, op)$	0.002 (0.012)	$Exp^3(oi, op)$	-0.024 (0.018)	$Exp^3(si, op)$	-0.014 (0.009)	$Exp^3(oi, op)$	-0.028 (0.015)
$Exp^4(si, op)$	-0.015 (0.014)	$Exp^4(oi, op)$	-0.029 (0.020)	$Exp^4(si, op)$	-0.024 (0.012)	$Exp^4(oi, op)$	-0.037 (0.018)
$Exp^{more}(si, op)$	0.011 (0.015)	$Exp^{more}(oi, op)$	-0.038 (0.023)	$Exp^{more}(si, op)$	-0.011 (0.009)	$Exp^{more}(oi, op)$	-0.052 (0.023)
Observations		44,776				44,776	
Fixed Effects:							
Exporter-Importer		✓				✓	
Product		✓				✓	
First stage: Savings				First stage: Savings $\times Large transaction$			
2 month lagged $\ln(er_{CO})$		-0.692 (0.148)					
2 month lagged $\ln(er_{CO})$		0.189 (0.007)				1.854 (0.005)	
$\times Large transaction$							
F		45.56				8313	

Importer-exporter clustered standard errors in parentheses. sp =same products, op =other goods, si =same importer, oi =other importers.

Figure 4: Testing Whether Learning Increases with Experience (si,sp)



Note: This figure shows whether the coefficients for experience are significantly increasing and different from zero. The vertical lines at each experience level give the 95% confidence intervals. The lines connecting the different experience levels are only drawn when the coefficients are significantly larger and the significance level is depicted by the coarseness of the lines.

same importer and the same product since this is where most of the action is coming from. We highlight whether the coefficients for each kind of experience are significantly different from zero and whether they are significantly larger as experience rises. We show our results in Figure 4. Panel (a) is for Argentina and Panel (b) is for Peru. The vertical lines at each experience depict the 95% confidence intervals for that estimate. We see, for example, that all of these are significantly positive and that estimates are larger and the confidence intervals are wider for Argentina. If the estimates of the coefficients on level of experience are increasing significantly, we connect the two estimates with a line. If not, there is no line connecting the estimates. The significance level is depicted by the coarseness of the lines. The more continuous the line, the higher the significance. For both countries, the coefficient on the third and fifth or more experience are significantly larger than that on the first. This suggests that there is some evidence that learning is increasing with experience.

5.2 Are Exporters Really Learning?

In this subsection, we present additional evidence that supports our interpretation that a learning mechanism is at work. The deepening of Colombia's preferential trade with Argentina around 2005 provides us with a natural experiment. We ask whether we see more learning in newly covered products compared to that in their counterparts that were already eligible for preference utilization (see Table 3). We might expect this to be the case.³⁵ In such a setting, all exporters are starting with zero experience for newly covered products and there will be less knowledge about how to meet preferences in the economy as a whole. As a consequence, learning from one's own experience would drive preference use to a greater extent for such recently incorporated products. Moreover, the lack of social learning to draw upon in these products can also make one's own experience with other products important. Accordingly, we might also see spillovers coming from other products in this case.

When we consider experience in other products, we need to further allow for whether the other products are newly covered product or not.³⁶ For this reason, we further break experience from other products into experience from other products which were covered before 2005 (other old products) and those that were newly covered in 2005 (other new products). Thus, experience from the same importer and other products is further decomposed into:

$$\underbrace{Exp_{et}(si, op)}_{\text{Same Importer (si), Other Products (op)}} = \underbrace{Exp_{et}(si, op; new)}_{\text{Same Importer (si), Other New Products (op; new)}} + \underbrace{Exp_{et}(si, op; old)}_{\text{Same Importer (si), Other Old Products (op; old)}} . \quad (7)$$

Let d_{new}^p be a dummy variable equal to one if product p is newly covered by the agreement in 2005. Experience variables interacted with this dummy allow

³⁵Recall that the estimates for learning are larger for Argentina than for Peru. Since the FTA has a longer history for Peru, we would have expected this as social learning would blur the effect of learning from individual experience.

³⁶This is clearly not an issue in the case of learning from the same product because we have an interaction with the "new" dummy variable.

us to estimate additional learning effects for newly covered products. Also, by decomposing experience in other products into new and old products, we can identify whether positive learning from other products observed for Argentina is coming from and exists only among new products. This results in the following specification:

$$\begin{aligned}
a_{eipt} = & \sum_{m=si,oi} \sum_{n=1}^5 (\beta_n^{m,sp} Exp_{et}^n(m, sp) + \gamma_n^{m,sp} d_{new}^p Exp_{et}^n(m, sp)) \\
& + \sum_{m=si,oi} \sum_{n=1}^5 (\beta_n^{m,op;new} Exp_{et}^n(m, op; new) + \gamma_n^{m,op;new} d_{new}^p Exp_{et}^n(m, op; new)) \\
& + \sum_{m=si,oi} \sum_{n=1}^5 (\beta_n^{m,op;old} Exp_{et}^n(m, op; old) + \gamma_n^{m,op;old} d_{new}^p Exp_{et}^n(m, op; old)) \\
& + X_{eipt} \beta + u_{eipt},
\end{aligned} \tag{8}$$

where $\beta_n^{m,op;new}$ and $\beta_n^{m,op;other}$ are learning effects from experience in other new and old products, respectively, common to both new and old products. In contrast, γ s capture additional learning effects occurred only for new products.

