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

Dongin Kim

Selected presentation for the International Agricultural Trade Research Consortium's (IATRC's) 2022 Annual Meeting: Transforming Global Value Chains, December 11-13, 2022, Clearwater Beach, FL.

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

Dongin Kim

Agricultural and Resource Economics
University of Connecticut

IATRC Annual Meeting, December 13, 2022

UCONN

Overview

- ✓ I study the impact of Preferential Trading Agreements (PTAs) in agriculture.
  - Focusing on **non-tariff provisions** that might create heterogeneity

- ✓ Using a machine learning selection algorithm to overcome the potential econometrics challenges dealing with many policy variables
  - I find the 7 most relevant provisions affecting agricultural trade
  - Which capture most of PTA's agricultural trade effects.

- ✓ In the structural gravity general equilibrium context, the PTA may increase the agricultural GDP:
  - If PTA contains the above-mentioned selected provisions
  - And/or if countries have more PTA partners.

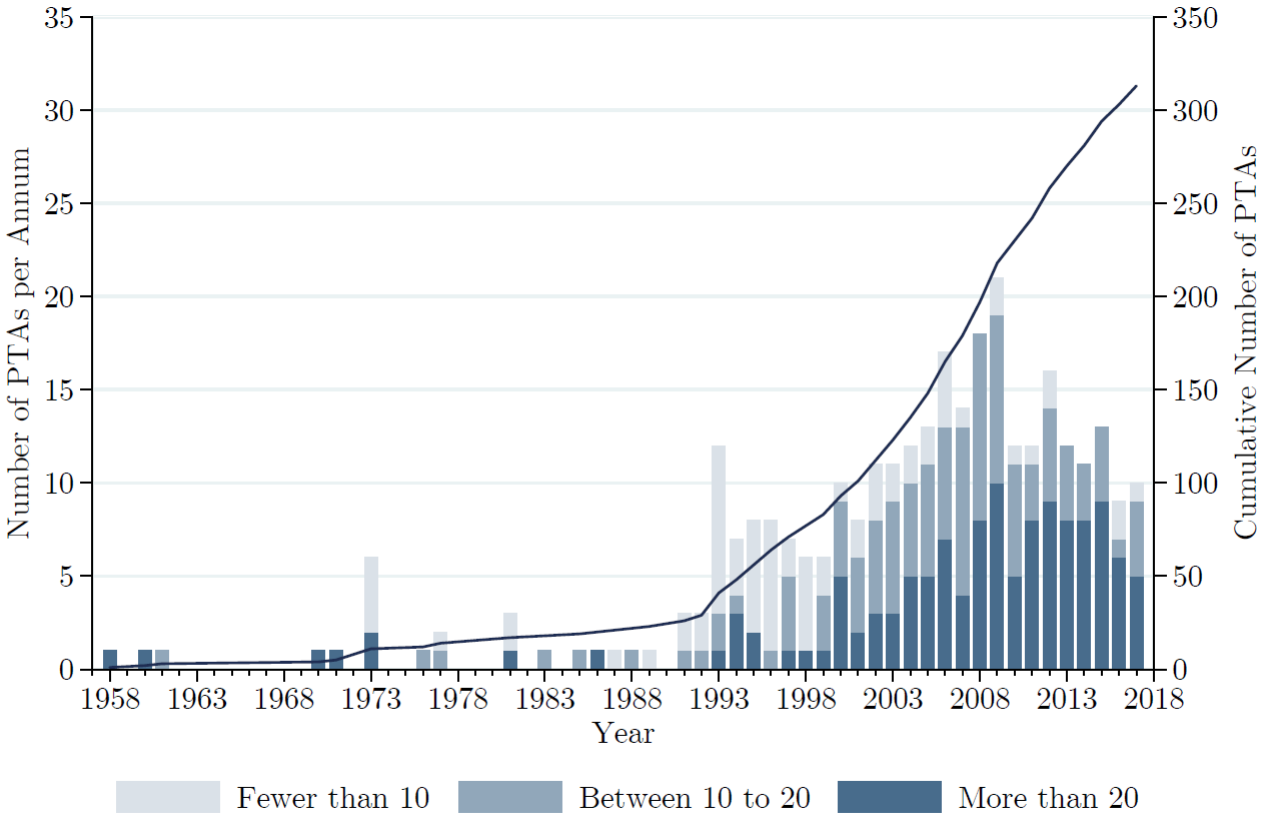
Motivation

- ✓ Countries have signed preferential trade agreements (PTAs)
  - Numbers have increased rapidly since 1990

- ✓ PTAs contain hundreds of provisions
  - Policy areas covered in PTAs increased
  - Number of provisions significantly varies across PTAs

- ✓ Non-tariff measures are more prominent in agriculture than in other sectors

PTA Enforcements by Year



Source: Author's calculation using the dataset published by Hofmann et al. (2017).

