Preferential Trading in Agriculture: New Insights from a Structural Gravity Analysis and Machine Learning

Dongin Kim*

Abstract

Preferential Trade Agreements (PTAs) may yield benefits in agricultural trade liberalization. Still, empirical findings quantifying these benefits are ambiguous, partly because of the complicated and diverse provisions that may promote or hinder agricultural trade. Modern PTAs contain hundreds of provisions in addition to tariff reductions in areas as diverse as competition policy, quantitative restrictions, Technical Barriers to Trade (TBT), Sanitary and Phytosanitary (SPS), or intellectual property rights. Existing research has struggled with overfitting and severe multicollinearity problems when trying to estimate the effects of these provisions on trade flows. This paper uses the plug-in Lasso, a machine learning approach, to select the most related provisions and quantify their impact on agricultural trade flows. The results show that PTA provisions related to anti-discriminatory policies, SPS and TBT measures, and geographical indication promoted agricultural trade between partner countries. However, these effects are inconsistent over different income levels, calling into question the current PTA regime and its development directions for sustainable integration of global agriculture.

Keywords: Preferential trade agreement, treaty provisions, agricultural trade, three-way gravity model, Lasso feature selection

JEL codes: Q17; F13; F14

October 26, 2022

* The latest version can be found here.

* Dongin Kim, Department of Agricultural and Resource Economics, University of Connecticut
Email: dong.i.kim@uconn.edu

1. Introduction

International trade plays a vital role in agriculture, and governments aim to shape their countries' sustainable export and import strategies. As a result, countries have rapidly increased their participation in preferential trade agreements (PTAs) in the last two decades. The lack of progress in multilateral trade negotiations creates the fear of losing market share by being excluded from PTAs, which partly explains the increasing trend in signing PTAs to sustain global market access (Rocha and Ruta, 2022). As of March 2022, the World Trade Organization (WTO) reports 354 active PTAs corresponding to 577 notifications from WTO members (WTO, 2022). Since the end of the 1990s, countries began signing bilateral and regional PTAs, and approximately 13 PTAs have been enforced each year for the last two decades. In the meantime, attention has shifted from the import tariff reduction to the role of non-tariff barriers and behind-the-border policies, such as mitigating differences in regulations and technical standards across nations or protecting intellectual property rights (Disdier et al., 2008; Duvaleix et al., 2021; Santeramo and Lamonaca, 2019). Consequently, the number of provisions has risen with the increase in the average number of policy areas covered in each PTAs. For illustration, most enforced PTAs covered more than 20 policy areas after the mid-2000s, in contrast to less than ten policy areas in the 1990s and before. [1] Since many of the provisions and regulations included in PTAs go beyond WTO commitments, it is fair to say that preferential liberalization shapes the global governance of trade in the twenty-first century (Horn et al., 2010).

This article studies the impact of PTAs on agricultural trade, accounting for heterogeneity caused by the diversity in the contents of each PTA. The role of PTA and its provision varies from deliberate to reactive depending on the country's unique agricultural strategy. Some governments actively find access to the global market to take advantage of scale economies in the larger market. In contrast, others consider agriculture an element of a broader set of complex trade-offs in trade agreements.[2]

[1] An illustration of this development is the comparison of the PTA that the EU signed with Egypt in 1972, 92 pages long, with the 2016 Comprehensive Economic and Trade Agreement (CETA) between Canada and the EU, which is 1,598 pages long (Baccini, 2019).

[2] For example, South Korea has aggressively sought and signed agreements to export more of its manufacturing goods, trading off its domestic agricultural competitiveness (Kawai and Wignaraja, 2011).

In addition, some countries use non-tariff barriers to compensate for tariff cuts (Cheong et al., 2018). Consequently, PTAs have an underlying rationale for contributing to the global agricultural and food market in various ways, and the provisions related to the agriculture sector became of interest (Chauffour and Maur, 2011). Thus, addressing preferential trade in agriculture centered on understanding whether the individual or the set of PTA provisions promotes or hinders agricultural trade.

PTAs cover a wide range of behind-the-border policies, namely the non-tariff barriers (NTBs), such as competition policies, product standards, regulatory regimes, investment codes, environmental policies, and intellectual property rights protection (Disdier et al., 2008; Duvaleix et al., 2021; Santeramo and Lamonaca, 2019). The trade effect of NTBs can be either positive or negative, while legally binding agreements can reduce uncertainty for traders arising from unilateral policy interventions (Bagwell and Staiger, 2004; Limão, 2006). The primary NTB provisions that affect agricultural trade are sanitary and phytosanitary (SPS) and technical barriers to trade (TBT). The SPS and TBT provisions raise trade costs with stricter regulation and negatively affect agricultural trade (Disdier et al., 2008; Peterson et al., 2013; Murina and Nicita, 2017). Subsidies and competition policy are also the core of agricultural elements in PTAs following the transparency and non-discrimination principle. Recently, geographical indication (GI) became the center of agricultural and food trade research as the rise of protecting intellectual property rights, led by European Union (EU) countries. Many studies have discussed its impact within and outside the EU member states (Josling, 2006; Moschini et al., 2008; Huysmans and Swinnen, 2019; Duvaleix et al., 2021; Curzi and Huysmans, 2022).

A substantial literature looks at the impact of PTAs on a multitude of economic, political, and social phenomena (see, for a review, Baccini, 2019). Political scientists have focused on a broad set of outcomes ranging from domestic economic reforms, behind-the-border protectionism, and human rights protection in PTA signatories (e.g., Mansfield and Pevehouse, 2000; Mansfield and Reinhardt, 2003; Chauffour and Maur, 2011). For trade economists, the international trade flow implications are of critical interest (e.g., Limão, 2006; Baier and Bergstrand, 2007; Egger et al., 2011). Some literature studies the association between PTAs and agricultural trade, looking, among others, at trade creation and diversion (e.g., Grant and Lambert, 2008; Sun and Reed, 2010; Scoppola et al.,

2018) and economic welfare effects (e.g., Koo and Kennedy, 2006; Croser and Anderson, 2011; Martin, 2018). However, what has been missing in the existing literature is the distinction between tariff and non-tariff changes under PTAs. Their findings are limited to explaining the effects of PTA presence, ignoring unobserved heterogeneity due to the diversity in the contents of each PTA.

Recent studies established the relative importance of individual provisions in determining the agreement's overall impacts (Kohl et al., 2016; Hofmann et al., 2017; Campi and Dueñas, 2019; Hofmann et al., 2019; Breinlich et al., 2022; Mattoo et al., 2022). Their attempts faced technical difficulties in estimating too many variables with the traditional econometric methods and suffered from controlling for the fact that similar provisions repeatedly appear in a single agreement. In sum, these concerns leave the overfitting and severe multicollinearity problems to researchers making strong assumptions on the individual provision's importance (e.g., Kohl et al., 2016; Santeramo and Lamonaca, 2021). Alternatively, researchers have grouped or aggregated provisions to make an index and avoid the econometric issues. For example, Campi and Dueñas (2019) grouped intellectual property right provisions to indicate differential effects of PTAs. Hofmann et al. (2017) and Mattoo et al. (2022) show the heterogeneous effects of PTA by using the count of provisions in an agreement as its depth measure, in which they implicitly give equal weight to each provision. Nevertheless, the proposed methods in these studies still have the disadvantage of requiring *ad hoc* assumptions on how to aggregate individual provisions.

This paper adopts a data-driven variable selection algorithm to single out the most effective provisions instead of making assumptions about the importance of individual provisions or aggregating them in arbitrary ways. I then quantify their impact on trade flows using traditional econometric methods. In particular, I use a machine learning (ML) technique combined with the parsimonious three-way gravity framework. Although trade and applied economists disregarded ML for a long time, there is a growing interest in its theoretical foundation and applications (Athey and Imbens, 2019; Storm et al., 2020). Several studies point toward the benefits of big data and ML systems for economic analysis (Einav and Levin, 2014; Bajari et al., 2015; Mullainathan and Spiess, 2017; Athey and Imbens, 2017). Recent ML applications in international economics are broadening into two primary subjects: forecasting trade flows (Gopinath et al., 2021) and trade policy evaluation (Batarseh et al., 2021; Breinlich et al., 2022).

From a functional perspective, one advantage of ML is that it uses data-driven approaches for model selection that can avoid the potential identification problem. In any ML method, the data determine the functional form rather than performing a single estimation which requires the researcher provides a list of covariates. Thus, ML is an "algorithm" rather than an estimation method because it estimates many alternative models and then selects among them to maximize a criterion. This approach contrasts with the traditional econometric approach, where a model is picked based on the researchers' hypotheses. Moreover, the ML can improve the possibility of being systematic and fully descriptive regarding the process of how it selects the final model (Athey and Imbens, 2019). These methodological advances offer applied and trade economists a comprehensive set of reliable techniques for variable selection and regularization to enhance the resulting statistical model's prediction accuracy and interpretability (Breinlich et al., 2022).

I use Lasso regression, a featured ML method for the regularized regression, to select the effective PTA provisions and measure their impact on agricultural trade flows. To consider the case of nonlinear models with high-dimensional fixed effects, which have become standard in the economic analysis of trade flows, I use the Poisson pseudo-maximum likelihood (PPML) version of Lasso regression. In particular, I use the plug-in approach of Belloni et al. (2016), which accounts for heteroskedasticity and clustered errors. I apply this PPML-Lasso model to a comprehensive PTA provisions data set of the World Bank provided by Mattoo et al. (2020). This data set contains very detailed mapped provision variables over 900. Because these variables are highly correlated with one another in nature, regularized regression is a proper estimation strategy. I also conduct the second-stage regression on unselected provisions to obtain the bundles of provisions that affect agricultural trade because the plug-in method may select too few variables.

This paper contributes to several existing literature. First, the results will add to the literature on NTBs in trade by empirically testing to what extent behind-the-border policies of PTAs have promoted agricultural exports. The existing literature has decomposed the PTA effects by individual provisions or studied the impact of individual non-tariff measures of WTO commitments. The impact of SPS and TBT measures have been popular topics in agricultural trade research, while GI has become the center of recent studies with the rise of intellectual property right in the modern PTAs. However, the existing literature is somewhat inconclusive on whether such policies promote or hinder agricultural trade. My findings provide robust empirical evidence of the positive impact of SPS, TBT, and GI in agricultural trade between partner countries. In addition, I found other effective provisions under diverse policy areas, such as export taxes, subsidies, and trade facilitation. The effects are estimated without *ad hoc* assumptions, so it avoids the potential identification problem. In addition, I incorporate the results with the comparison of food and non-food manufacturing. This comparison sheds light on the differential effects of PTAs across sectors, which have been treated the same and thus ignored in the previous empirical literature. I obtained similar provisions selected for the food manufacturing sector, but the provisions chosen for the non-food manufacturing trade are significantly different. Although the bundles of provisions in the second-stage analysis cover more policy areas that duplicate over sectors, discrepancies are still considerable across sectors.

I also contribute to the growing literature on the PTA's heterogeneous effects on trade flows. Some literature estimate agreement-specific effects to account for differential effects of PTAs (Baier and Bergstrand, 2007; Bergstrand et al., 2015) and considers a two-stage methodology to obtain the average trade agreement effects using the variance calculated from the country-pair-specific agreement estimates (Baier et al., 2019). This branch of study account for the unobserved heterogeneity across agreements, but their empirical findings do not explain what determines the heterogeneity. Other studies rely on the fact that each PTA contains a different set of provisions. For instance, Hofmann et al. (2017) propose a depth measure of PTAs based on principal components analysis. However, the suggested method requires the researcher's assumption on assigning values to each provision, so it is not free from the identification problem. My data-driven approach based on an ML technique avoids such issues. I modified Breinlich et al. (2022)'s framework to make it consistent with the structural gravity model and obtained robust estimations of individual PTA provision effects using the new data on domestic and international trade, recently made available by Borchert et al. (2021). Because the estimate of the dummy variable approach on the PTA presence disappears once jointly estimated with the set of selected provisions, I found that the inclusion of effective provisions causes the differential effects of PTA.