For concreteness consider a world with four products. Products A and B were previously covered (*old*), while products C and D were newly covered (*new*). An exporter has exported each of the four products once with preferences before time t . For exporting product A, the exporter's experience in using preferences with the same product A is one, his experience with other *new* products is two (coming from C and D), and his experience with other *old* products is one (coming from B). The effects of such experience are captured by the β s associated with them. We allow for the effects for new products to differ from those for old products. These additional effects are captured by coefficients (γ s) on the experience variables interacted with the dummy for new products.

The results from estimating equation (8) are presented in Table 5 for IV and Table 6 for OLS. Note that Table 5 differs from Table 3 for Argentina in a number of ways. To begin with, it allows the products newly covered in 2005 to behave

differently in terms of learning from the products that were covered before 2005. This is done by interacting a dummy “new” with the (si,sp) and (oi,sp), that is, same importer same product, other importer same product, classifications. The dummy new is one if the product is newly covered in 2005. From Table 5 we will see that products that are newly covered are where most of the learning is coming from. This helps explain the greater learning we saw in Argentina in Table 3. Also, we will see that the estimates for learning for old products for Argentina are similar to those for Peru.³⁷

In Table 5, the upper left block (labelled 11) gives the coefficients on experience with the same importer and the same product for old products and the additional effect for new products. For old products, for example, single experience results in a 3.3% increase in using preferences. The additional effect for new products is given in the adjoining column to be 4.6%. In other words, single experience in a new product gives rise to a 7.9% increase. The upper right block (labelled 12) gives the same coefficients but for experience with other importers in the same product. In this block, there is one significantly positive effect for the fourth experience, though the additional effects for new products are by not significantly positive.

The middle left block (labelled 21) reports the effects for the same importer and other products when the other products are new. Common effects are not significantly positive. When the product in the transaction is new, the additional effects are positive and almost always significant. As a result, learning effects from experience in other new products (the sum of the first and second columns) are positive and roughly constant (around 0.16). The middle right block (labelled 22) gives the analogous estimates for experience with other importers. These estimates are similar to those in block 21.

³⁷In essence, we are running a separate regression for these two kinds of products while constraining the coefficients on all other controls to be the same in the two regressions. The estimated coefficients for the new products are the sum of the coefficients with and without the dummy and the coefficients on the old and new products are significantly different from each other if the coefficient on the interaction is significantly different from zero.

The bottom left block (labelled 31) presents the effects for the same importer and other products when the other products are old. These are uniformly insignificant. The bottom right block (labelled 32) shows the analogous estimates for experience with other importers. Note that common effects are sometimes positive and significant, though additional effects for new products are never significantly positive.

To sum up, learning effects with the same product and the same importer in Tables 3 and 4 are higher for Argentina than Peru. Arguably, this is most likely because the FTA between Argentina and Colombia was relatively new compared to that between Peru and Colombia. Evidence presented in Table 3 indicates that this is indeed the case. While estimated effects point to stronger learning for newly covered products, those for old products are much closer to those for Peru in Table 4. As far as experience with other products goes, recall that in Table 3, learning effects from other products were positive for Argentina but not for Peru. Evidence in Table 5 suggests that this is driven by the newly covered products for Argentina. More precisely, estimates differ across old and newly covered products only when the other products are themselves newly covered.³⁸

³⁸It may be argued that exporters are learning about the importer rather than the process of using preferences. Note that there are additional learning effects for the new products covered by the new trade agreement between Argentina and Colombia. These effects cannot be explained by relationship-specific learning, which should be captured by the effects common to both new and old trade agreements.

Table 5: Interaction with Post Agreement Preferences for Argentina (IV)