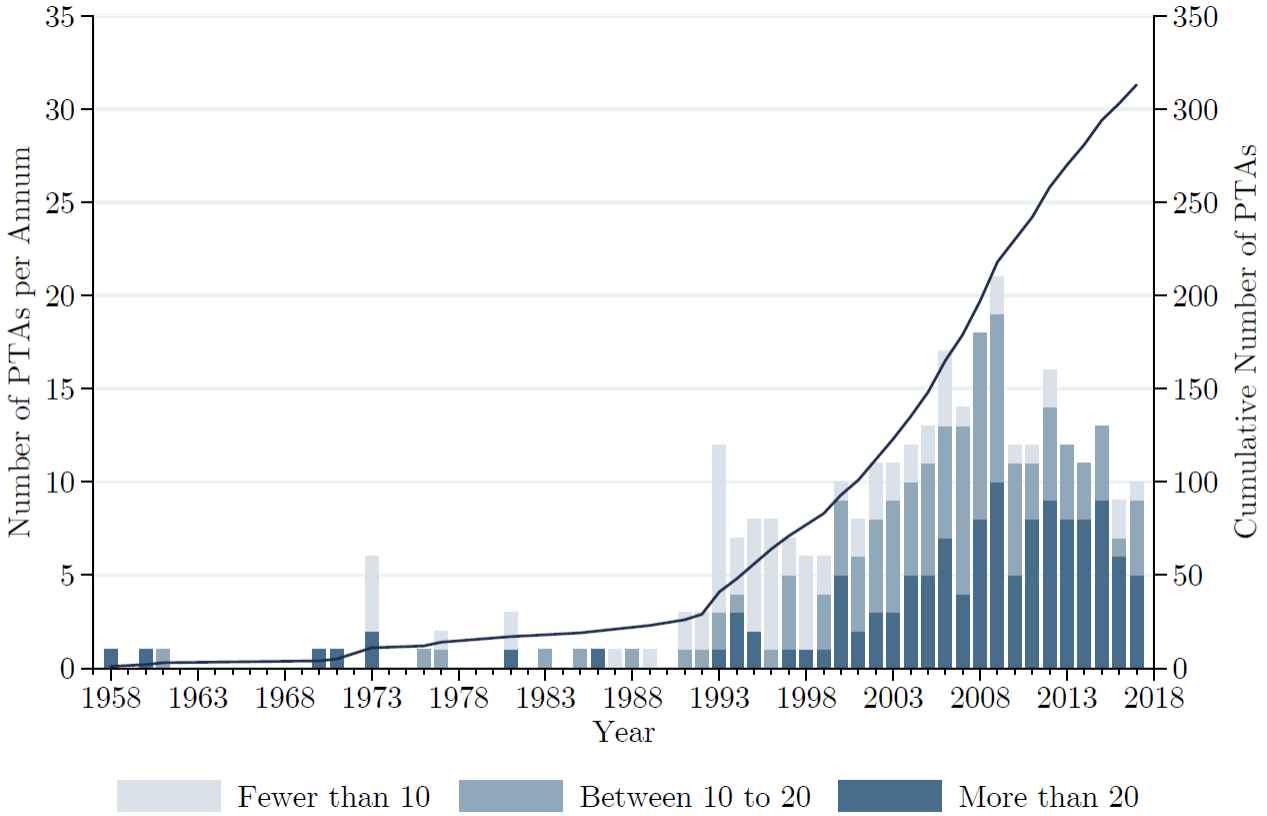
Motivation

✓ Countries have signed preferential trade agreements (PTAs)
- Numbers have increased rapidly since 1990

✓ PTAs contain hundreds of provisions
- Policy areas covered in PTAs increased
- Number of provisions significantly varies across PTAs