Finally, I contribute to development studies in the international trade policy context. Although rarely designed as explicitly protectionist measures, NTBs nonetheless have the potential to raise trade costs. The proliferation of NTBs, such as SPS or environment-related measures, became a threat to exports from developing countries, which have adaptability constraints by technical and financial capacity (Disdier et al., 2008). Studies found the asymmetry in the cost effects on trade flows, but their conclusions vary. Disdier et al. (2008) suggest that developing countries suffer more from the SPS and TBT measures, while Shepherd and Wilson (2013) claim that developed countries are hurt more by EU standards than developing countries. My findings suggest another aspect. The selected PTA provisions decrease agricultural exports from developed to developing countries but increase trade between countries with similar income levels. However, the effects are inconclusive on exports from developing to developed countries. These findings suggest that the selected PTA provisions, such as SPS measures, may facilitate agricultural trade under a specific condition that they are willing to adopt the new standards. Once they adopt and pass the regional standards, they can signal their quality of goods to foreign consumers, so they export more of their products (Disdier et al., 2008). However, this potential promotion effect is only practical between similar-income countries, where both incentive and capability exist for the exporters. Instead, NTBs can be used in a protectionist way between different income countries, and if so, decrease trade.

The remainder of the paper is organized as follows. Section 2 illustrates the development of PTA and the provisions that might affect agricultural trade. Section 3 develops the theoretical framework based on the structural gravity model and presents the estimating framework using the regularization methods. Section 4 describes the data. Section 5 provides the empirical findings. Section 6 concludes.

2. The Development of PTA and the Treatment of Agriculture

Since the early 1990s, governments have increasingly looked to preferential trade agreements (PTAs) in bilateral or regional contexts to promote trade between the partnered countries. The purpose is to deepen economic and political integration among the contracted partners beyond the multilateral commitment, such as WTO. Figure 1 shows that the number of agreements has grown relatively steadily over the course of the last two decades, reaching a total of 313 active PTAs in 2017.[3] Reflecting the shifting priorities of governments, the PTAs' content and configuration have changed, and their increased uptake has been accompanied by a shift in the characteristics of the agreements themselves (Dür et al., 2014; Hofmann et al., 2017). Accordingly, the nature of the commitments made for agriculture may also change over time. This section explores such changes in PTAs and presents provisions possibly affecting agricultural trade flows.

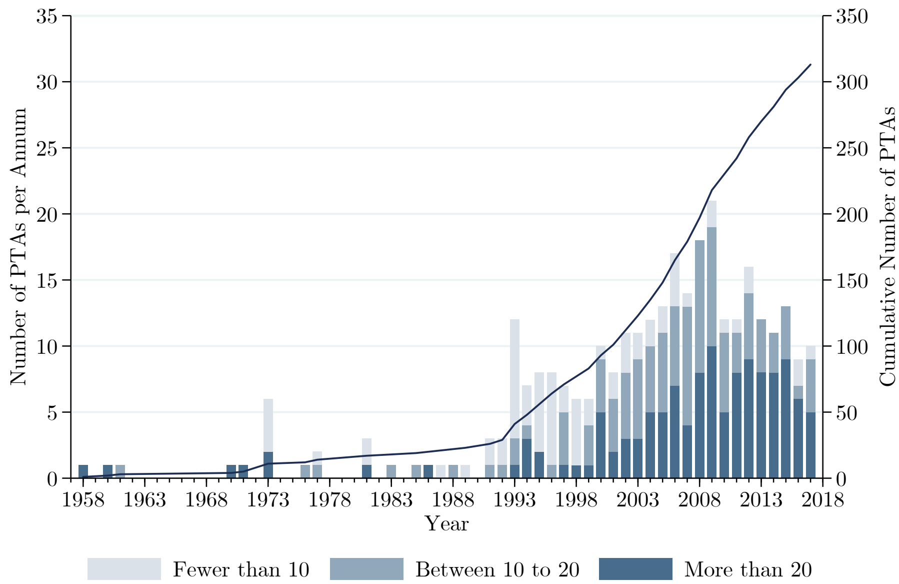

Figure 1: Number of Policy Areas Covered in PTAs

Notes: The figure shows the number of PTA enforced annually from 1958 to 2021. Colors within the bars indicate the number of policy areas covered within those PTAs enforced each year. The policy areas defined by Hofmann et al. (2017) are based on the most frequently appeared provisions in active PTAs from 1958 to 2017. The maximum number of the policy area is 52.

The number of policy areas covered by PTAs has increased in the last two decades as the partner countries seek deep integration. As shown in Figure 1, most agreements covered fewer than ten policy areas until the late 1990s. However, since the 2000s, most new PTAs have covered between 10 and 20 policy areas, with some having even more than 20. [4] The inclusion of new policy areas in

[3] As of March 2022, active PTAs reached 354, corresponding to 577 notifications from WTO members, counting goods, services, and accessions separately (WTO, 2022).

[4] The policy areas in Figure 1 and Figure 2 are based on the most frequently appeared provisions in active PTAs from 1958 to 2017, as defined in (Hofmann et al., 2017). Specifically, the database includes 52 policy areas classified based on the active PTAs.

PTAs is not random. As shown in Figure 2, trade agreements covering fewer policy areas generally focus on traditional trade policy, such as tariff liberalization or customs, accustomed by WTO commitment. Agreements with broader coverage (between 10 and 20 policy areas) tend to include trade-related regulatory issues, such as subsidies or technical barriers to trade. Finally, agreements with more than 20 provisions often include policy areas that are not directly related to trade, such as labor, environment, and movement of people.

WTO commitment strongly influenced to structure of early PTA provisions. The policy areas under the current mandate of WTO include customs regulations, export taxes, anti-dumping, SPS measures, countervailing measures, or TBT measures (see Figure 2). As the regime shifts to a globalized and more integrated economy, Thompson-Lipponen and Greenville (2019) found that most trade agreements have included tariff cuts and other market access concessions exceeding individual countries' WTO commitments for agricultural products. For example, over 90% of PTAs in their study reduced tariff rates below the most favored nation rate of WTO, often scheduled to be duty-free. Moreover, increasing numbers of PTAs also reportedly include norms for SPS, TBT, and trade remedies since the mid-1990s (WTO, 2022).

The SPS and TBT measures are most likely to be included in agreements between developed and developing countries and are least likely to feature in agreements among developed countries (Baccini, 2019). If these measures are used in a protectionist way, they impede trade flows (Peterson et al., 2013; Murina and Nicita, 2017). However, in the case of incomplete information on traded products, the SPS and TBT measures can facilitate trade by signaling that products are safe to consumers (Disdier et al., 2008). They are also being used for greater transparency and equivalence-based measures, thus potentially facilitating trade flows (Thompson-Lipponen and Greenville, 2019).

The elimination of intra-regional tariffs has increased the appeal of trade remedies, including agriculture. Of the three forms of remedies explored, safeguards are the most commonly-observed remedy, followed by anti-dumping and countervailing duties. These measures are often believed to protect domestic agricultural products, while some studies find them ineffective due to the trade diversion effects (Carter and Gunning-Trant, 2010; Carter and Steinbach, 2018).

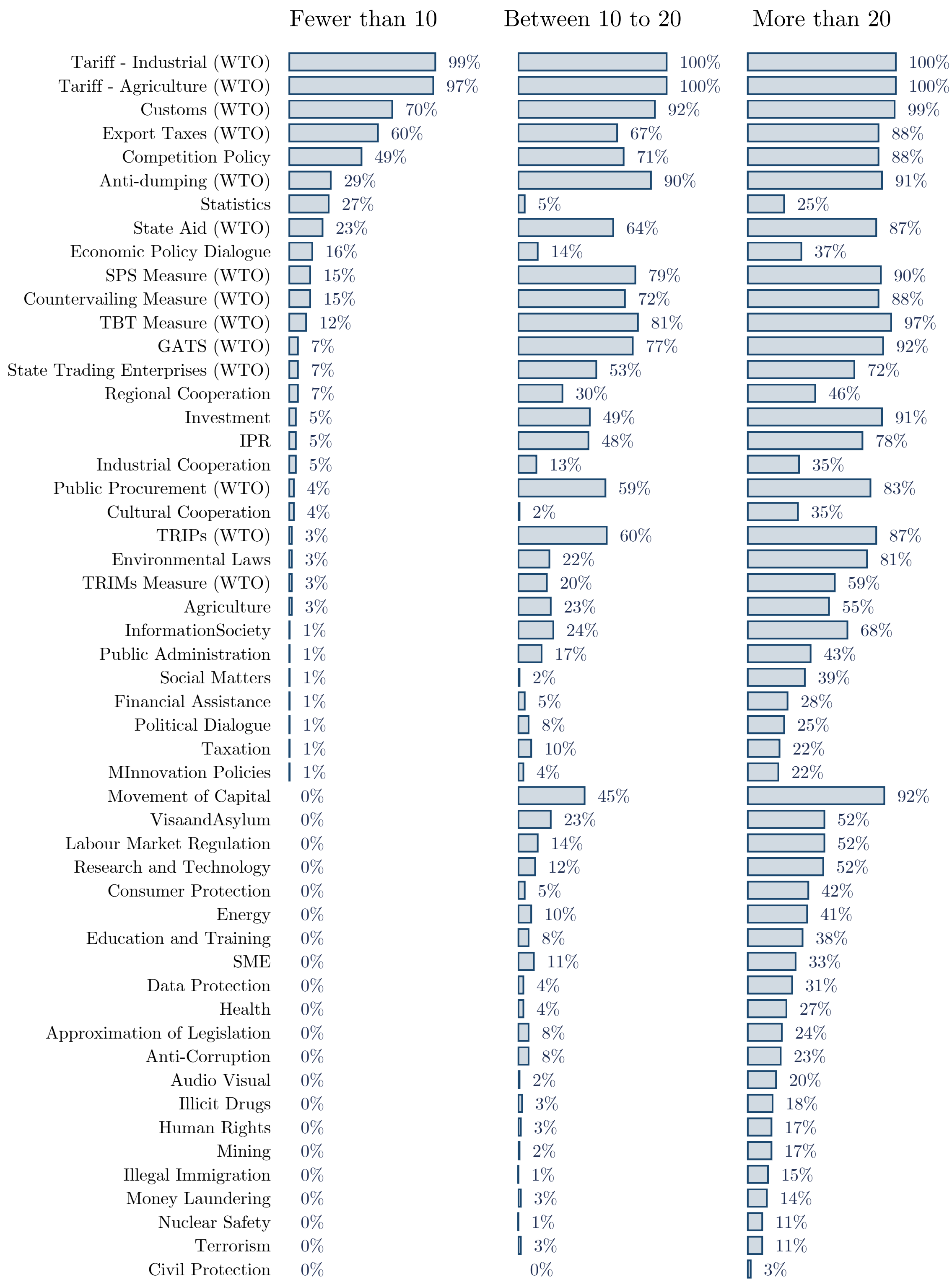

Figure 2: Share of Policy Areas for PTAs

Notes: The figure shows the share of policy areas for different PTAs. The policy areas defined by Hofmann et al. (2017) are based on the most frequently appeared provisions in active PTAs from 1958 to 2017. The policy areas under the current mandate of WTO are indicated in parentheses.