	Same importer				Other importers			
Same product		(11)				(12)		
	$Exp^1(si, sp)$	0.033 (0.027)	$d_{new}Exp^1(si, sp)$	0.046 (0.034)	$Exp^1(oi, sp)$	0.029 (0.037)	$d_{new}Exp^1(oi, sp)$	-0.020 (0.047)
	$Exp^2(si, sp)$	0.037 (0.035)	$d_{new}Exp^2(si, sp)$	0.106 (0.038)	$Exp^2(oi, sp)$	0.074 (0.052)	$d_{new}Exp^2(oi, sp)$	-0.052 (0.066)
	$Exp^3(si, sp)$	0.089 (0.035)	$d_{new}Exp^3(si, sp)$	0.072 (0.048)	$Exp^3(oi, sp)$	0.065 (0.071)	$d_{new}Exp^3(oi, sp)$	-0.038 (0.078)
	$Exp^4(si, sp)$	0.041 (0.046)	$d_{new}Exp^4(si, sp)$	0.107 (0.055)	$Exp^4(oi, sp)$	0.108 (0.061)	$d_{new}Exp^4(oi, sp)$	-0.152 (0.075)
	$Exp^{more}(si, sp)$	0.031 (0.044)	$d_{new}Exp^{more}(si, sp)$	0.191 (0.058)	$Exp^{more}(oi, sp)$	0.036 (0.058)	$d_{new}Exp^{more}(oi, sp)$	-0.085 (0.062)
Other new product		(21)				(22)		
	$Exp^1(si, op, new)$	-0.164 (0.079)	$d_{new}Exp^1(si, op, new)$	0.300 (0.082)	$Exp^1(oi, op, new)$	-0.060 (0.102)	$d_{new}Exp^1(oi, op, new)$	0.197 (0.103)
	$Exp^2(si, op, new)$	0.027 (0.057)	$d_{new}Exp^2(si, op, new)$	0.137 (0.064)	$Exp^2(oi, op, new)$	0.141 (0.090)	$d_{new}Exp^2(oi, op, new)$	0.012 (0.089)
	$Exp^3(si, op, new)$	0.032 (0.081)	$d_{new}Exp^3(si, op, new)$	0.173 (0.089)	$Exp^3(oi, op, new)$	0.030 (0.124)	$d_{new}Exp^3(oi, op, new)$	0.087 (0.126)
	$Exp^4(si, op, new)$	0.104 (0.083)	$d_{new}Exp^4(si, op, new)$	0.076 (0.091)	$Exp^4(oi, op, new)$	0.131 (0.130)	$d_{new}Exp^4(oi, op, new)$	0.003 (0.109)
	$Exp^{more}(si, op, new)$	0.084 (0.057)	$d_{new}Exp^{more}(si, op, new)$	0.082 (0.064)	$Exp^{more}(oi, op, new)$	-0.013 (0.071)	$d_{new}Exp^{more}(oi, op, new)$	0.167 (0.068)
Other old product		(31)				(32)		
	$Exp^1(si, op, old)$	0.006 (0.044)	$d_{new}Exp^1(si, op, old)$	0.005 (0.057)	$Exp^1(oi, op, old)$	0.028 (0.057)	$d_{new}Exp^1(oi, op, old)$	-0.020 (0.069)
	$Exp^2(si, op, old)$	-0.043 (0.044)	$d_{new}Exp^2(si, op, old)$	0.029 (0.053)	$Exp^2(oi, op, old)$	0.113 (0.051)	$d_{new}Exp^2(oi, op, old)$	-0.083 (0.079)
	$Exp^3(si, op, old)$	-0.036 (0.069)	$d_{new}Exp^3(si, op, old)$	-0.025 (0.068)	$Exp^3(oi, op, old)$	0.023 (0.063)	$d_{new}Exp^3(oi, op, old)$	-0.025 (0.090)
	$Exp^4(si, op, old)$	-0.027 (0.059)	$d_{new}Exp^4(si, op, old)$	0.009 (0.069)	$Exp^4(oi, op, old)$	0.116 (0.082)	$d_{new}Exp^4(oi, op, old)$	-0.226 (0.133)
	$Exp^{more}(si, op, old)$	-0.043 (0.056)	$d_{new}Exp^{more}(si, op, old)$	-0.064 (0.053)	$Exp^{more}(oi, op, old)$	0.055 (0.069)	$d_{new}Exp^{more}(oi, op, old)$	-0.057 (0.072)
Other variables	<i>Savings</i>	0.249 (0.134)	2 month lagged $\ln(er_o)$	0.008 (0.026)	<i>Margin</i>	-0.043 (0.021)		
	<i>Age</i>	0.031 (0.010)	<i>Savings</i> \times Large transaction	-0.033 (0.018)				
Fixed effects	Exporter-Importer	✓						
	Product	✓						
	Observations	19,579						
First stage	Savings		Savings \times Large transactions					
	2 month lagged $\ln(er_{CO})$	-0.525 (0.136)						
	2 month lagged $\ln(er_{CO})$	0.216		1.614				
	\times Large transaction	(0.008)		(0.009)				
F		20.78		837.9				

Table 6: Interaction with Post Agreement Preferences for Argentina (OLS)