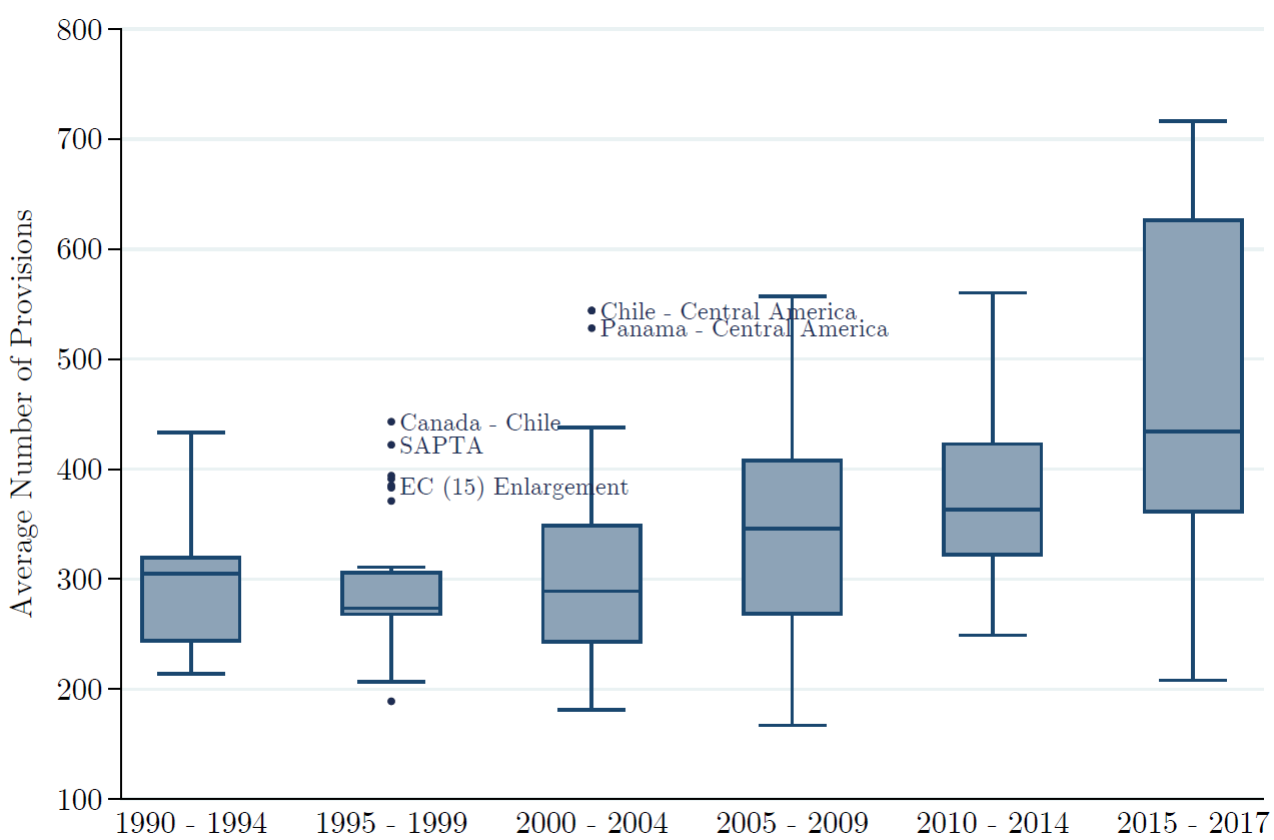
✓ Non-tariff measures are more prominent in agriculture than in other sectors

Policy Coverage



Source: Author's calculation using the dataset published by Hofmann et al. (2017).

Number of Non-tariff Provisions in PTAs



Source: Author's calculation using the dataset published by Mattoo et al. (2020).

Motivation

- ✓ Countries have signed preferential trade agreements (PTAs)
  - Numbers have increased rapidly since 1990
- ✓ PTAs contain hundreds of provisions
  - Policy areas/Provisions covered in PTAs increased
  - Number of provisions significantly varies across PTAs
- ✓ Non-tariff measures are more prominent in agriculture than in other sectors

Motivation

✓ Non-tariff provisions that might affect agricultural trade

- **SPS and TBT**:
  Disdier et al. (2008), Peterson et al. (2013), Murina and Nicita (2017)

- **Subsides and Competition Policies** (c.f., non-discrimination principle):
  Bagwell and Staiger (2004), Limao (2006)

- **Intellectual Property Rights** (i.e., geographical indication):
  Huysmans and Swinnen (2019), Duvaleix et al. (2021), Curzi and Huysmans (2022)

- **Others** (e.g., environmental law)

Research Questions

- ✓ Are there heterogeneous trade effects of enforcing PTAs
  - Due to the different provisions?
  - What provisions affect agricultural trade?