The later PTAs expand to the policy areas beyond WTO, such as intellectual property rights, environmental law, agricultural R&D, or export measures. Recent features related to investment, competition, and intellectual property appear to extend to agriculture. For IP provisions, agriculture is never explicitly excluded–on the contrary, agriculture-related provisions appear to have increased in recent years (Thompson-Lipponen and Greenville, 2019). Most agreements entering into force in the last decade provide for the protection of plant varieties or refer to GI. Of the latter, the majority take the form of commitments to protect specified GI. Cheeses are among the most controversial GIs in trade because cheese GIs such as *Feta, Asiago*, or *Fontina* are typically considered generic types of cheeses outside of the EU. Several studies have discussed its impact within and outside the EU member states, but the effects are still ambiguous (Moschini et al., 2008; Huysmans and Swinnen, 2019; Duvaleix et al., 2021; Curzi and Huysmans, 2022).

Environmental protection appears in a few modern PTAs as agricultural components. The agricultural reference in the environment chapter appears relatively narrow in most agreements. However, in some regional agreements, such as Common Market for Eastern and Southern Africa (COMESA), parties pledge to take measures to control air and water pollution from agricultural activities, encourage the manufacture and use of biodegradable pesticides and herbicides, and discourage the excessive use of agricultural chemicals and fertilizers.

Roughly one-third of active PTAs include provisions that explicitly refer to agricultural research, development, and training (see Figure 3). Given that agriculture may also be implicitly included in broader provisions for cooperation on R&D and training, the share of explicit references to agriculture is not insignificant and may reflect the importance of international cooperation in this area. Examples found in the agreements include references to cooperation on vocational training and research in the context of the standard agricultural policies of the agreements (e.g., EEC Treaty, Economic Community of West African States, and Caribbean Community and Common Market). In some cases, particular provisions are included. For example, COMESA provides for establishing data banks and journals to disseminate agricultural research and extension information.

Finally, concerning export restrictions and other export measures such as taxes and duties, Fulponi et al. (2011) found that most PTAs include chapters prohibiting quantitative export restrictions,

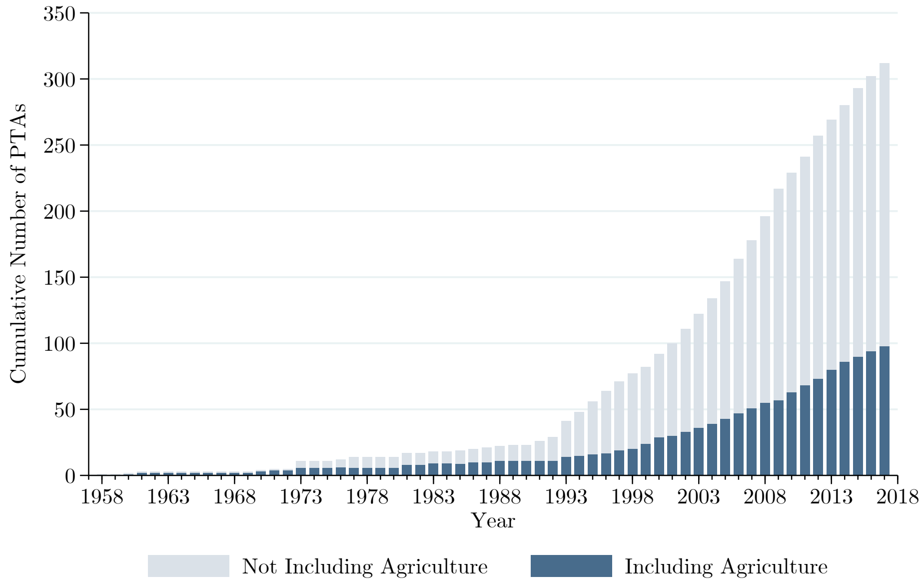

Figure 3: Number of PTAs Containing Agricultural Components

Notes: The figure shows the presence of agricultural components in enforced PTAs over year. The agricultural component defined by Hofmann et al. (2017) is technical assistance to conduct agricultural modernization projects or the exchange of information explicitly stated for agriculture.

except for reasons falling under Article XI of GATT. However, a few agreements exempt certain agricultural products from the prohibition of export restrictions. Moreover, in some agreements, like EFTA-Turkey, Egypt-Turkey, and EU-Turkey, export restrictions are permitted if severe shortages of an essential product arise, subject to certain conditions. Finally, few agreements were found to go beyond WTO rules by prohibiting export duties, even though these can be as detrimental to the functioning of markets as quantitative restrictions (Thompson-Lipponen and Greenville, 2019).

3. Methods

3.1 The Structural Gravity Model

The typical trade models generate relationships in bilateral trade flows as follows. For each exporter i and importer j, the bilateral trade flows Y_{ij}:

$$Y_{ij} = \omega_{ij}^{-\theta} * \frac{Q_i}{\rho_i^{-\theta}} * \frac{E_j}{\pi_j^{-\theta}}, \tag{1}$$

where Q_i represents the total output in countries i and E_j represents the total expenditure in country j. ω_{ij} captures the trade costs between i and j, where θ reflects the elasticity of trade flows to the costs. This elasticity may have different structural interpretations as it appears again with the multilateral resistance terms ρ_i and π_j (Anderson and Van Wincoop, 2003). These two resistance terms can be defined under the structural gravity system: $\rho_i^{-\theta} = \left[\sum_{j=1}^{N} \frac{E_j \omega_{ij}^{-\theta}}{\pi_j^{-\theta}}\right]$ and $\pi_j^{-\theta} = \left[\sum_{i=1}^{N} \frac{Y_i \omega_{ij}^{-\theta}}{\rho_i^{-\theta}}\right]$. Given output Q_i , expenditures E_j and trade costs ω_{ij}, the solution in ρ_i and π_j to this system of two equations is unique (Anderson and Van Wincoop, 2003). As demonstrated by Anderson and Van Wincoop (2003), the gravity system holds separately for each sector k under the separability property of structural gravity (Anderson and Yotov, 2010). Thus, the system of equations can contain sector k as the following:

$$Y_{ijk} = \omega_{ijk}^{-\theta} * \frac{Q_{ik}}{\rho_{ik}^{-\theta}} * \frac{E_{jk}}{\pi_{jk}^{-\theta}}, \quad \rho_{ik}^{-\theta} = \left[\sum_{j=1}^{N} \frac{E_{jk}\omega_{ijk}^{-\theta}}{\pi_{jk}^{-\theta}}\right], and \quad \pi_{jk}^{-\theta} = \left[\sum_{i=1}^{N} \frac{Y_{ik}\omega_{ijk}^{-\theta}}{\rho_{ik}^{-\theta}}\right]. \tag{2}$$

This system of equations is consistent with models based on Anderson (1979) and Krugman (1980) with a constant elasticity of substitution in consumer preferences (Redding and Venables, 2004). In these models, $\theta + 1$ corresponds to the elasticity of substitution. Ricardian models of trade are also fully consistent with gravity (Eaton and Kortum, 2002). In this model, θ corresponds to one of the coefficients of the Frechet distribution of productivity across product varieties. Models based on Melitz (2003) can also generate gravity equations, where the equivalent of θ would be the coefficient of the Pareto distribution of firm productivity (Chaney, 2008). In all of the previously mentioned models, the inward multilateral resistance term π_{jk} can be expressed as a function of the price index of industry k in the importing country j. In turn, ρ_{ik} captures the degree of competition faced by exporter i in industry k (Fally, 2015).

My research focuses on the PTA provision's effects on trade flows. With the general belief of PTA effects on trade flows, each provision is set to lower or raise the trade costs, meaning ω_{ijk} of Equation 2 to be affected. To estimate such changes correctly, the inclusion of intra-national trade data along with bilateral trade flows is desirable for several reasons. Intuitively, it ensures consistency with gravity theory, where consumers choose among and consume domestic as well as foreign varieties (Yotov et al., 2016). Thus, it enables to capture the effects of globalization on international trade and to correct for biases in the estimation of the impact of PTAs on trade (Bergstrand et al., 2015). To complement Breinlich et al. (2022)'s discussions on the gravity-consistent empirical model, I use the domestic trade data in the analysis (Borchert et al., 2021).[5]

3.2 Empirical Model

This section discusses how I set the empirical model to identify the particular agreement provisions that affect agricultural trade flows most. The multiplicative gravity model that explains the relationship between the trade flows y_{ijk} and the vector of PTA provisions x'_{ijk} is as follows:

$$\mu_{ijkt} := \mathrm{E}\left(y_{ijkt}|x'_{ijkt}, \alpha_{ikt}, \gamma_{jkt}, \delta_{ijk}\right) = \exp\left(x'_{ijkt}\beta'_k + \alpha_{ikt} + \gamma_{jkt} + \delta_{ijk}\right) , \tag{3}$$

where i, j, k, and t index exporter, importer, sector, and year, respectively. The trade variable y_{ijkt} include both international and intra-national trade. I lag the covariates x_{ijkt} by one period (t-1) for estimations but keep the notation for simplicity. $\alpha_{ikt}, \gamma_{jkt}$, and δ_{ijk} are sector-level analogs of the fixed effects from the parsimonious three-way gravity model. They account for the multilateral trade resistances effects and the unobserved trade costs to vary by sector.[6] According to the structural gravity theory, bilateral expenditures across countries are separable at the sectoral level from output and expenditure at the country level (Larch and Yotov, 2016). This separability means the multilateral resistances to be sector-specific (Yotov et al., 2016). Thus, the inclusion of the triplet fixed effects may account for all possible omitted variable bias and heterogeneity across the two sectors (Weidner and Zylkin, 2021). Although various measures under PTA provisions do not target particular products, I assume the policy variable of interest (x_{ijkt}) is sector-specific because some provisions are more associated with a specific group of products, such as SPS measures on agriculture products(Disdier et al., 2008). Therefore, I let the coefficient β vary by k so that the PTA effects between agriculture, food, and manufacturing are estimated separately.

[5] The intra-national trade data is absent from Breinlich et al. (2022), which might not be free from the critics whether their approach is theoretically-consistent.

[6] Anderson and Van Wincoop (2003) call multilateral resistance, a theoretical measure of each country's connectedness to the overall trade network. The country-pair fixed effects was suggested by Baier and Bergstrand (2007), who pointed out the potential estimation bias due to omitted cross-sectional heterogeneity, motivated as coming from unobserved trade costs.

I estimate Equation 3 using PPML. Silva and Tenreyro (2006) pointed out the ordinary least squares (OLS) estimator with log transformation is inconsistent without strict assumptions and cannot deal with zero trade flows, which are a common feature of agricultural trade data. Given the exponential mean form, the PPML estimator allows the research to include zero trade in the estimation that accounts for the potential bias due to censoring the observations (Silva and Tenreyro, 2006). In addition, I use a modified version of the iteratively re-weighted least-squares (IRLS) algorithm that is robust to statistical separation and convergence issues arise from many high-dimensional fixed effects (Correia et al., 2020).[7] This recent computational innovation allows the research to estimate the three-way fixed effects model with the promise of the consistency and asymptotic distribution of the three-way PPML estimator (Weidner and Zylkin, 2021).