	Same importer				Other importers			
	(11)				(12)			
Same product	$Exp^1(si, sp)$	0.044 (0.022)	$d_{new}Exp^1(si, sp)$	0.061 (0.030)	$Exp^1(oi, sp)$	0.044 (0.035)	$d_{new}Exp^1(oi, sp)$	-0.007 (0.045)
	$Exp^2(si, sp)$	0.067 (0.027)	$d_{new}Exp^2(si, sp)$	0.106 (0.035)	$Exp^2(oi, sp)$	0.111 (0.041)	$d_{new}Exp^2(oi, sp)$	-0.067 (0.054)
	$Exp^3(si, sp)$	0.104 (0.029)	$d_{new}Exp^3(si, sp)$	0.081 (0.040)	$Exp^3(oi, sp)$	0.109 (0.054)	$d_{new}Exp^3(oi, sp)$	-0.069 (0.062)
	$Exp^4(si, sp)$	0.074 (0.037)	$d_{new}Exp^4(si, sp)$	0.116 (0.047)	$Exp^4(oi, sp)$	0.145 (0.053)	$d_{new}Exp^4(oi, sp)$	-0.140 (0.066)
	$Exp^{more}(si, sp)$	0.052 (0.035)	$d_{new}Exp^{more}(si, sp)$	0.202 (0.051)	$Exp^{more}(oi, sp)$	0.084 (0.045)	$d_{new}Exp^{more}(oi, sp)$	-0.095 (0.054)
Other new product	(21)				(22)			
	$Exp^1(si, op, new)$	-0.124 (0.069)	$d_{new}Exp^1(si, op, new)$	0.258 (0.074)	$Exp^1(oi, op, new)$	-0.083 (0.094)	$d_{new}Exp^1(oi, op, new)$	0.199 (0.094)
	$Exp^2(si, op, new)$	0.046 (0.046)	$d_{new}Exp^2(si, op, new)$	0.098 (0.050)	$Exp^2(oi, op, new)$	0.126 (0.058)	$d_{new}Exp^2(oi, op, new)$	-0.021 (0.065)
	$Exp^3(si, op, new)$	0.003 (0.078)	$d_{new}Exp^3(si, op, new)$	0.164 (0.088)	$Exp^3(oi, op, new)$	0.120 (0.102)	$d_{new}Exp^3(oi, op, new)$	0.007 (0.107)
	$Exp^4(si, op, new)$	0.074 (0.082)	$d_{new}Exp^4(si, op, new)$	0.084 (0.090)	$Exp^4(oi, op, new)$	0.021 (0.089)	$d_{new}Exp^4(oi, op, new)$	0.111 (0.068)
	$Exp^{more}(si, op, new)$	0.078 (0.055)	$d_{new}Exp^{more}(si, op, new)$	0.080 (0.060)	$Exp^{more}(oi, op, new)$	-0.067 (0.058)	$d_{new}Exp^{more}(oi, op, new)$	0.216 (0.055)
Other old product	(31)				(32)			
	$Exp^1(si, op, old)$	0.011 (0.042)	$d_{new}Exp^1(si, op, old)$	0.028 (0.051)	$Exp^1(oi, op, old)$	0.041 (0.052)	$d_{new}Exp^1(oi, op, old)$	-0.069 (0.063)
	$Exp^2(si, op, old)$	-0.044 (0.043)	$d_{new}Exp^2(si, op, old)$	0.063 (0.049)	$Exp^2(oi, op, old)$	0.073 (0.042)	$d_{new}Exp^2(oi, op, old)$	-0.087 (0.081)
	$Exp^3(si, op, old)$	-0.019 (0.064)	$d_{new}Exp^3(si, op, old)$	-0.020 (0.069)	$Exp^3(oi, op, old)$	0.032 (0.050)	$d_{new}Exp^3(oi, op, old)$	-0.051 (0.087)
	$Exp^4(si, op, old)$	0.020 (0.047)	$d_{new}Exp^4(si, op, old)$	-0.029 (0.055)	$Exp^4(oi, op, old)$	0.132 (0.058)	$d_{new}Exp^4(oi, op, old)$	-0.157 (0.085)
	$Exp^{more}(si, op, old)$	-0.033 (0.053)	$d_{new}Exp^{more}(si, op, old)$	-0.051 (0.053)	$Exp^{more}(oi, op, old)$	0.057 (0.067)	$d_{new}Exp^{more}(oi, op, old)$	-0.100 (0.060)
Other variables	<i>Savings</i>	0.025 (0.005)	2 month lagged $\ln(er_o)$	-0.023 (0.023)	<i>Margin</i>	-0.012 (0.005)		
	<i>Age</i>	0.046 (0.007)		-0.003 (0.001)				
Fixed effects	Exporter-Importer	✓						
	Product	✓						
	Observations	19,579						

6 Conclusion

This paper is the first to cast light on the shape of costs of meeting ROOs using a model-based empirical approach and addressing first-order endogeneity concerns. Our results suggest that costs of using preferences fall with experience. Consequently, policies targeted to new exporters should have large payoffs in terms of preference utilization and exports. Not only would their current use rise, but so would their future use. As larger firms tend to have more experience as well as larger orders, they are more likely to use preferences, so that FTAs could negatively impact competition. Policies encouraging the exports of small young firms would help mitigate this.