- ✓ How much do countries gain after enforcing PTAs?
  - Partnered/non-partnered countries
  - Containing/not-containing relevant provisions

- ➢ I test with theoretically-consistent estimators using recently published datasets and a data-driven estimation strategy

Literature Review

- ✓ PTA's trade effects on agriculture
  - Grant and Lambert (2008), Sun and Reed (2010), Scoppola et al. (2018), etc.
  - ➢ What has been missing: Identify non-tariff impacts of PTAs

- ✓ Heterogeneous impacts of PTAs
  - Recent studies established the importance of non-tariff provisions in determining PTAs' overall impact (Hofmann et al., 2017; Hofmann et al., 2019)
  - Overfitting and multicollinearity issues arise due to many policy variables (Kohl and Garrestsen, 2016; Santeramo and Lamonaca, 2021)
  - ➢ Reduce dimension with *ad hoc* assumptions
    - Select the provisions (e.g., Campi et al., 2019)
    - Create indices (e.g., Hofmann et al., 2017; Mattoo et al., 2022)
  - ➢ Rely on data-driven approach
    - Breinlich et al., 2021

Estimation Strategy

✓ Estimation Strategy

1. Reduce the policy variable dimension using data-driven method (Regularization regression)
   - Select the most relevant PTA provisions affecting agricultural trade

2. Quantify the impact of PTA and the provisions on agricultural trade

Data

- ✓ Bilateral trade data
  - ➢ **ITPD-E** from USITC by Borchert et al. (2021)
    - From 1988 to 2017
    - 249 exporters and 251 importers
    - Include intra-national trade needed to obtain theory-consistent estimators
    - Used in many recent studies (e.g., Campos et al., 2021; Rideley et al., 2022; Timini et al., 2022)

- ✓ Policy data
  - ➢ **Deep Trade Agreements** from World Bank by Mattoo et al. (2020)
    - 937 provisions under 18 most frequently covered policy areas for 282/317 PTAs
    - All provisions are dummy variables (only indicate the presence)
    - Excluded observations for the countries associated with the 35 unmapped-PTAs
      → the final data includes 213 exporters and importers.

Gravity Model

✓ Armington-CES Gravity Model:

$$X_{ij} = \tau_{ij}^{-\theta} \frac{Y_i}{\pi_i^{-\theta}} \frac{E_j}{\rho_j^{-\theta}}$$

- X_{ij} is agricultural exports from country i to j
- τ_{ij} is the bilateral trade cost
- Y_i is country i's total agricultural production
- E_j is country j's aggregate expenditure
- π_i and ρ_j are the multilateral resistance terms
- θ represents the elasticity of substitution

Structural Gravity Model

✓ Theoretically-consistent Empirical Model

$$\mu_{ijt} \coloneqq E(X_{ijt}|\boldsymbol{PTA}'_{ijt-1}, \alpha_{it}, \gamma_{jt}, \delta_{ij}) = \exp(\boldsymbol{PTA}'_{ijt-1}\boldsymbol{\beta}' + \alpha_{it} + \gamma_{jt} + \delta_{ij})$$

- X_{ijt} is bilateral agricultural trade from country i to j at year t
- α_{it} and γ_{jt} encompass the time-varying country dummy variables
- δ_{ij} denotes the set of country-pair fixed effects
- $\boldsymbol{PTA}'_{ijt-1}$ represents the first-differenced vector of PTA provisions
- Poisson PML to estimate $\boldsymbol{\beta}'$
  - BUT we don't want to estimate with 937 variables
  - We need to reduce the dimension $PTA_l \subseteq \boldsymbol{PTA}'$

How Regularization Regression Works?