3.3 The Lasso Regression

The typical econometric methods cannot estimate Equation 3 if the number of covariates x'_{ijkt} is significantly large. The number of provisions in my data is 378, of which only a subset x' will have a non-zero effect on trade flows. However, I do not know x' beforehand, so a proper statistical technique is needed to find x'. I adopt a regularization approach suggested by Athey and Imbens (2017) and Athey and Imbens (2019) that involves appending a penalization term to the PPML model to purge PTA variables with coefficients equal to zero. The minimization problem that defines the sectoral three-way Lasso PPML system is therefore given by:

$$\left(\hat{\alpha}, \hat{\gamma}, \hat{\delta}, \hat{\beta}\right) := \arg \min_{\alpha,\gamma,\delta,\beta} \frac{1}{n}\left(\sum_{i,j,k,t} \left(\mu_{ijkt} - y_{ijkt} \ln \mu_{ijkt}\right)\right) + \frac{1}{n}\sum_{l=1}^{m} \hat{\phi}_{kl}\lambda_k|\beta_{kl}|\,, \tag{4}$$

where n is the number of observations. Note that I leave out k dimension for simplicity in the following discussions. The first part of the Equation 4 reflects the standard PPML minimization problem using pseudo-likelihood function. The second part is the Lasso term which consists of two tuning parameters, $\hat{\phi}_l \geq 0$ and $\lambda \geq 0$. I include diagonal matrix $\hat{\phi}_l$ to account for regressor-specific

[7] Correia et al. (2020) developed an algorithm for estimating nonlinear fixed effect (FE) models based on IRLS that exploits the linearity of the weighted least squares step in the IRLS algorithm to eliminate the FEs in each iteration. It then relies on the Frisch-Waugh-Lovell theorem to update the weights, which is repeated until convergence is achieved.

penalty weights in addition to the standard Lasso penalty λ (Belloni et al., 2016). Along with the standard tuning parameter λ, $\hat{\phi}_l$ refines the model iteratively across PTA variables. Larger penalties shrink the β until zero, so it identifies non-zero PTA variables to be included in the final model.[8] The fixed effects parameters are not penalized because I want to include them in every iteration to keep the structural gravity framework. In other words, for any given β, α, γ, and δ are obtained by solving the standard PPML regression.[9]

There are two widely suggested approaches to determine the tuning parameters λ and $\hat{\phi}_l$. One is the more traditional "cross-validation" approach, whose strength comes from the predictive performance. Cross-validation is a re-sampling procedure that repeatedly holds out a subset of the data and chooses λ to maximize the model's predictive fit. However, $\hat{\phi}_l$ does not play a role here but is set to be 1. There are two major concerns about this approach. First, due to the nature of this method, it may select too many irrelevant variables. Second, as it ignores the regressor-specific penalty $\hat{\phi}_l$, it does not take into account any heteroskedasticity and within-cluster correlation featured in the data. Next is the "plug-in" approach that specifies appropriate functional forms for the penalty parameters based on statistical theory and uses estimates for these parameters (Ahrens et al., 2020). It is more theory-driven than the cross-validation approach because the researcher must choose the number of sub-samples when conducting cross-validation. The main advantage of this approach is that it is very parsimonious in selecting the variables, so it offers superior performance versus cross-validation approaches in finite samples (Breinlich et al., 2022). Thus, the "post-Lasso" estimates, or the typical PPML estimates using the selected variables through the plug-in approach, have a "near-oracle" property that ensures it captures the correct model if the sample is sufficiently large Belloni et al. (2012). This advantage is critical to answering my research question, so I adopt the plug-in method.

The plug-in Lasso also has several technical advantages over other methods. The Lasso approach finds the best fitted model by selecting the correct β_l. To do so, it updates the fit score by a small

[8] In practice, if $\lambda \to \infty$, the system selects none of the PTA variables as it is the only way to solve this minimization equation. On the other hand, if $\lambda = 0$, it is identical to solving the standard PPML under the entire system.

[9] This means I use the original high-dimensional fixed effects IRLS algorithm of Correia at al. 2020 to compute the fixed effects but replace the weighted linear regression step with the suggested Lasso approach.

change in β_l only if the improvement in fit is large relative to the penalty. The plug-in Lasso's improvement lies in the regressor-specific penalty $\hat{\phi}_l$ to adjust to reflect the standard error of the score. Including this penalty prevents a case where I select wrong regressors due to estimation noise, so it plays a critical role in taking account of heteroskedasticity in trade data. Thus, the values for λ and $\hat{\phi}_l$ must be set high enough so that the value of the score for β_l becomes large relative to its standard error for regressor $x_{ijt,l}$ to be selected.[10]

Due to the strict rule in the plug-in approach, potential weakness comes from selecting only a few variables. A problem may arise when there is a substantial number of highly correlated variables because the plug-in algorithm may penalize the correlated variables of a selected variable. Thus, the plug-in Lasso may wrongly penalize relevant provisions because the PTA provisions in my data are correlated if they have similar legal contexts. Because this paper aims to identify all relevant PTA provisions that affect agricultural trade, I complement the shortcomings by performing a second-stage analysis regressing the selected estimator on all excluded provisions. Thus, it identifies bundles of provisions highly correlated with the ones selected in the first stage.[11] These auxiliary regressions aim to construct bundles of provisions, at least if combined, likely to have a causal impact on agricultural trade when included in trade agreements (Breinlich et al., 2022).

4. Data

The analysis rely on the novel data set combining the new database containing the international and intra-national trade from Borchert et al. (2021) with the data on the content of PTAs collected by Mattoo et al. (2020). First, I obtained agricultural export data from 1991 to 2017 for 249 exporters and 251 importers. This includes intra-national trade, the difference between the value of total production and total exports. The output and trade for all industries in Agriculture, besides Forestry and Fishing, are from the Food and Agriculture Organization of the United Nations Statistics Division (FAOSTAT). I aggregated 28 sub-agricultural product trade, considering zero

[10] I cluster all observations belonging to directional country pairs to take account of the potential serial correlation of errors and correlation within each pair when I calculate standard errors. I use the same cluster Lasso approach as in Belloni et al. (2016).

[11] Because the selected variables from the previous stage may be the tip of the iceberg of a bundle of variables that affect agricultural trade, this additional regression may help to identify the rest of the iceberg. As such, Breinlich et al. (2022) called this second stage an 'iceberg Lasso.'

values from the source as true zero trade between the country pairs. I also obtained food and non-food manufacturing trade data from the same database to observe sectoral differences.[12]

Table 1: Distribution of Provisions by Policy Area

Policy Area	WTO	Number of Provisions	Essential Provisions	Share
Anti-dumping (AD)	✓	39	9	23%
Competition Policy (CP)		35	33	94%
Countervailing Duties (CVD)	✓	14	8	57%
Environmental Laws (ENV)		49	20	41%
Export Taxes (ET)	✓	45	21	47%
Investment (INV)		57	14	25%
Intellectual Property Rights (IPR)		120	8	7%
Labor Market Regulations (LM)		18	7	39%
Migration (MIG)		30	17	57%
Movement of Capital (MOC)		94	43	46%
Public Procurement (PP)		100	24	24%
Rules of Origin (ROR)		38	27	71%
Services (SER)		64	38	59%
Sanitary and Phytosanitary (SPS)	✓	59	17	29%
State-Owned Enterprises (STE)	✓	53	19	36%
Subsidies (SUB)	✓	36	24	67%
Technical Barriers to Trade (TBT)	✓	34	17	50%
Trade Facilitation (TF)		52	32	62%
Total		937	378	40%

Notes. This table lists the PTAs' 18 policy areas covered for the analysis. The second column indicates whether WTO commitment has influenced the policy area. The following columns report the number of provisions by policy area from Mattoo et al. (2020) and the essential provision used for actual analysis.

The data on the content of PTAs includes information on 282 out of 317 active PTAs between 1988 and 2017. Therefore, I excluded the 35 unmapped PTAs and the applicable country pairs from the Lasso regression.[13] The data focus on the 18 policy areas that are most frequently covered in trade agreements, compared to the 52 policy areas of Hofmann et al. (2017).[14] These policy areas only include NTBs, ranging from those WTO extensions to beyond the WTO commitments. Table 1 shows detailed provisions covering stated objectives under each policy area, a total of 937 provisions. However, I use the most frequently appeared 378 provisions for the analysis to alleviate

[12] The food manufacturing export data are from 251 exporters to 251 importers from 1988 to 2017, and the non-food manufacturing export data are from 252 exporters and 252 importers for the same period.

[13] I keep all information on the presence of PTA for the post-Lasso and secondary analyses using dummy and dummy interaction models.

[14] These policy areas are equal to or above the 20 percent share of the trade agreements mapped by Hofmann et al. (2017).

calculation difficulties caused by high dimensionality.[15]

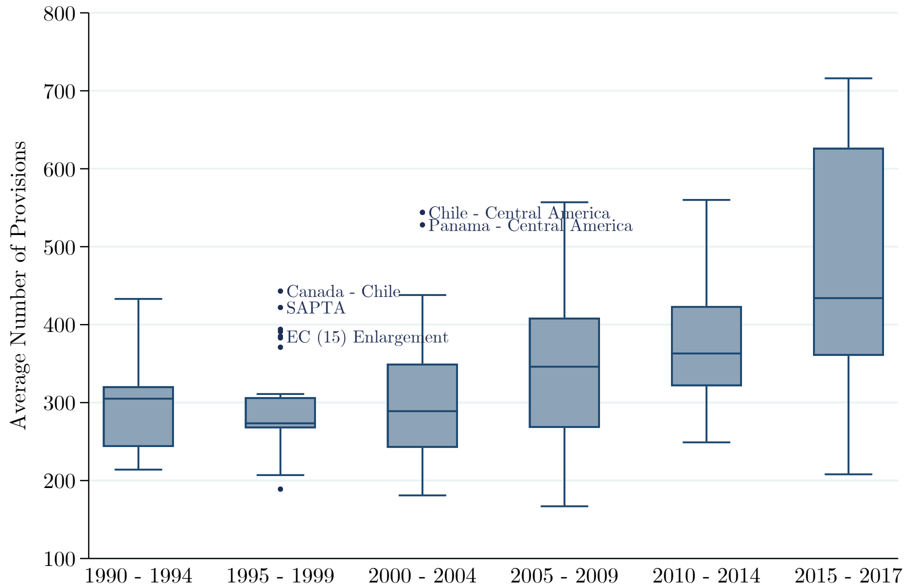

Figure 4: Average Number of Provisions in PTAs

Notes: This boxplot shows the range and average of the number of provisions of PTAs enforced from 1990 to 2019.

The number of PTA provisions has grown steadily since the mid-1990. Figure 4 shows that average number of provisions appear to have dropped from the early-1990 to the mid-1990, but regained its momentum to grow since then. The average number of provisions was grown from 274 to 440 in the last two decades. The variation of the provision coverage across the enforced PTAs has also enlarged. Excluding a few exceptions at extreme, the range of the number of PTA provisions enforced from 1995 to 1999 is about 100.[16] However, the range increased to 520 for the PTAs enforced after 2015. This figure suggests that each PTA differs regarding characteristics and what NTB measures have been offered, implying a high possibility of heterogeneous PTA effects.

I classify countries by income level to distinguish trade flows between developed and developing countries. Using the World Bank's classification, I grouped countries into North and South. I

[15] Breinlich et al. (2022) classified essential and non-essential provisions based on experts' knowledge, but I rely on the frequency level of appearance in all agreements. The threshold is 0.093 where the average frequency level splits.

[16] Those extremes are mostly regional integration or bilateral agreements with the formed regional integration members, such as EC Enlargement in 1995, South Asian Preferential Trade Agreement (SAPTA) in 1995, or Chile - Central America 2002.

treated upper-middle and high-income countries as 'North' and lower-middle and low-income countries as 'South'. Figure 5 shows that the share of agricultural exports between PTA partners has gradually increased for 'North-South' and 'South-North' pairs while it has decreased for 'North-North' and 'South-South' pairs. This dissimilar trend across each regional pair suggests a need for an estimation strategy to consider potential differential effects.