References

- Anson, José, Olivier Cadot, Antoni Estevadeordal, Jaime de Melo, Akiko Suwa-Eisenmann, and Bolormaa Tumurchudur. 2005. "Rules of Origin in North–South Preferential Trading Arrangements with an Application to NAFTA." *Review of International Economics*, 13(3): 501–517.
- Benguria, Felipe. 2022. "Do US exporters take advantage of free trade agreements? Evidence from the US-Colombia free trade agreement." *Review of International Economics*, 30(4): 1148–1179.
- Cadot, Olivier, Antoni Estevadeordal, Akiko Suwa-Eisenmann, Akiko Suwa Eisenmann, and Thierry Verdier. 2006a. *The origin of goods: Rules of origin in regional trade agreements*. Oxford University Press on Demand.
- Cadot, Olivier, Céline Carrère, Jaime De Melo, and Bolormaa Tumurchudur. 2006b. "Product-specific rules of origin in EU and US preferential trading arrangements: an Assessment." *World Trade Rev.*, 5: 199.
- Cherkashin, Ivan, Svetlana Demidova, Hiau Looi Kee, and Kala Krishna. 2015. "Firm heterogeneity and costly trade: A new estimation strategy and policy experiments." *Journal of International Economics*, 96(1): 18–36.
- Conconi, Paola, Manuel García-Santana, Laura Puccio, and Roberto Venturini. 2018. "From final goods to inputs: the protectionist effect of rules of origin." *American Economic Review*, 108(8): 2335–65.
- Cornejo, Rafael, and Jeremy Harris. 2007. *Convergence in the rules of origin Spaghetti Bowl: a methodological proposal*. Vol. 34, BID-INTAL.
- Demidova, Svetlana, Hiau Looi Kee, and Kala Krishna. 2012. "Do trade policy differences induce sorting? Theory and evidence from Bangladeshi apparel exporters." *Journal of International Economics*, 87(2): 247–261.
- Devlin, Robert, and Antoni Estevadeordal. 2001. "What's New in the New Regionalism in the Americas?" *INTAL-ITD-STA Working Paper 6*.
- Dinh, Duy, Dzmitry Kniahin, Mondher Mimouni, and Xavier Pichot. 2019. "Global Landscape of Rules of Origin: Insights From The New Comprehensive Database."
- Estevadeordal, Antoni. 2000. "Negotiating preferential market access." *Journal of World Trade*, 34(1).
- Gopinath, Gita, and Jeremy C Stein. 2021. "Banking, Trade, and the making of a Dominant Currency." *The Quarterly Journal of Economics*, 136(2): 783–830.
- Harris, Jeremy Tyler. 2007. *Measurement and determination of rules of origin in preferential trade agreements (PTA's)*. University of Maryland, College Park.
- Head, Keith, Thierry Mayer, and Marc Melitz. 2021. "Moving parts: When

- more restrictive content rules backfire.” *Mimeo*.
- Kasteng, Jonas, Nils Norell, and Patrik Tingvall.** 2022. *Learning by Using Free Trade Agreements: A Firm and Transaction-Level Analysis of the EU–South Korea FTA*. National Board of Trade Sweden.
- Kirk, Robert.** 2007. “Rules of Origin in East Asia.” In *Trade Issues in East Asia : Preferential Rules of Origin*. Washington, DC.. Chapter 1, 3–22. World Bank.
- Krishna, Kala.** 2006. “Understanding Rules of Origin.” In *The Origin of Goods: Rules of Origin in Regional Trade Agreements*. , ed. Olivier Cadot, Antoni Esteveordal, Akiko Suwa-Eisenmann, Akiko Suwa Eisenmann and Thierry Verdier, 19–34. Oxford University Press.
- Krishna, Kala, and Anne Krueger.** 1995. “Implementing free trade areas: Rules of origin and hidden protection.” In *In New Directions in Trade Theory*. , ed. J Levinsohn, A Deardorf and A Stern, 149–198. Ann Arbor: The University of Michigan Press.
- Ludema, Rodney D., Anna Maria Mayda, and Christian Volpe Martincus.** 2021. “Uneven Stumbling Blocks: Latin American Regionalism Revisited.” *mimeo, Georgetown University*.
- Mesquita Moreira, Mauricio, Juan Blyde, Paolo Giordano, Marisol Rodriguez Chatruc, Kun Li, Jeremy Harris, and Krista Lucenti.** 2018. “Connecting the Dots: A Road Map for Better Integration in Latin America and the Caribbean.”
- Pelkmans-Balaoing, Annette O, and Miriam Manchin.** 2007. *Rules of origin and the web of East Asian free trade agreements*. The World Bank.
- UNCTAD.** 2018. “The Use of the EU’s Free Trade Agreements Exporter and Importer Utilization of Preferential Tariffs.”

A Construction of the Dataset

A.1 Data Matching

As explained in Subsection 3.2, the Colombian import database reports both the importing firm’s tax ID and name and the exporting firm’s name, whereas the Argentinean and Peruvian export databases include the exporting firm’s tax ID and name, also for each transaction. Hence, the exporting and importing sides must be merged. We do so primarily using the exporting firms’ names, supplemented with information on the dates, the origin/destination country, the product (code), the value, and the weight.