- ✓ How Lasso regression selects the variables?
  - It is a method for penalizing extra variables that are not significantly improve the model specification

$$\arg\min_{\beta_l} \quad RSS + \lambda \underbrace{\sum_{l=1}^{m} |\beta_l|}_{Penalty\ Term}$$

"Adding one more variable": $-$ $+$

- ✓ I use PPML version of Lasso suggested by Breinlich et al. (2021)

Results: Selected PTA Provisions (Agriculture)

	PPML	Lasso	Post-Lasso	PPML
PTA - Presence of PTA	0.240^{***} (0.036)			0.022 (0.028)
ENV.19 - Prevent Pollution by Ships		0.052	-0.048 (0.042)	-0.040 (0.045)
ET.06 - Require Scheduling of EQ/QR		0.014	0.176^{***} (0.061)	0.174^{***} (0.062)
SPS.44 - Include SPS Chapter		0.052	0.079^{***} (0.018)	0.096^{***} (0.027)
STE.39 - Include Sector-specific Discipline on STE		0.107	0.162^{***} (0.056)	0.175^{***} (0.060)
SUB.16 - Require for Prior Notification on Subsidies		0.004	0.189^{***} (0.030)	0.191^{***} (0.030)
TBT.06/33 - Use Regional Standards		0.091	0.197^{**} (0.095)	0.191^{**} (0.096)
TF.32 - Mutual Recognition of AOs		0.251	0.451^{**} (0.058)	0.448^{**} (0.058)
Observation	798,239	786,621	786,821	786,821
Pseudo R^2	0.994		0.994	0.994

Note: Heteroskedasticity-robust standard errors are in the parentheses, clustered by exporter-importer pair. For the estimated parameters, the semi-elasticity is given by 100*(exp(β) - 1)%.

Results: Other Sectors (Processed Food)

	PPML	Lasso	Post-Lasso	PPML
PTA - Presence of PTA	0.189^{***} (0.036)			-0.011 (0.023)
CP.31 - Create Regional/Agreement-specific Authority		0.058	0.104^{*} (0.062)	0.100 (0.062)
ENV.19 - Prevent Pollution by Ships		0.018	0.052 (0.034)	0.056 (0.036)
ET.06 - Require Scheduling of EQ/QR		0.131	0.198^{***} (0.047)	0.201^{***} (0.048)
IPR.43 - Designate the List of GI Products		0.012	0.203^{***} (0.019)	0.195^{***} (0.023)
SPS.44 - Include SPS Chapter		0.014	0.111^{***} (0.016)	0.119^{***} (0.023)
TBT.06/33 - Use Regional Standards		0.438	0.397^{***} (0.057)	0.398^{***} (0.057)
TF.32 - Mutual Recognition of AOs		0.069	0.239^{***} (0.042)	0.241^{***} (0.043)
Observation	957,886	946,580	946,580	946,580
Pseudo R^2	0.996		0.996	0.996

Note: Heteroskedasticity-robust standard errors are in the parentheses, clustered by exporter-importer pair. For the estimated parameters, the semi-elasticity is given by 100*(exp(β) - 1)%.

Results: Other Sectors (Non-food Manufacturing)

	PPML	Lasso	Post-Lasso	PPML
PTA - Presence of PTA	0.109 (0.077)			0.013 (0.091)
CP.31 - Create Regional/Agreement-specific Authority		0.031	0.294*** (0.055)	0.287*** (0.075)
ET.06 - Require Scheduling of EQ/QR		0.008	0.188*** (0.051)	0.184*** (0.057)
Observation	1,217,498	1,206,138	1,206,138	1,206,138
Pseudo R^2	0.996		0.991	0.991

Note: Heteroskedasticity-robust standard errors are in the parentheses, clustered by exporter-importer pair. For the estimated parameters, the semi-elasticity is given by 100*(exp(β) - 1)%.

General Equilibrium Analysis

✓ Structural Gravity System (*conditional GE described in Yotov et al., 2016):

$$X_{ij} = \tau_{ij}^{-\theta} \frac{Y_i}{\pi_i^{-\theta}} \frac{E_j}{\rho_j^{-\theta}}; \quad \pi_i^{-\theta} = \sum_{j=1}^{N} \frac{E_j \tau_{ij}^{-\theta}}{\rho_j^{-\theta}}; \quad \rho_j^{-\theta} = \sum_{i=1}^{N} \frac{Y_i \tau_{ij}^{-\theta}}{\pi_i^{-\theta}}$$

- In the above partial equilibrium estimates, we considered $Y_i, E_j, \pi_i^{-\theta}$, and $\rho_j^{-\theta}$ being constant when $\tau_{ij}^{-\theta}$ got affected by $\boldsymbol{PTA}'_{ijt}$
- So, it assumes that enforcing PTA only affects the partnered countries;
  - No impact on the other countries
- But $\pi_i^{-\theta}$ and $\rho_j^{-\theta}$ can be defined by E_j, Y_i, and $\tau_{ij}^{-\theta}$;
  - Accounts for changes in all other trading partners