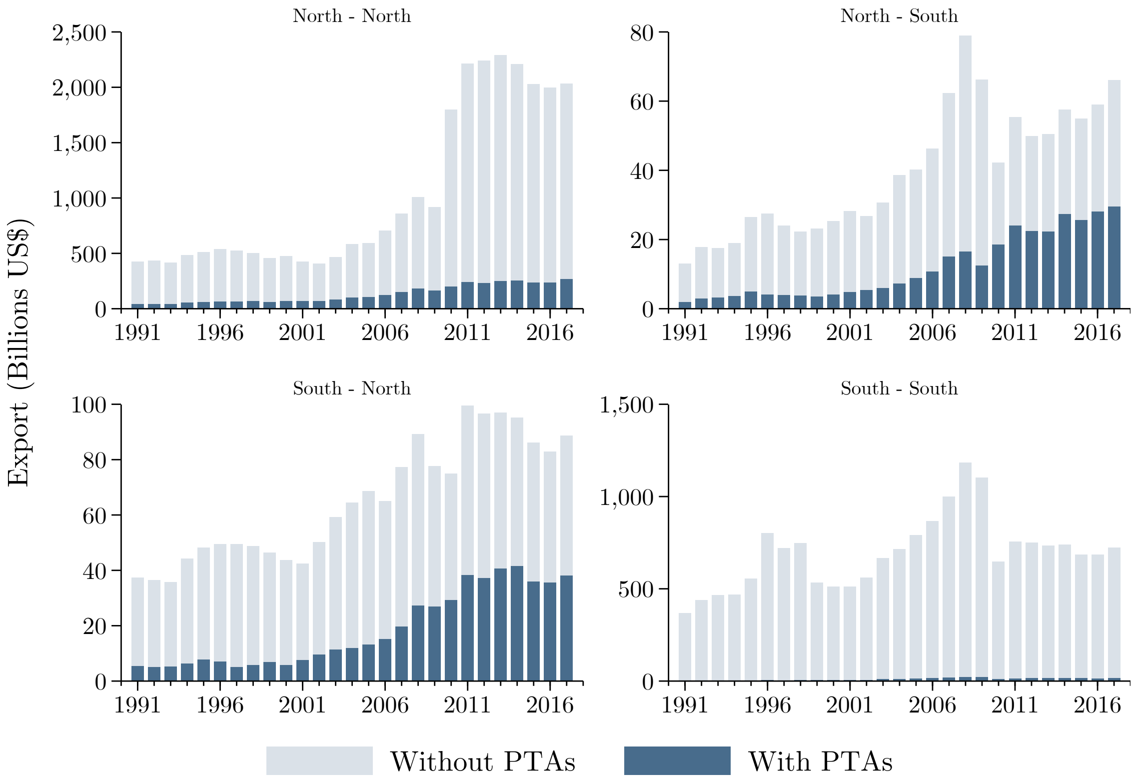

Figure 5: Agricultural Trade across High and Low-income Countries

Notes: The figure shows the agricultural trade from 1991 to 2017 across region pair classified by income level. I use the world bank's classification of high (including upper middle) and low-income (including lower middle) countries to distinguish North and South countries, respectively. The highlighted portion represents the amount of agricultural trade between PTA partners.

5. Results

5.1 Effective PTA Provisions on Agriculture Trade

Table 2 shows the main results of the plug-in Lasso and post-Lasso regressions. In column (1), I start by presenting the results of a PPML structural gravity estimation with a dummy of the presence of PTAs between the trading partners. The results replicates earlier findings that enforcing

PTA promotes agricultural trade. PTAs increased agricultural trade by 27.1 percent.[17] Column (2) presents the plug-in Lasso regression results, showing the selected provisions and their coefficients found to be non-zero. Then, column (3) shows the post-Lasso regression results using the selected provisions. Finally, column (4) presents the estimated results of the comprehensive model, adding a dummy of the presence of PTAs to the regression model of (3).

Table 2: The Selected PTA Provisions for Agricultural Trade

	PPML (1)	Lasso (2)	Post-Lasso (3)	PPML (4)
PTA - Presence of PTA	0.240***			−0.0218
	(0.036)			(0.028)
ENV.19 - Prevent Pollution by Ships		0.052	−0.048	−0.040
			(0.042)	(0.045)
ET.06 - Require Scheduling of EQ/QR		0.014	0.176***	0.174***
			(0.061)	(0.062)
SPS.44 - Include SPS Chapter		0.052	0.079***	0.096***
			(0.018)	(0.027)
STE.39 - Include Sector-specific Discipline on STE		0.107	0.162***	0.175***
			(0.056)	(0.060)
SUB.16 - Require for Prior Notification on Subsidies		0.004	0.189***	0.191***
			(0.030)	(0.030)
TBT.06/33 - Use Regional Standards		0.091	0.197**	0.191**
			(0.095)	(0.096)
TF.32 - Mutual Recognition of AOs		0.251	0.451**	0.448**
			(0.058)	(0.058)
Observations	798,239	786,821	786,821	786,821
Pseudo R-squared	0.994		0.994	0.994

Notes. This table presents the plug-in Lasso regression results and the post-estimation results for agriculture. The clustered robust standard errors by directional country pair are in the parenthesis for PPML and Post-Lasso regressions.

For agricultural trade, the plug-in Lasso selected seven provisions related to policy areas in Environmental Law (ENV), Export Taxes (ET), SPS, State Trading Enterprise (STE), Subsidies (SUB), TBT, and Trade Facilitation (TF). There are intuitive explanations for most of these variables' potential effects on agricultural trade. For example, the environmental law might increase the cost of transport, in particular, as ENV.19 aims to prevent pollution by shipping. Thus, this provision generally has adverse effects on trade. Although the post-Lasso results show that the estimated coefficient is not statistically significant, its sign is reasonable.

[17] The estimated coefficient b in semi-elasticity form can be transformed into the elasticity form by $\exp(b) - 1 = \exp(0.24) - 1 = 0.271$.

Other selected provisions are estimated to affect agricultural trade positively. Most PTAs contain non-discrimination provisions that can promote competitive neutrality, consequently, create an accessible border. For example, preventing subsidies and the intervention of state enterprises (STEs) may promote non-discriminative environments. Unlike other sectors, WTO commitments do not require eliminating agricultural subsidies, so the inclusion of such provisions in PTAs critically affects agricultural trade patterns between members. The inclusion of STE.39, which includes sector-specific discipline on STEs, can define the clear role of STE in each country, aiming to restrict their intervention (Willemyns, 2016). SUB.16 require prior notification on any measures related to subsidies. This system is designed to prevent unexpected subsidies that might distort the market (Kowalski et al., 2013). I found that STE.39 and SUB.16 increased agricultural trade by 17.6% and 20.8%, respectively.

The trade facilitation policy may not be agricultural-specific, but it helps to promote trade flows by supporting and arranging different local standards and regulations in international legal terms. TF.32 states the mutual recognition of authorized operators that can facilitate trade by allowing certified traders and supply chain operators would be given a smoother passage through customs procedures (Sá Porto et al., 2015). TF.32 increases agricultural trade by 57.0% in my data. The export taxes policy area incorporates GATT Article XI prohibiting any quantitative restrictions between the parties. In particular, ET.06, under the chapter that prohibits all export quotas, requires both parties to schedule mutually exclusive export quotas. Thus, imposing this regulation prevents unnecessary trade distortions. The results show that ET.06 increased agricultural trade by 19.2%.

I found that including SPS chapters in PTA increased agricultural trade by 8.2%, and allowing regional standards increased agricultural trade by 21.8%.[18] The trade promotional effects of SPS and TBT measures can be controversial because these measures have been believed to have trade-impeding effects (Peterson et al., 2013; Murina and Nicita, 2017). However, there's another view that PTAs facilitate trade by signaling the imported goods' quality to domestic consumers once

[18] I grouped TBT.06 and TBT.33 as one variable because they are identical in their description. In addition, they are almost perfectly collinear, where the correlation between the two variables is 0.98, and TBT.06 is only a subset of TBT.33.

passed the regional standards (Disdier et al., 2008). My findings support the second idea, but there is a need to investigate if these effects are consistent over different countries by income levels because technical and financial capacity determines the adaptability constraints toward such standards.

In column (4), I re-estimated the model using the same covariates as column (3) but now re-adding my original PTA dummy from column (1). In this case, the coefficient on PTA captures any effect on trade flows that is not already captured by the provision variables that the Lasso regression selected. With this in mind, I take the insignificant and near-zero coefficient on PTA in column (4) as an encouraging indication that the selected provisions thoroughly explain the average PTA effect reported in column (1).

5.2 Comparison with Manufacturing Trade

To investigate sectoral differences, I conduct the same analyses for the food and non-food manufacturing sectors. Table 3 shows the results of the plug-in Lasso and post-Lasso regressions, where each panel presents the outcomes of individual sectors. In column (1), the results show that the presence of PTA increased the food manufacturing trade by 20.8%. The estimated coefficient is lower than agriculture, but the PTA's trade promotional effects are solid. The plug-in Lasso selected a similar set of provisions for the food manufacturing trade to that of agriculture. It selects ENV.19, ET.06, SPS.44, TBT06/33, TF.32, and two new provisions under Competition Policy and Intellectual Property Rights, but excludes SUB.16 and STE.39.

Similar to the Subsidies (SUB) and State-owned Enterprises (STE) areas I covered above section, Competition Policy (CP) can promote trade by prohibiting anti-competitive action. In particular, CP.31 requires each party to establish regional and agreement-related competition authority to handle conflicts and complement international settlement authorities, if needed. I found that CP.31 increased the food manufacturing trade by 11.0%. I also found that IPR.43 increased trade flows by 22.5%. This provision designates the list of GI products, meaning that the legal protection of GI products increased food trade in total value. This result adds to the recent discussion on the economic consequences of rising GI-related regulations. My findings only suggest the aggregated amount of agricultural trading has increased, so it is unclear whether the effects come from the increment of GI products trade or the deflection of other products. In addition, the effects might

Table 3: The Selected PTA Provisions for Manufacturing Trade

	PPML (1)	Lasso (2)	Post-Lasso (3)	PPML (4)
Panel A. Food Manufacturing				
PTA - Presence of PTA	0.189***			−0.011
	(0.036)			(0.023)
CP.31 - Create Regional/Agreement-specific Authority		0.058	0.104*	0.100
			(0.062)	(0.062)
ENV.19 - Prevent Pollution by Ships		0.018	0.052	0.056
			(0.034)	(0.036)
ET.06 - Require Scheduling of EQ/QRs		0.131	0.198***	0.201***
			(0.047)	(0.048)
IPR.43 - Designate the List of GI Products		0.012	0.203***	0.195***
			(0.019)	(0.023)
SPS.44 - Include SPS Chapter		0.014	0.111***	0.119***
			(0.016)	(0.023)
TBT.06/33 - Use Regional Standards		0.438	0.397***	0.398***
			(0.057)	(0.057)
TF.32 - Mutual Recognition of AOs		0.069	0.239***	0.241***
			(0.042)	(0.043)
Observations	957,886	946,580	946,580	946,580
Pseudo R-squared	0.996		0.996	0.996
Panel B. Non-food Manufacturing				
PTA - Presence of PTA	0.109			0.013
	(0.077)			(0.091)
CP.23 - Promote Transparency		0.031	0.294***	0.287***
			(0.055)	(0.075)
ET.02 - Prohibits all EQ/QRs		0.008	0.188***	0.184***
			(0.051)	(0.057)
Observations	1,217,498	1,206,138	1,206,138	1,206,138
Pseudo R-squared	0.991		0.991	0.991

Notes. Each panel presents the separate estimation results for food and non-food manufacturing. The clustered robust standard errors by directional country pair are in the parenthesis for PPML and Post-Lasso regressions.

be limited to a particular area, such as the EU, because only European countries have insisted on protecting GI products, while other countries oppose the expansion of EU GI protection (Curzi and Huysmans, 2022). Thus, I explore its regional effects in the later section.