Firms’ names generally differ in both databases. This could be due to the type of business structure or due to spelling. In the first case, a firm could appear, for instance, as an S.R.L. (Sociedad de Responsabilidad Limitada –the equivalent of a Limited Liability Company in the U.S.–) in one database and as a S.A. (Sociedad Anonima –the equivalent of publicly traded company in the U.S.–) in the other database. In the second case, data can be subject to typos, abbreviations, or missing words in one or both datasets. To address this issue, we first harmonized firms’ names in each dataset separately. Specifically, we removed special and punctuation characters and conjunctions, we then replaced business structures by their acronyms, and finally we abbreviated common words in firms (e.g., Exportadora –Exporter- o Exportaciones –Exports- are replaced by EXP) before eliminating them.

Second, we resorted to a fuzzy matching algorithm (i.e., probabilistic linking) to compare and match the harmonized firms’ names in both pairs of databases (i.e., considering the specific origin/destination countries). This algorithm found the best match (or group of matches) in the standardized data, up to a similarity score of 85%. In the final step we performed a manual review to validate the matches that are a 100% similar and to decide on the matches that are in a range of 85% to 99% of similarity, using, in addition the data on the dates, the product (code), the value, and the weight.

The match is very good. For Argentina, we are able to match 95.1% of exporting firms, 98.9% of transactions, and 99.3% of the value of the transactions. For Peru, we are able to match 94.9% of exporting firms, 99.7% of transactions, and 99.9% of the value of the transactions.

A.2 Data Cleaning

The challenge in the data comes from the absence of clear information on the MFN tariff relevant for the transaction. In order to construct two key variables used in the regressions, namely savings and the preference utilization dummy, we need information on both the MFN tariff and the preferential tariff as well as whether preferences were invoked. There is no field that gives us the MFN tariff. However for each transaction we do know the tariff applied. Hence, we imputed the MFN tariff to be the tariff applied in the given month of the transaction on the same product at HS 10 digit level on Colombia's imports from countries that do not have trade agreements with it at that time. If no such transaction exists, we use the tariff paid on the previous transaction. In the data, some products have more than one MFN tariff in a given month. Since MFN tariffs are very slow to change, we treat these observations as suspect and drop all observations on these products. This loses us 0.19% of the data.

In addition, as transactions of a size below “the de minimis level” are exempt from paying tariffs. We drop all transactions below 200 dollars, which was the de minimis level for Colombia in 2016.³⁹ This loses us 11.6% of the data.

After these two cleaning processes, we construct a preference utilization dummy by comparing the tariff paid to the MFN rate. If the tariff paid is below the MFN rate, we infer that preferences were used in the transaction. By definition, all members of the WTO are given MFN status so that any tariff below the MFN tariff must come from preferential trade agreements. There is a field for the trade agreement used in the transaction, but this field is missing for 84.9% of the data.

³⁹See https://global-express.org/assets/files/Customs%20Committee/de-minimis/GEA-overview-on-de-minimis_April-2016.pdf.

For this reason we could not use this variable. The preferential tariff for a given product exported by a given country is constructed as the most recent tariff paid below the MFN tariff to date on the 10 digit product by exporters from the particular country.

Once we know when preferences were used, we can construct for each exporter a history of experience in using preferences in past transactions. These experiences can also be broken down into four categories used in the paper. The preference margin is defined as the difference in the MFN tariff and the preferential tariff. This allows us to construct the savings variable as the preference margin times the FOB value of the transaction. Perceptive readers might be concerned about the products where preferences exist but are not utilized. In our construction, this would show up as the MFN tariff being equal to the tariff paid in all transactions of the product for a particular year and constructed preference margin would be zero, while the true margin could be positive. But this does not matter because we control country-product-year fixed effects.

We also drop observations for which tariff paid exceeded the MFN tariff. This can happen if, for example, anti-dumping or countervailing tariffs were being imposed. Since these instances occur when other forces are in play, we dropped these transactions. This loses us 2.6% of the data.

Preferential tariffs are usually phased in over time, in other words, they are negotiated to fall over the period until they reach the negotiated level. As a result, when we see preferential tariffs that rise over a period, we are concerned. To be cautious, we drop the products in the entirety from the data. This loses us 31.1% of the data. Note however that as the histories were constructed before dropping these observations, we are not concerned about this affecting our experience variables.

B A Simple Model

Suppose an exporter and an importer meet randomly at period t . Importer i has the following constant elasticity demand system

$$q_{eipt} = \left((\tau_{pt}^{\text{pref}})^{a_{eipt}} (\tau_{pt}^{\text{mfn}})^{1-a_{eipt}} p_{eipt} \right)^{-\eta} x_{eipt}, \quad (\text{A1})$$

where p is the product of the transaction, a_{eipt} is the dummy variable which is equal to one if preference is applied in the transaction, τ_{pt}^h for $h = \text{pref}, \text{mfn}$ summarizes one plus the tariff to pay with and without preferences, p_{eipt} is the price of the transaction, x_{eipt} is the demand shock of the importer, and η is the constant elasticity of demand.