General Equilibrium Analysis

✓ By solving the structural system (Baier et al., 2019):

$$\hat{X}_{ij} = \frac{\hat{w}_i^{-\theta} \exp(\beta' PTA'_{ij})}{\hat{P}_j^{-\theta}} \hat{E}_j;$$

$$\hat{W}_i = \frac{\hat{E}_i}{\hat{P}_i}$$

- $\hat{w}$ and $\hat{P}$ are predicted wage and price levels decided endogenously in the system.
- I iterate my calculation over the years and provide average annual effects of PTA enforcement in agricultural welfare from 1990 to 2017.

Results: Annual Welfare Gain/Loss

	No PTA	PTA with no Relevant Provisions	PTA with Fewer Relevant Provisions	PTA with More Relevant Provisions
Worldwide	-0.13%	$+0.62\%$	$+0.62\%$	$+1.41\%$
North	-0.04%	$+0.90\%$	$+1.11\%$	$+2.27\%$
South	-0.15%	$+0.41\%$	$+0.37\%$	$+0.46\%$

Note: This table shows the annual gain/loss ($\%\Delta\widehat{W}_i$) of agricultural welfare.

Results: Top 10 Agricultural Exporters

Country	PTA (Count)	Effective PTA (Count)	Welfare Gain/Loss (%)	ΔExports (%)
USA	20	0	+0.65	-1.33
BRA	43	0	+0.04	-1.71
NLD	94	50	+2.21	+1.86
CAN	14	1	+0.09	-2.53
CHN	24	1	+1.59	-0.77
AUS	27	0	-1.22	-0.21
ESP	90	50	+5.25	-0.03
IND	52	0	-5.09	-0.24
DEU	98	50	+9.33	+0.16
FRA	96	50	+10.84	+0.22

Note: This simulation is based on 2016.

Summary

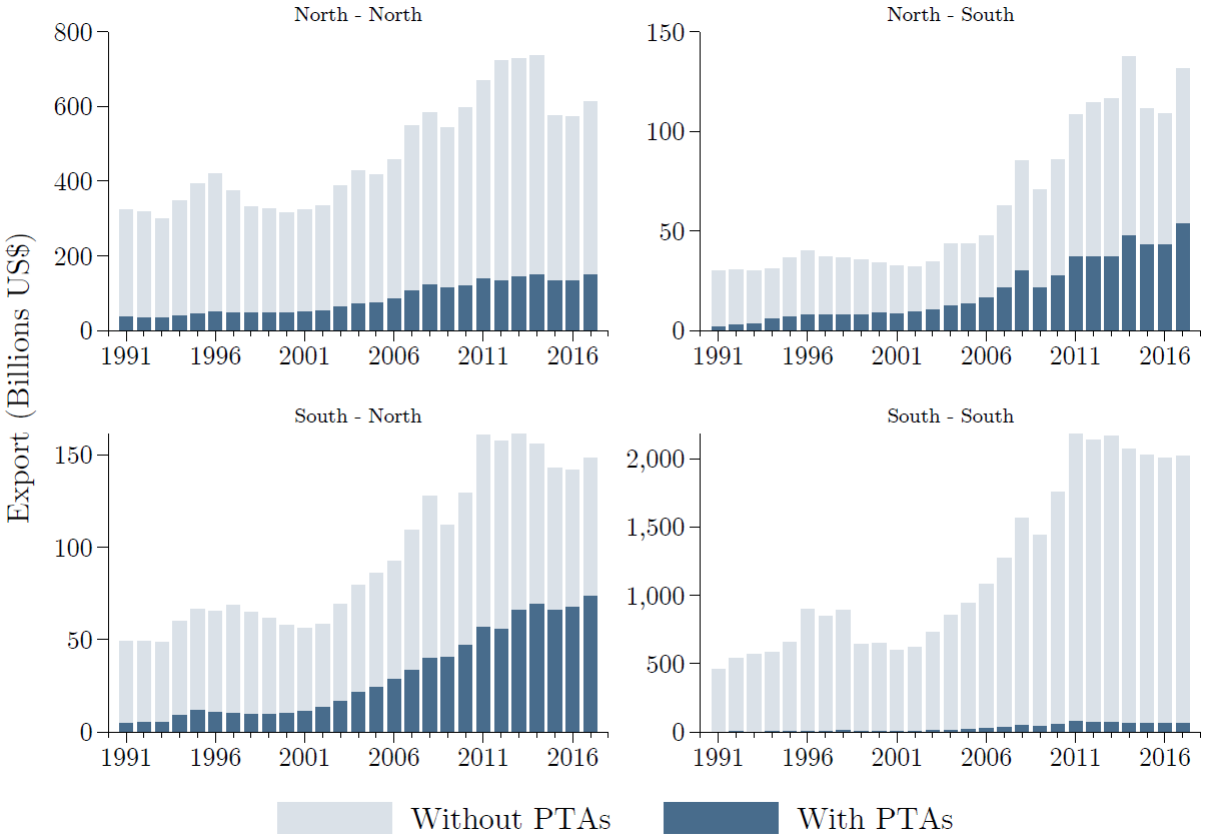
- ✓ PTA's heterogeneous trade effects caused by non-tariff provisions
  - 6 most relevant provisions with policy areas: Anti-discriminatory (Export Taxes, Subsidy, State Enterprises), SPS and TBT measures, and Trade Facilitations
  - PTAs without these provisions have no impact on agricultural trade