The results are significantly different in the non-food manufacturing sector. The presence of PTA is still positively associated with non-food manufacturing trade flows, but the estimated coefficient is relatively trivial to other sectors and not statistically significant. The plug-in Lasso regression selected only two provisions for the non-food manufacturing trade. It selects one provision each from Competition Policy and Export Taxes. CP.23 promotes transparency by informing of competition concerns, providing consultation with authority, and publishing the progress and decisions. It increased manufacturing trade by 34.2%. ET.02 states that the parties prohibit all export quotas, which prevents unnecessary trade distortions. The results show that ET.02 increased manufacturing trade by 20.7%.

5.3 The Second-stage Lasso: Extra Bundles of Provisions

The plug-in Lasso selects only a few variables due to the strict rule in its selection algorithm. In particular, a problem may arise when a substantial number of provisions are highly correlated, like in my policy data. In this case, the plug-in Lasso may wrongly penalize relevant variables due to the high correlation with the selected variables. Therefore, I run an additional plug-in Lasso regression that regresses each selected variable on all other provisions excluded from the first stage. These auxiliary second-stage regressions aim to construct bundles of provisions, at least if combined, likely to have a causal impact on agricultural trade when included in trade agreements (Breinlich et al., 2022).

Table 4 presents the second-stage Lasso results for agriculture.[19] Each column presents the second-stage Lasso results of the seven provisions selected in the previous stage, and the second row presents their estimated impact on agricultural trade flows, also available in column (3) of Table 2. The subsequent rows report all provisions not selected in the first stage but identified in the second stage with their correlation with the provision in the first row. The results show that the second-stage regression identifies 35 (=28+7) provisions associated with agricultural trade flows.

In most cases, the second-stage Lasso regression selected highly correlated provisions under the same policy area. For example, column (2) presents two highly correlated provisions under the same policy area, selected as a bundle with ET.06. This selection is not surprising because these provisions target the same objective of prohibiting quantitative restrictions. However, the plug-in Lasso selected three more provisions of other related policy areas. CVD.01 fulfills a similar goal as it disallows the countervailing duties. Moreover, STE.32 and SUB.14 may share a broader goal of

[19] This section discusses the results of the agricultural trade. Please see Table S2 for the results in manufacturing sectors.

prohibiting market distortions as they limit the state-owned enterprises' role and subsidies. The results represent the synergy effects between groups of provisions that aim for one broader outcome. I could observe the same pattern from SPS.44, STE.39, SUB.16, and TF.32.[20]

In other cases, there is no intuitive explanation of how the algorithm selected seemingly unrelated provisions for ENV.19 and TBT.06/33. For example, I found that ENV.19 is highly correlated with two selected provisions under the Competition Policy area: CP.07 and CP.15. Because CP.07 and CP.15 are designed to prohibit anti-competitive behavior between undertakings and abuse of market position by dominant parties. Compared to ENV.19's goal of regulating environmental standards, there seems to be no intuitive explanation for how the three provisions are highly correlated and on a similar dimension to identify agricultural trade. Similarly, there is a lack of strong relationship between TBT.06/33 and ENV.06. These correlations seem to be subject to the template effect, that is, the tendency of essential trading blocs such as the EU and the United States to use similar provisions in all their agreements (Breinlich et al., 2022). For example, PTAs are significantly similar, up to 89% between US-oriented and 80% between EU-oriented agreements (Mattoo et al., 2020).

[20] I do not repeat the explanation for these provisions to avoid redundancy. Instead, I provide Table S1 for the list of all provisions and their descriptions so that readers may match the descriptions to the results of other provisions in Table 4.

Table 4: The Second-stage Lasso Results for Agriculture

(1) ENV.19	(2) ET.06	(3) SPS.44	(4) STE.39	(5) SUB.16	(6) TBT.06/33	(7) TF.32
-4.7%	19.2%	8.2%	17.6%	20.8%	21.8%	57.0%
CP.07 (0.62)	CVD.01 (0.62)	CP.03 (0.71)	STE.27 (0.72)	SUB.26 (0.65)	ENV.06 (0.54)	TBT.18 (0.67)
CP.15 (0.60)	ET.03 (0.55)	CP.04 (0.67)	STE.28 (0.65)	SUB.34 (0.75)		TF.29 (0.93)
	ET.15 (0.61)	ROR.06 (0.81)	STE.46 (0.62)	TBT.22 (0.65)		
	STE.32 (0.58)	SPS.56 (0.62)	SUB.03 (0.67)	TF.11 (0.59)		
	SUB.14 (0.53)	TF.20 (0.60)	SUB.29 (0.88)	TF.38 (0.60)		
		TF.26 (0.67)	TBT.16 (0.66)			
			TBT.21 (0.58)			

Notes. Each column presents the second-stage Lasso results of the selected PTA provisions associated with agricultural trade flows. The estimated percent changes in agricultural trade by the provision selected in the first-stage Lasso are in the second row. The subsequent rows report all provisions not identified at the first-stage but in the second-stage Lasso regressions. The numbers in brackets are correlations with the provisions selected by the first stage.

5.4 Cross-regional Differences in PTA Effects

As Figure 5 shows, most of increasing number of PTAs were between developed and developing countries. Existing studies provide only a little economic rationale for these agreements, but they lump North-South PTAs together with other types of trade pacts. In this section, I discuss the potential differences in PTA effects on agricultural trade between regions with different income levels.

To account for the cross-regional differences, I first modify my baseline model as

$$\mu_{ijct} := \mathrm{E}\left(y_{ijct} | x'_{ijct}, \alpha_{ict}, \gamma_{jct}, \delta_{ijc}\right) = \exp\left(x'_{ijct}\beta' + \alpha_{ict} + \gamma_{jct} + \delta_{ijc}\right) \ , \tag{5}$$

where c identifies the directional region pair, either North-North, North-South, South-North, or South-South. All other specifications are the same as Equation 3 but excluded the industry (k) dimension because this section studies agricultural trade only. Instead, the model includes new set of fixed effects, α_{ict}, γ_{jct}, and δ_{ijc}, now interacting with directional region pairs (c) to capture unobserved pair-specific multilateral resistances. Therefore, this model accounts for any heterogeneity between the region pairs.

I re-estimated the plug-in Lasso regression using the modified model with region pair fixed effects. The results in Table 5 show that only four provisions are selected to be associated with agricultural trade. ET.06 and SPS.44 appear again, but the estimates are more significant. As previously explained, these two provisions have promotional effects on trade by prohibiting export quotas and signaling the quality and safety of imported goods. The newly selected provisions, CP.31 and IPR.43, also have trade promotional effects. CP.31 requires to creation of regional/agreement-specific authority, so it expedites handling conflicts at the border, consequently, facilitates trade. IPR.43 is about designating the list of GI products. This provision acts similarly to SPS.44 in that it can reduce information asymmetry between producers and consumers, thus signaling imported goods' quality and potentially enhance trade (Sorgho and Larue, 2014).

Although I controlled potential heterogeneity across region pairs and obtained robust estimates for the average effects, the questions remain about how PTAs affect agricultural trade differently in each

Table 5: The Plug-in Lasso Results: Adding Cross-regional Fixed Effects

	PPML (1)	Lasso (2)	Post-Lasso (3)	PPML (4)
PTA - Presence of PTA	0.185***			0.034
	(0.049)			(0.028)
CP.31 - Create Regional/Agreement-specific Authority		0.100	0.434***	0.422***
			(0.067)	(0.067)
ET.06 - Require Scheduling of EQ/QR		0.095	0.329***	0.316***
			(0.064)	(0.064)
IPR.43 - Designate the List of GI Products		0.054	0.074***	0.098***
			(0.029)	(0.026)
SPS.44 - Include SPS Chapter		0.047	0.154***	0.126***
			(0.026)	(0.026)
Observations	711,609	700,106	700,106	700,106
Pseudo R-squared	0.995		0.996	0.996

Notes. This table presents the plug-in Lasso regression results and the post-estimation results after adding cross regional fixed effects. The clustered robust standard errors by directional country pair are in the parenthesis.

region pair. Thus, I also estimated β for each region-pair (c) and provided the individual regression results in Table 6. First, I start by presenting the estimation results of the simple dummy model indicating the presence of PTA in Panel A. The results show that the PTAs promoted agricultural exports from similar-income partners but not from partners with gaps. These results are surprising because the recent PTA developments are between North and South countries.

In Panel B, I provide the estimates for the effects of the four selected provisions. The selected PTA provisions generally increased trade between countries with similar income levels and decreased trade flows between countries with income gaps. For instance, SPS.44 positively affects North-North and South-South agricultural trade by 16.4% and 52.5%, respectively, but negatively in North-South trade by 8.5%. I also found its negative impact on the South-North trade, but it is not statistically significant. These results contradict Disdier et al. (2008); Murina and Nicita (2017)'s claim that developing countries may suffer more from the SPS measures due to the adaptability constraints by low technical and financial capacity. They also claim that only developed countries can benefit from implementing SPS standards because they may promote exports once they pass the importer's regional standards, which also contradicts those results.

Table 6: The PTA Effects on Agricultural Trade across Regions

	North-North (1)	North-South (2)	South-North (3)	South-South (4)
Panel A.				
PTA - Presence of PTA	0.230***	0.004	0.017	0.377***
	(0.035)	(0.040)	(0.088)	(0.049)
Observations	185,177	162,281	187,875	176,276
Pseudo R-squared	0.996	0.948	0.950	0.998
Panel B.				
CP.31 - Create Regional/Agreement-specific Authority	0.854***	0.005	−0.148	0.386***
	(0.164)	(0.157)	(0.103)	(0.130)
ET.06 - Require Scheduling of EQ/QR	0.052	0.245	−0.123	1.206***
	(0.149)	(0.190)	(0.085)	(0.407)
IPR.43 - Designate the List of GI Products	0.069*	−0.074	0.052	
	(0.041)	(0.086)	(0.066)	
SPS.44 - Include SPS Chapter	0.164***	−0.085*	−0.020	0.525***
	(0.043)	(0.046)	(0.026)	(0.037)
Observations	179,745	159,957	185,609	174,795
Pseudo R-squared	0.996	0.949	0.950	0.998

Notes. This table compares the differential effects of PTA provisions across regions, classified by income levels. Each column presents the post-Lasso regression results for each region pair using the selected provisions from Table 5. The clustered robust standard errors by directional country pair are in the parenthesis.

One extra condition may complement their claims to explain my findings. The exporters must have both the 'ability' and 'incentive' to adopt the new standard. Assuming that meeting the SPS standards means meeting the local quality and taste, exporters must expand or transform the production line to differentiate product quality vertically or horizontally. However, if the cost of expanding the production line is high, exporters are less incentivized to differentiate their products. It happens when the targeting quality is far from the home consumer's taste, usually when there are significant income gaps. Thus, exporters have less incentive to meet the standards of countries with significant income differences. This is one reason there is more trade between similar-income countries, and the logic applies to both developed and developing countries.[21] Thus, if SPS measures are applied, exporters adopt the new standard and expand their production line only if their trading partners have similar income levels or tastes, where the cost to adjust their production line is relatively small. Consequently, combined with Disdier et al. (2008)'s theory, the SPS measure can increase North-North or South-South trade but decrease trade between North and South.