Exporter e maximizes the profit of the transaction by choosing price and determines if preferences are used or not in the transaction. The profit of the transaction is

$$\pi_{eipt} = (p_{eipt} - R_j^{a_{eipt}} c_{eipt}) q_{eipt} - a_{eipt} \varepsilon_{eipt} F_{eijt}. \quad (\text{A2})$$

Using preferences is costly in two ways. First, the exporter needs to meet the rules of origin that may increase the marginal costs of production $R_p \geq 1$. Second, the exporter needs to pay the fixed costs of documentation for the certificate of origin, $F_{eipt} > 0$. This fixed cost may depend on their experience in using preferences in the past transactions. We also assume that this costs includes some shocks $\varepsilon_{eipt} > 0$, which rationalizes why some transactions with large value do not apply preference.

Assume that exporters solve the problem in two stages. In the first stage, exporters decide whether they use preferences or not. In the second stage, they set the price. We can solve the problem backwards. Given the preference application, the exporter sets the price that maximizes equation (A2). The first order condition

implies

$$p_{eipt}(a_{eipt}) = \frac{\eta}{\eta - 1} R_p^{a_{eipt}} c_{eipt}. \quad (\text{A3})$$

Substituting this into equation (A2), we can derive the profit as a function of preference application

$$\pi_{eipt}(a_{eipt}) = \frac{1}{\eta} r_{eipt}(a_{eipt}) - a_{eipt} \varepsilon_{eipt} F_{eipt}, \quad (\text{A4})$$

where $r_{eipt}(a_{eipt})$ is the value of the transaction

$$\begin{aligned} r_{eipt}(a_{eipt}) &= \left(\left(\frac{\tau_{pt}^{\text{mfn}}}{\tau_{pt}^{\text{pref}}} \right)^\eta R_p^{-(\eta-1)} \right)^{a_{eipt}} (\tau_{pt}^{\text{mfn}})^{-\eta} \left(\frac{\eta}{\eta - 1} c_{eipt} \right)^{1-\eta} x_{eipt} \\ &= \left(\left(\frac{\tau_{pt}^{\text{mfn}}}{\tau_{pt}^{\text{pref}}} \right)^\eta R_p^{-(\eta-1)} \right)^{a_{eipt}} r_{eipt}(0) \end{aligned} \quad (\text{A5})$$

$$\equiv B_{pt}^{a_{eipt}} r_{eipt}(0), \quad (\text{A6})$$

where $B_{pt}^{a_{eipt}}$ is the benefits of using preferences, which we assume to be more than one. Given the profits with and without the preference, the exporter chooses to use the preference if the profit with preference is larger than the other

$$\begin{aligned} a_{eipt} &= \mathbb{1} \{ \pi_{eipt}(1) - \pi_{eipt}(0) > 0 \} \\ &= \mathbb{1} \left\{ \frac{1}{\eta} (r_{eipt}(1) - r_{eipt}(0)) > \varepsilon_{eipt} F_{eipt} \right\}. \end{aligned} \quad (\text{A7})$$

We can then rewrite the problem as follows⁴⁰

$$\begin{aligned}
a_{eipt} &= \mathbb{1} \left\{ \ln \left(\left(\frac{\tau_{pt}^{\text{mfn}}}{\tau_{pt}^{\text{pref}}} \right)^\eta R_p^{-(\eta-1)} - 1 \right) + \ln r_{eipt}(0) - \ln \eta > \ln \varepsilon_{eipt} + \ln F_{eipt} \right\} \\
&= \mathbb{1} \left\{ \ln(\tau_{pt}^{\text{mfn}} - \tau_{pt}^{\text{pref}}) r_{eipt}(0) - \ln F_{eipt} \right. \\
&\quad \left. + \ln \left(\frac{\left(\frac{\tau_{pt}^{\text{mfn}}}{\tau_{pt}^{\text{pref}}} \right)^\eta R_p^{-(\eta-1)} - 1}{\tau_{pt}^{\text{mfn}} - \tau_{pt}^{\text{pref}}} \right) - \ln \eta > \ln \varepsilon_{eipt} \right\} \\
&= \mathbb{1} \{ \ln s_{eipt} - \ln F_{eipt} + \chi_{pt} > \ln \varepsilon_{eipt} \}, \tag{A8}
\end{aligned}$$

where

$$s_{eipt} \equiv (\tau_{pt}^{\text{mfn}} - \tau_{pt}^{\text{pref}}) r_{eipt}(0), \tag{A9}$$

$$\chi_{pt} \equiv \ln \left(\frac{\left(\frac{\tau_{pt}^{\text{mfn}}}{\tau_{pt}^{\text{pref}}} \right)^\eta R_p^{-(\eta-1)} - 1}{\tau_{pt}^{\text{mfn}} - \tau_{pt}^{\text{pref}}} \right) - \ln \eta. \tag{A10}$$