- ✓ The welfare gain by PTA enforcement is more significant
  - If PTA contains more relevant provisions
  - For the developed countries
  - For the countries have more PTA partners

Next Step

- ✓ The model cannot fully control endogeneity of policy
  - Adding reliable confounding variables
- ✓ Tariff impacts
  - Different time-lags (tariff reduction effects)
  - Sub-industry level analysis with average tariff
- ✓ Prediction
  - Full endowment model
  - More scenarios -> policy implications

Supplementary: Bilateral Agricultural Trade between PTA Partners



Supplementary: Selected PTA Provisions and Bundles

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

Supplementary: Including/excluding Intra-national Trade

	Including Domestic Sales		**Excluding Domestic Sales**	
PTA - Presence of PTA	0.240^{***}	0.022	0.035	-0.020
	(0.036)	(0.028)	(0.031)	(0.066)
ENV.19 - Prevent Pollution by Ships		-0.040		0.018
		(0.045)		(0.090)
ET.06 - Require Scheduling of EQ/QR		0.174^{***}		0.099
		(0.062)		(0.126)
SPS.44 - Include SPS Chapter		0.096^{***}		0.101
		(0.027)		(0.068)
STE.39 - Include Sector-specific Discipline on STE		0.175^{***}		0.321^{***}
		(0.060)		(0.119)
SUB.16 - Require for Prior Notification on Subsidies		0.191^{***}		-0.107^{*}
		(0.030)		(0.058)
TBT.06/33 - Use Regional Standards		0.191^{**}		0.153
		(0.096)		(0.122)
TF.32 - Mutual Recognition of AOs		0.448^{**}		0.312^{***}
		(0.058)		(0.130)
Observation	798,239	786,821	790,978	779,560
Pseudo R^2	0.994	0.994	0.963	0.964

Note: Heteroskedasticity-robust standard errors are in the parentheses, clustered by exporter-importer pair. For the estimated parameters, the semi-elasticity is given by 100*(exp(β) - 1)%.

Supplementary: Different Time-lags

	Simple Dummy	w/ Provisions
PTA$_{t-1}$	0.094^{***}	-0.028
	(0.036)	(0.043)
PTA$_{t-3}$	0.036	0.040
	(0.036)	(0.025)
PTA$_{t-5}$	0.251^{***}	-0.005
	(0.029)	(0.030)
ENV.19 - Prevent Pollution by Ships		0.031
		(0.065)
ET.06 - Require Scheduling of EQ/QR		0.069
		(0.131)
SPS.44 - Include SPS Chapter		0.095^{*}
		(0.052)
STE.39 - Include Sector-specific Discipline on STE		0.330^{***}
		(0.116)
SUB.16 - Require for Prior Notification on Subsidies		-0.150^{**}
		(0.058)
TBT.06/33 - Use Regional Standards		0.165
		(0.102)
TF.32 - Mutual Recognition of AOs		0.202
		(0.157)
Observation	$672{,}041$	$657{,}870$
Pseudo R^2	0.995	0.966

Note: Heteroskedasticity-robust standard errors are in the parentheses, clustered by exporter-importer pair. For the estimated parameters, the semi-elasticity is given by 100*(exp(β) - 1)%.