Other provisions have similar patterns. For example, CP.31 positively impacts North-North and South-South trade but has no impact on trade between North and South countries. ET.06 only positively impacts South-South trade. Lastly, IPR.43 impacts only North-North trade because EU countries lead GI regulation while other countries oppose it. Note that there is no estimate of IPR.43 for South-South trade because GI-related regulations do not apply to the agreements with developed countries.

5.5 Conditional General Equilibrium Effects

6. Conclusion

As the number of agreements grows significantly and their content becomes deeper, many PTAs contain complicated and interrelated provisions that might affect agricultural trade. An important economics question is whether the PTA promotes or hinders agricultural trade and, if so, what

[21] The home market effects well explain how the economies of scale work in a system that countries with greater demand for some products at home tend to have larger sales of the same products abroad (Hallak, 2010). A key prediction is that richer (poorer) countries with higher demand for high (low) quality goods will specialize in producing and exporting them. Consequently, there is more trade between similar-income countries (Fajgelbaum et al., 2011).

provisions are responsible for this. In this paper, I proposed a machine learning method in the structural gravity framework to find the most effective policy variables (PTA provisions) that affect the agricultural trade flows. The motivation for using this method is due to the limitation of traditional econometrics when estimating a large set of variables. The model selected 8 out of 378 provisions from the essential PTA provisions (column 2, Table 2). I then quantify the selected provisions' effects on agricultural trade using the parsimonious three-way PPML estimator (column 3, Table 2).

The estimation results show that six provisions increased agricultural trade under various policy areas, including Export Taxes, SPS, State Trading Enterprise, Subsidies, TBT, and Trade Facilitation. Some are intuitive and reasonable, but others need a more profound interpretation. For instance, most PTAs contain non-discrimination provisions that can promote competitive neutrality, consequently, create an accessible border. Most provisions under Export Taxes, State Trading Enterprise, and Subsidies are related to this vision. In addition, Trade Facilitation is also an essential chapter of PTAs to support arranging different local standards and regulations in international legal terms. Less intuitive provisions are under policy areas like SPS and TBT. The SPS and TBT measures can facilitate trade by signaling that products are safe for consumers, but these measures can also impede trade in a protectionist manner. My findings support the first idea because including SPS chapters in PTAs and allowing to use of regional standards increased agricultural trade by almost 70% in my data.

The promotional effects of PTA and the above provisions are inconsistent in countries with different income levels. I classified North and South countries by income level, then considered the cross-regional fixed effects in the second model. The model selected four provisions that have promotional effects on agricultural trade (Table 5). A new provision not captured in the first model is about the GI regulations. By designating the list of GI products, the agricultural trade increased by 7.7% in my data. However, only the EU and small participants from developed countries adopted GI-related regulations, so the effects particularly hold in the transaction between developed countries. I provided the regional-specific estimates of these four provisions and confirmed the GI-related provision only affects North-North trade (Panel B, Table 6). In addition, the SPS measures increased North-North and South-South but decreased North-South agricultural trade. The ability and incentive play a vital role in exporters' decision on whether to adopt new standards to access foreign consumers, resulting in these regional-specific differences.

More importantly, I found that PTAs are only effective in agricultural trade between similar income countries and did not affect the trade between developing and developed countries (Panel A, Table 6). Although I did not provide further analyses on this topic, the policy implication of this finding is already critical, considering that the recent development of PTAs is between North and South countries. Moreover, I did not obtain any provisions selected by the plug-in Lasso if I sub-sampled North and South countries, implying that no PTA provision consistently affects agricultural trade between developing and developed countries. In sum, the combined results call into question the current PTA regime and its development directions for sustainable integration of global agriculture.

References

Ahrens, A., Hansen, C. B., and Schaffer, M. E. (2020). lassopack: Model selection and prediction with regularized regression in Stata. *The Stata Journal*, 20(1):176–235.

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *The American economic review*, 69(1):106–116.

Anderson, J. E. and Van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American economic review*, 93(1):170–192.

Anderson, J. E. and Yotov, Y. V. (2010). The changing incidence of geography. *American Economic Review*, 100(5):2157–86.

Athey, S. and Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic perspectives*, 31(2):3–32.

Athey, S. and Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11:685–725.

Baccini, L. (2019). The economics and politics of preferential trade agreements. *Annual Review of Political Science*, 22(1):75–92.

Bagwell, K. and Staiger, R. W. (2004). *The economics of the world trading system.* MIT press, Cambridge, MA.

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

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

Bajari, P., Nekipelov, D., Ryan, S. P., and Yang, M. (2015). Machine learning methods for demand estimation. *American Economic Review*, 105(5):481–85.

Batarseh, F. A., Gopinath, M., Monken, A., and Gu, Z. (2021). Public policymaking for international agricultural trade using association rules and ensemble machine learning. *Machine Learning with Applications*, 5:100046.

Belloni, A., Chen, D., Chernozhukov, V., and Hansen, C. (2012). Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica*, 80(6):2369–2429.

Belloni, A., Chernozhukov, V., Hansen, C., and Kozbur, D. (2016). Inference in high-dimensional panel models with an application to gun control. *Journal of Business & Economic Statistics*, 34(4):590–605.

Bergstrand, J. H., Larch, M., and Yotov, Y. V. (2015). Economic integration agreements, border effects, and distance elasticities in the gravity equation. *European Economic Review*, 78:307–327.

Borchert, I., Larch, M., Shikher, S., and Yotov, Y. V. (2021). The international trade and production database for estimation (ITPD-E). *International Economics*, 166:140–166.

Breinlich, H., Corradi, V., Rocha, N., Ruta, M., Santos Silva, J., and Zylkin, T. (2022). Machine learning in international trade research-evaluating the impact of trade agreements.

Campi, M. and Dueñas, M. (2019). Intellectual property rights, trade agreements, and international trade. *Research Policy*, 48(3):531–545.

Carter, C. A. and Gunning-Trant, C. (2010). US trade remedy law and agriculture: Trade diversion and investigation effects. *Canadian Journal of Economics/Revue canadienne d'économique*, 43(1):97–126.

Carter, C. A. and Steinbach, S. (2018). Trade diversion and the initiation effect: A case study of US trade remedies in agriculture. *NBER Working Paper*, (24745).

Chaney, T. (2008). Distorted gravity: The intensive and extensive margins of international trade. *American Economic Review*, 98(4):1707–21.

Chauffour, J. and Maur, J. (2011). *Preferential Trade Agreement Policies for Development: A Handbook*. Trade and Development. World Bank, Washington, DC.

Cheong, J., Tang, K. K., et al. (2018). The trade effects of tariffs and non-tariff changes of preferential trade agreements. *Economic Modelling*, 70:370–382.

Correia, S., Guimarães, P., and Zylkin, T. (2020). Fast Poisson estimation with high-dimensional fixed effects. *The Stata Journal*, 20(1):95–115.

Croser, J. and Anderson, K. (2011). Changing contributions of different agricultural policy instruments to global reductions in trade and welfare. *World Trade Review*, 10(3):297–323.

Curzi, D. and Huysmans, M. (2022). The impact of protecting EU geographical indications in trade agreements. *American Journal of Agricultural Economics*, 104(1):364–384.

Disdier, A.-C., Fontagné, L., and Mimouni, M. (2008). The impact of regulations on agricultural trade: Evidence from the SPS and TBT agreements. *American Journal of Agricultural Economics*, 90(2):336–350.

Dür, A., Baccini, L., and Elsig, M. (2014). The design of international trade agreements: Introducing a new dataset. *The Review of International Organizations*, 9(3):353–375.

Duvaleix, S., Emlinger, C., Gaigné, C., and Latouche, K. (2021). Geographical indications and trade: Firm-level evidence from the french cheese industry. *Food Policy*, 102:102118.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5):1741–1779.

Egger, P., Larch, M., Staub, K. E., and Winkelmann, R. (2011). The trade effects of endogenous preferential trade agreements. *American Economic Journal: Economic Policy*, 3(3):113–43.

Einav, L. and Levin, J. (2014). Economics in the age of big data. *Science*, 346(6210):1243089.

Fajgelbaum, P., Grossman, G. M., and Helpman, E. (2011). Income distribution, product quality, and international trade. *Journal of political Economy*, 119(4):721–765.

Fally, T. (2015). Structural gravity and fixed effects. *Journal of International Economics*, 97(1):76–85.

Fulponi, L., Shearer, M., and Almeida, J. (2011). Regional trade agreements - Treatment of agriculture. *OECD Food, Agriculture and Fisheries Papers*, (44).

Gopinath, M., Batarseh, F. A., Beckman, J., Kulkarni, A., and Jeong, S. (2021). International agricultural trade forecasting using machine learning. *Data & Policy*, 3.

Grant, J. H. and Lambert, D. M. (2008). Do regional trade agreements increase members' agricultural trade? *American journal of agricultural economics*, 90(3):765–782.

Hallak, J. C. (2010). A product-quality view of the Linder hypothesis. *The Review of Economics and Statistics*, 92(3):453–466.

Hofmann, C., Osnago, A., and Ruta, M. (2017). *Horizontal Depth: A New Database on the Content of Preferential Trade Agreements*. World Bank, Washington, DC.

Hofmann, C., Osnago, A., and Ruta, M. (2019). The content of preferential trade agreements. *World Trade Review*, 18(3):365–398.

Horn, H., Mavroidis, P. C., and Sapir, A. (2010). Beyond the WTO? An anatomy of EU and US preferential trade agreements. *The World Economy*, 33(11):1565–1588.

Huysmans, M. and Swinnen, J. (2019). No terroir in the cold? A note on the geography of geographical indications. *Journal of agricultural economics*, 70(2):550–559.

Josling, T. (2006). The war on terroir: Geographical indications as a transatlantic trade conflict. *Journal of agricultural economics*, 57(3):337–363.

Kawai, M. and Wignaraja, G. (2011). Asian FTAs: Trends, prospects and challenges. *Journal of Asian Economics*, 22(1):1–22.

Kohl, T., Brakman, S., and Garretsen, H. (2016). Do trade agreements stimulate international trade differently? Evidence from 296 trade agreements. *The World Economy*, 39(1):97–131.

Koo, W. W. and Kennedy, P. L. (2006). The impact of agricultural subsidies on global welfare. *American Journal of Agricultural Economics*, 88(5):1219–1226.

Kowalski, P., Büge, M., Sztajerowska, M., and Egeland, M. (2013). State-owned enterprises: Trade effects and policy implications. *OECD Trade Policy Papers*, (147).

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *The American Economic Review*, 70(5):950–959.

Larch, M. and Yotov, Y. V. (2016). General equilibrium trade policy analysis with structural gravity.

Limão, N. (2006). Preferential trade agreements as stumbling blocks for multilateral trade liberalization: Evidence for the United States. *American Economic Review*, 96(3):896–914.

Mansfield, E. D. and Pevehouse, J. C. (2000). Trade blocs, trade flows, and international conflict. *International organization*, 54(4):775–808.

Mansfield, E. D. and Reinhardt, E. (2003). Multilateral determinants of regionalism: The effects of GATT/WTO on the formation of preferential trading arrangements. *International organization*, 57(4):829–862.

Martin, W. (2018). A research agenda for international agricultural trade. *Applied Economic Perspectives and Policy*, 40(1):155–173.