C Measurement Error

One potential problem is that, in the data, we cannot observe $r_{eipt}(0)$ if $a_{eipt} = 1$. We therefore use $r_{eipt}(1)$ as a proxy to construct the saving variable. However, if the simple model is actually the data generating process, this proxy generates the measurement error bias. Specifically, our saving variable can be expressed as follows

$$\ln s_{eipt}^* = \ln s_{eipt} + a_{eipt} \Delta \ln r_{eipt}, \tag{A11}$$

⁴⁰Both sides of equation (A7) are positive so that we can take logs. We have already assumed that ε_{eipt} , F_{eipt} , and $r_{eipt}(1) - r_{eipt}(0) = (B_{pt} - 1)r_{eipt}(0)$ are positive.

where

$$\Delta \ln r_{eipt} \equiv \ln r_{eipt}(1) - \ln r_{eipt}(0) = \ln \left(\left(\frac{\tau_{pt}^{\text{mfn}}}{\tau_{pt}^{\text{pref}}} \right)^\eta R_p^{-(\eta-1)} \right) \equiv \frac{\kappa_{pt}}{\gamma} > 0, \quad (\text{A12})$$

where the second equality comes from equation (A5) in Appendix B. Notice that the gain in the log of revenue from using preferences depends only on product-time specific parameters.

Let $\ln \varepsilon_{it}$ be independently drawn from a distribution $G(\cdot)$. Then we have

$$\Pr(a_{eipt} = 1 | X_{eipt}) = G(\ln s_{eipt} - \ln F_{eipt} + \chi_{pt}), \quad (\text{A13})$$

where X_{eipt} summarizes $(c_{eipt}, x_{eipt}, F_{eipt}, \tau_{pt}^{\text{mfn}}, \tau_{pt}^{\text{pref}}, R_p)$. Suppose $G(\cdot)$ is a uniform distribution with mean μ and density γ

$$G(x) \equiv \gamma \left(x + \frac{1}{2\gamma} - \mu \right). \quad (\text{A14})$$

Then we can derive the probability of using preference conditional on the saving from the transaction, fixed cost, and product-time specific parameter as

$$\Pr(a_{eipt} = 1 | X_{eipt}) = \gamma \ln s_{eipt} - \gamma \ln F_{eipt} + \gamma \chi_{pt} + \frac{1}{2} - \gamma \mu, \quad (\text{A15})$$

which can be summarized as

$$\Pr(a_{eipt} = 1 | X_{eipt}) = \gamma \ln s_{eipt} - \mathcal{F}_{eipt} + \xi_{pt}. \quad (\text{A16})$$

Substituting the definition of our saving variable into equation (A16), we have

$$\Pr(a_{eipt} = 1 | X_{eipt}) = \gamma \ln s_{eipt}^* - \mathcal{F}_{eipt} + \xi_{pt} - a_{eipt} \kappa_{pt}, \quad (\text{A17})$$

where X_{eipt} summarizes $(c_{eipt}, x_{eipt}, F_{eipt}, \tau_{pt}^{\text{mfn}}, \tau_{pt}^{\text{pref}}, R_p)$. Define the difference

between the realized and expected values of a_{eipt} conditional on X_{eipt} as follows

$$u_{eipt} \equiv a_{eipt} - E[a_{eipt}|X_{eipt}]. \quad (\text{A18})$$

Note that the conditional expectation of u_{eipt} is zero ($E[u_{eipt}|X_{eipt}] = 0$) and $E[a_{eipt}|X_{eipt}] = \Pr(a_{eipt} = 1|X_{eipt})$. By substituting equation (A18) into (A17) and rearranging it, we obtain the empirical specification we used in the previous sections.

$$a_{eipt} = \frac{\gamma}{1 + \kappa_{pt}} \ln s_{eipt}^* - \frac{\mathcal{F}_{eipt}}{1 + \kappa_{pt}} + \frac{\xi_{pt}}{1 + \kappa_{pt}} + \frac{u_{eipt}}{1 + \kappa_{pt}}. \quad (\text{A19})$$

There are a few implications of equation (A19) to highlight here. Note that the bias introduced by mis-measurement of the savings variable creates another source of endogeneity bias in addition to any endogeneity biases arising from omitted variable biases or reverse causation discussed in the body of the paper. Since we have an instrument for savings, this will take care of endogeneity biases irrespective of their source. Hence, we know that our estimate of the coefficient of savings using the IV will be downward biased due to mis-measurement ($\kappa_{pt} > 0$).⁴¹

Our interest in the paper is primarily on the coefficients of experience. Although the levels of these coefficients are biased downward, it is worth noting that their ratios are unbiased. Thus, an increasing pattern in the coefficient of the experience variables still indicates falling fixed costs with experience.

⁴¹When we look at equation (A19), we see κ_{pt} is likely to be small as long as the preference margin is small, the marginal cost of meeting ROOs is small, the demand elasticity is not too large, and the variance of the shock, ε_{eipt} is large. For example, suppose the tariff ratio is 1.1 (the average preference margin (in level) is 4.7% for Argentina and 14.4% for Peru), R is 1.1 (R needs to be small enough relative to the preference margin for preference use to be profitable), the demand elasticity is 4 (average in the literature), and $\gamma = 0.2$ (recall that γ is the coefficient of saving). In this example, we have $\kappa \approx 0.02$. In this case, the bias is small.