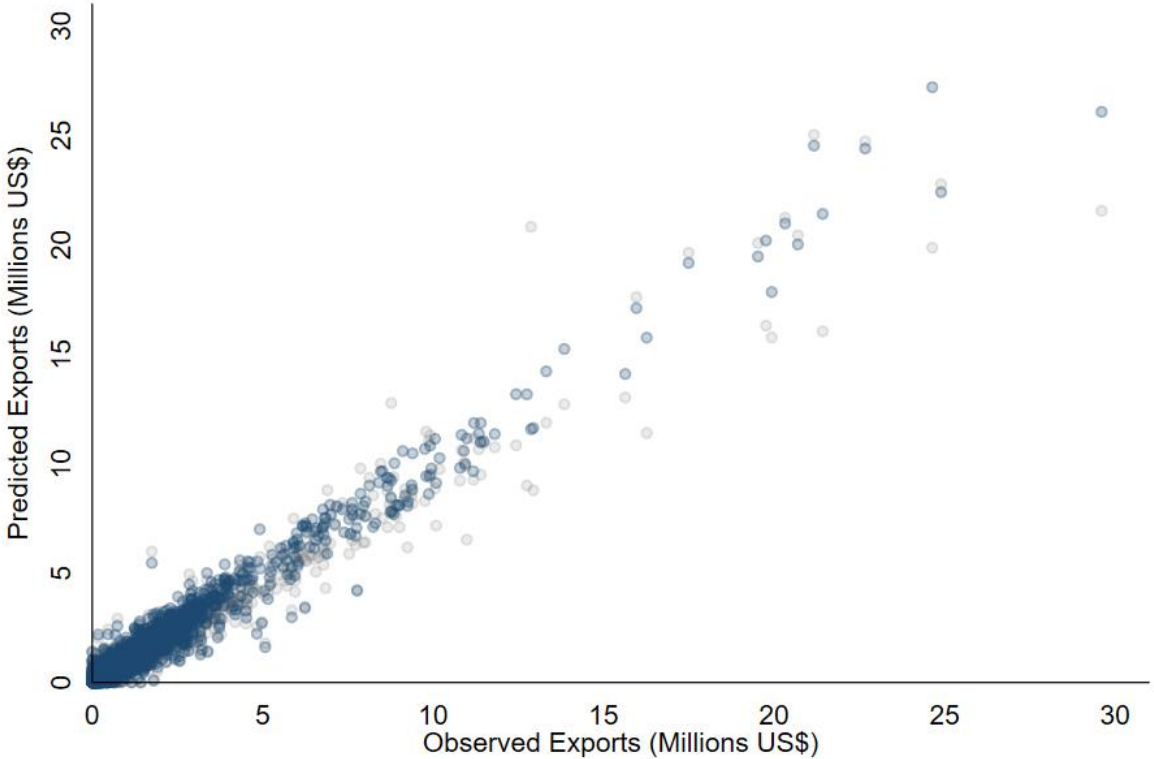
Supplementary: Variable Selection Method in Detail

✓ PPML-Lasso regression model (Breinlich et al., 2021):

$$\left(\hat{\boldsymbol{\beta}}, \hat{\alpha}, \hat{\gamma}, \hat{\delta}\right) \coloneqq \arg\min_{\beta,\alpha,\gamma,\delta} \underbrace{\frac{1}{N}\sum_{i,j,t}\left(\mu_{ijt} - X_{ijt}\ln\mu_{ijt}\right)}_{-1 \times PPML\ Pseudo\ Likelihood} + \underbrace{\frac{1}{N}\sum_{l=1}^{m}\hat{\phi}_l\lambda|\beta_l|}_{Penalty\ Term}$$

- μ_{ijt} is the conditional mean, defined in the previous slide
- n is the number of observations (relies on countries and year)
- $\lambda \geq 0$ is the standard tuning parameter for any regularization regression
- $\hat{\phi}_l \geq 0$ is the regressor-specific penalty term suggested by Belloni et al. (2016)
- Used the plug-in algorithms to find λ and $\hat{\phi}_l$

Supplementary: Predicted vs. Observed Trade



Note: The gray dots represent the model fit of the simple dummy approach. The blue dots represent the model fit of including the selected provisions.