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

Mattoo, A., Rocha, N., and Ruta, M. (2020). *Handbook of Deep Trade Agreements*. World Bank, Washington, DC.

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

Moschini, G., Menapace, L., and Pick, D. (2008). Geographical indications and the competitive provision of quality in agricultural markets. *American Journal of Agricultural Economics*, 90(3):794–812.

Mullainathan, S. and Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106.

Murina, M. and Nicita, A. (2017). Trading with conditions: The effect of sanitary and phytosanitary measures on the agricultural exports from low-income countries. *The World Economy*, 40(1):168–181.

Peterson, E., Grant, J., Roberts, D., and Karov, V. (2013). Evaluating the trade restrictiveness of phytosanitary measures on US fresh fruit and vegetable imports. *American Journal of Agricultural Economics*, 95(4):842–858.

Redding, S. and Venables, A. J. (2004). Economic geography and international inequality. *Journal of international Economics*, 62(1):53–82.

Rocha, N. and Ruta, M. (2022). *Deep Trade Agreements: Anchoring Global Value Chains in Latin America and the Caribbean*. World Bank, Washington, DC.

Sá Porto, P. C. d., Canuto, O., and Morini, C. (2015). The impacts of trade facilitation measures on international trade flows. *World Bank Policy Research Working Paper*, (7367).

Santeramo, F. G. and Lamonaca, E. (2019). The effects of non-tariff measures on agri-food trade: A review and meta-analysis of empirical evidence. *Journal of Agricultural Economics*, 70(3):595–617.

Santeramo, F. G. and Lamonaca, E. (2021). Standards and regulatory cooperation in regional trade agreements: What the effects on trade? *Applied Economic Perspectives and Policy*.

Scoppola, M., Raimondi, V., and Olper, A. (2018). The impact of eu trade preferences on the extensive and intensive margins of agricultural and food products. *Agricultural economics*, 49(2):251–263.

Shepherd, B. and Wilson, N. L. (2013). Product standards and developing country agricultural exports: The case of the European Union. *Food Policy*, 42:1–10.

Silva, J. S. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and statistics*, 88(4):641–658.

Sorgho, Z. and Larue, B. (2014). Geographical indication regulation and intra-trade in the European Union. *Agricultural Economics*, 45(S1):1–12.

Storm, H., Baylis, K., and Heckelei, T. (2020). Machine learning in agricultural and applied economics. *European Review of Agricultural Economics*, 47(3):849–892.

Sun, L. and Reed, M. R. (2010). Impacts of free trade agreements on agricultural trade creation and trade diversion. *American Journal of Agricultural Economics*, 92(5):1351–1363.

Thompson-Lipponen, C. and Greenville, J. (2019). The evolution of the treatment of agriculture in preferential trade agreements. *OECD Food, Agriculture and Fisheries Papers*, (126).

Weidner, M. and Zylkin, T. (2021). Bias and consistency in three-way gravity models. *Journal of International Economics*, 132:103513.

Willemyns, I. (2016). Disciplines on state-owned enterprises in international economic law: Are we moving in the right direction? *Journal of International Economic Law*, 19(3):657–680.

WTO (2022). Regional Trade Agreements Database. http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx.

Yotov, Y. V., Piermartini, R., Larch, M., et al. (2016). *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*. WTO iLibrary.

Supplementary Tables

Table S1: Description of All PTA Provisions Selected by the Lasso Regressions

Provisions	Description
Competition Policy	
CP.03	Does the agreement promote open markets either (1) economy wide, (2) sector specific or (3) both?
CP.04	Does the agreement promote the principle of transparency?
CP.07	Does the agreement has as objective promoting fair competition and curbing anti-competitive practices?
CP.12	Does the agreement require the establishment/existence of competition law/measures?
CP.15	Does the agreement prohibits/regulates cartels/concerted practices?
CP.16	Does the agreement prohibits/regulates abuse of market dominance?
CP.18	Does the agreement regulates monopolies?
CP.20	Does the agreement regulates state aid?
CP.23	Does the agreement contain provisions that promote transparency?
CP.31	Does the agreement provide for the creation of a regional/agreement-relatedl competition authority?
CP.33	What general exceptions are included in the agreement? (GATS Article XIV list)
CP.34	Does the agreement allow for security exceptions?
Countervailing duties	
CVD.01	Countervailing duties disallowed
CVD.03	Countervailing duties allowed and with specific provisons
Environmental Laws	
ENV.03	General obligation of environmental cooperation
ENV.04	Does the agreement call for regulatory cooperation or harmonization in environmental regulation?
ENV.06	Does the agreement provide for a general exception to other obligations for environmental reasons?
ENV.19	Does the agreement require states to prevent pollution by ships?
ENV.20	Does the agreement require states to implement fisheries management?
ENV.24	Does the agreement require measures to prevent deforestation and/or require sustainable trade practices in forest products?
ENV.25	Does the agreement prohibit of dumping hazardous and toxic wastes?
ENV.28	Does the agreement require states to implement water management?

Continued on next page

Table S1: *Continued from previous page*

Provisions	Description
Export Taxes	
ET.02	Prohibits all export quotas / QRs between the Parties, but with reference to certain exceptions mentioned in the provision that are WTO-plus
ET.03	Prohibits new export quotas / QRs between the Parties
ET.06	Requires scheduling of export quota / QR between the Parties
ET.15	Prohibits an increase in the rate of any existing export tax
Intellectual Property Rights	
IPR.43	Designates list of GIs subject to protection by both Parties, subject to limited exceptions
Visa and Asylum - migration	
MIG.18	Does the agreement specifically allow parties to bar entry of natural persons based on public security/order reasons?
MIG.19	Does the agreement allow parties to undertake temporary safeguard measures to bar entry of natural persons?
MIG.24	Does the agreement encourage parties to undertake mutually agreed cooperation activities?
MIG.28	Does the agreement address the movement of migrant workers already employed by a company in the country of destination?
MIG.29	Does the agreement address the movement of migrant workers seeking employment in the country of destination?
MIG.30	Does the agreement positively address or facilitate persons obtaining residency in either party?
Movement of Capital	
MOC.02	Does the agreement distinguish new restrictions from existing restrictions?
MOC.27	Does the transfer provision explicitly exclude 'good faith and non-discriminatory application of its laws' governing capital account regulations?
MOC.57	Does the safeguard provision explicitly state that restrictions must be temporary with a maximum time limit specified?
MOC.113	Does the transfer provision apply specifically to transfers relating to contributions to capital?
Rules of Origin	
ROR.06	What is the length of the record keeping period?
ROR.19	Does the agreement contain drawback rules?
Sanitary and Phytosanitary	
SPS.44	Is there an SPS chapter or provision?
SPS.56	Is there a general IRC clause/Common policy/standardization programme (beyond trade-related objectives)?
State-Owned Enterprises	
STE.24	Does the agreement expressly regulate/exclude state enterprises pursuing public services?

Continued on next page

Table S1: *Continued from previous page*

Provisions	Description
STE.27	Does the agreement regulate ownership or property regimes, or liberalization processes?
STE.28	Does the agreement prohibit discrimination by state enterprises?
STE.32	Does the agreement require state enterprises not to distort trade?
STE.39	Does the agreement include any other specific discipline for certain sectors or objectives?
STE.46	Does the agreement provide for a body or a committee to deal with transparency or enforcement of the disciplines of state enterprises?
Subsidies	
SUB.03	Does the agreement prohibit or regulate export subsidies?
SUB.06	Does the agreement provide for exemptions for legitimate subsidies (such as environmental, R&D, regional)?
SUB.13	Does the agreement include any national treatment obligation (goods) for subsidies?
SUB.14	Does the agreement include any national treatment obligation (services or establishment) for subsidies?
SUB.16	Does the agreement provide for notification requirements (e.g. individual measures, annual reporting)?
SUB.26	Does the agreement regulate the imposition of countervailing duties?
SUB.29	Does the agreement provide for special treatment in favour of any party?
SUB.34	Does the agreement cover support granted by sub-central, regional and/or local authorities?
Technical Barriers to Trade	
TBT.06/33	Standards - Is the use or creation of regional standards promoted?
TBT.14	Conformity Assessment - Is the use or creation of regional standards?
TBT.16	Is the time period allowed for comments specified?
TBT.18	Contact points/consultations for exchange of information
TBT.19	Is a regional body established?
TBT.21	Are there regional consultations foreseen to resolve disputes?
TBT.22	Is there a mechanism to issue recommendations?
TBT.33	Standards: Is the use or creation of regional standards promoted?
TBT.34	Standards: Is the use of international standards promoted?
Trade Facilitation and Customs	
TF.11	Pre-arrival processing
TF.14	Post Clearance Audits

Continued on next page

Table S1: *Continued from previous page*

Provisions	Description
TF.20	Use of international standards
TF.21	Single Window
TF.26	Exchange of customs-related information
TF.29	Advance lodging in electronic format
TF.32	Mutual recognition of AOs
TF.38	Cooperate on law enforcement
TF.39	Exchange of information on best practices
TF.45	Issuance of proof of origin

Notes: This table lists all provisions selected as outcomes of any Lasso approach from the manuscript. The description is as it is from Mattoo et al. (2020), but abbreviated if necessary.

Table S2: The Second-stage Lasso Results for Manufacturing

Food Manufacturing							Non-food Manufacturing	
(1) CP.31	(2) ENV.19	(3) ET.06	(4) IPR.43	(5) SPS.44	(6) TBT.06/33	(7) TF.32	(8) CP.23	(9) ET.02
11.0%	5.3%	21.5%	22.5%	11.7%	48.7%	27%	34.2%	20.7%
CP.12 (0.59)	ENV.20 (0.75)	CVD.01 (0.72)	MOC.57 (0.68)	CP.03 (0.73)	CP.20 (0.64)	MIG.19 (0.68)	ENV.04 (0.69)	CP.18 (0.64)
ENV.03 (0.62)	ENV.25 (0.85)	ET.15 (0.62)	SPS.56 (0.60)	CP.04 (0.71)	TBT.14 (0.91)	MOC.27 (0.96)	ENV.24 (0.71)	ET.03 (0.76)
MIG.18 (0.78)	ENV.28 (0.81)	STE.32 (0.56)	STE.24 (0.61)	CP.16 (0.77)	TBT.19 (0.62)	TF.29 (0.94)	MIG.29 (0.73)	ET.06 (0.65)
MIG.24 (0.86)	MIG.28 (0.64)	SUB.13 (0.55)		CP.33 (0.84)	TBT.34 (0.52)		MOC.02 (0.60)	SUB.29 (0.63)
MIG.30 (0.83)	MIG.29 (0.59)	SUB.14 (0.52)		CP.34 (0.62)	TF.14 (0.66)		MOC.113 (0.57)	TF.26 (0.71)
TF.21 (0.52)	STE.24 (0.64)			CVD.03 (0.69)	TF.39 (0.70)		TBT.33 (0.81)	
TF.32 (0.54)	STE.32(0.72)			ROR.06 (0.83)				
TF.45 (0.60)	SUB.06(0.70)			ROR.19 (0.65)				
				SPS.56 (0.59)				

Notes. Each column presents the second-stage Lasso results of the selected PTA provisions associated with manufacturing trade flows. Columns (1) - (7) report the results for food manufacturing and (8) - (9) for non-food manufacturing. The estimated percent changes in manufacturing trade by the provision selected in the first-stage Lasso are in the second row. The subsequent rows report all provisions not identified at the first-stage but in the second-stage Lasso regressions. The numbers in brackets are correlations with the provisions selected by the first stage.