

Deep Trade Agreements and Development Heterogeneous Effects Staggered Difference-in-Differences 2000-2023

Shahadat Hossain¹; Dr. Ji Jiali²

^{1,2}Business School, China West Normal University Nanchong, Sichuan, China

Publication Date: 2026/08/14

Abstract: Contemporary trade governance is marked by an increasing number of Deep Trade Agreements (DTAs); however the developmental implications for emerging economies are subject to some level of contestation due to the existence of methodological biases and a lack of consideration of heterogeneity. This is an investigation that addresses four key deficiencies: (1) the bias of staggered Difference-in-Differences (DiD) designs is overcome using a multi-estimator strategy (Callaway & Sant'Anna 2021; Borusyak, Jaravel, & Portal 2021); (2) triple heterogeneity, which encompasses DTA type, depth and domestic institutional capacity, is systematically addressed; (3) institutional thresholds are estimated instead of postulated; (4) inclusive growth implications are taken into consideration beyond aggregate trade indicators. Empirical analysis of a panel consisting of 140 developing economies between the years 2000-2023 has shown that DTAs have led to an increase in exports by 12.7-14.2 per cent and foreign direct investment (FDI) by 2.3-2.5 percentage points and a decline in the Human Development Index (HDI) in rural areas by 0.7-0.8 per cent. The size of increases is about 2.3 times greater for countries that cross a Rule-of-law threshold of -0.5 on the World Governance Indicators scale. North-South deep agreements have the highest export benefits (28.2 percent growth). Vertical depth (i.e. granularity of commitments in core policy categories) outperforms horizontal depth (i.e. simple count of policy areas) in terms of manufacturing productivity and including labor standards in DTAs reduce rural-urban disparities by 50 per cent. Reduction of Trade Policy Uncertainty (TPU), instrumented on the number of veto players in partner governments, is the main channel, representing 33 per cent of the total effect. These findings recognize the need to align DTA accession with domestic institutional reforms and consciously draw policies to design inclusive policies to achieve equitable development gains.

Keyword: Deep Trade Agreements Staggered Difference-in-Differences Threshold Regression Inclusive Growth Developing Economies Trade Policy Uncertainty Global Value Chain Integration.

How to Cite: Shahadat Hossain; Dr. Ji Jiali (2026) Deep Trade Agreements and Development Heterogeneous Effects Staggered Difference-in-Differences 2000-2023. *International Journal of Innovative Science and Research Technology*, 11(7), 3713-3724. <https://doi.org/10.38124/ijisrt/26jul175>

I. INTRODUCTION

A. Introduction and Research Motivation

The architecture of the global trade governance has changed fundamentally from multilateral tariff negotiations to preferential integration based on 'depth', i.e. regulatorily aligned nonboarder policy areas (WTO 2023). As of 2023, more than 350 Preferential Trade Agreements (PTAs) are still in place and 82 per cent of developing economies are parties to at least one DTA that cover issues such as trade in investments, intellectual property rights, competition, labor standards and environmental regulation (World Bank 2023). As opposed to the rather shallow agreements of the twentieth century (i.e., averaging eight policy areas), the modern DTAs cover an average of seventeen domains, reflecting the increasingly regulatory nature of modern trade barriers (UNCTAD 2023). For developing economies, the rationale for the use of DTAs is strategic - to unlock access to high-income markets, to attract FDI, associated technology

spillovers, institutionalize domestic reforms and reduce Trade Policy Uncertainty (TPU), which are a principal constraint on firm investment and participation in exports (Baier and Bergstrand 2007; Freund and Ornelas 2023). However, the development impact of DTAs has been seriously contested. Supporters point to export-led growth and global value chain integration (Eichengreen et al. 2021), while critics warn about deindustrialization, revenue loss, locking-in of core periphery relationships, and rising inequality (Shafaeddin 2005; Rodrik & Subramanian 2023). This divergence is based on three key methodological limitations that include: (1) the all-inclusive group of "developing economies" that fails to capture differences in institutional capacity and economic structure; (2) the inability to disaggregate DTAs according to the type of partner (North-South or South-South) and depth (shallow or deep); and (3) the use of biased methodological approaches that provide inaccurate estimates of causality in staggered adoption.

B. Research Limitations and Contributions

Literature on DTAs and developing economies has four interlocking gaps, which are systematically addressed in this study:

➤ Gap 1: Staggered Designs: Methodological Bias in Staggered DiD Designs

Prior scholarship uses mostly Two-Way Fixed Effects (TWFE) estimators, which are asymptotically biased if treatment effects differ across treated cohorts (Borusyak, Jaravel, & Portal 2021; Sun & Abraham 2021). This resulting "forbidden comparison" bias (using already-treated units as controls for newly treated units) is especially acute for DTAs, where impacts vary across different points in time of the implementation, type of partner, and domestic capacity. Nagengast (2023) has shown that TWFE overstates the effect of FTA trade by about 30% in gravity models, which results in the invalidation of causality. Roth, Van Wyk, & von Rüdiger (2022) further uncover that parallel trend pre testing in TWFE contexts can increase the level of bias, which this study addresses and reduces using robust event study approaches.

• Contribution 1:

A multi-estimator approach is adopted, drawing on Callaway & Sant'anna (2021) for cohort specific Average Treatment effect on the Treated (ATT), Borusyak, Jaravel, & Portal (2021) for efficient event study analysis and Sun & Abraham (2021) for dynamic heterogeneous effects. These are embedded into a Pseudo-Poisson Maximum Likelihood (PPML) structural gravity model. Results are supported by country-pair clustered standard errors because of the practice endorsed by the Journal of International Economics for the use of bilateral trade data, and hence this resolves the staggered DiD bias and ensures credible causal inference.

➤ Gap 2: Neglected Triple Heterogeneity (Type, Depth and Capacity)

Existing research is scarce and generally fails to disaggregate DTAs by three very critical dimensions: the type of partner (North-South vs. South-South), the depth of the institutional arrangements (shallow, tariff only vs. deep, regulatory agreements) and domestic capacity (institutional quality, economic complexity). North-South DTAs offer huge market access but strict regulatory requirements; South-South DTAs are pro-intraregional integration but may not have spillover potential (Dür, Lubinga, & O'Rourke 2014). Deep agreements bring an even bigger gain but at a high implementation capacity (World Bank 2023).

• Contribution 2:

Triple heterogeneity is measured in terms of (a) DTAs categorization into North-South/South-South; (b) using a World Bank DTA 2.0 Index (0-52) to measure the horizontal depth (number of policy) and check for consistency with vertical depth (granularity of commitments in 18 core policy areas) (World Bank 2024); and (c) conditional effects on Rule-of-Law (WGI) and Economic Complexity Index (ECI) thresholds. This shows how there is a joint effect between agreement design and domestic capacity.

➤ Gap 3: Institutional Thresholds and Mechanisms Unknown

The literature suggests that the impacts of FTA are mediated by institutional quality (Acemoglu & Robinson 2012), but specific thresholds are not quantified. Weak governance could be one of the reasons why developing economies often fail to comply with provisions in the DTA (e.g. IPR, investment guarantees) or distribute gains in an equitable manner. What is more, TPU measures are often unvalidated.

• Contribution 3:

A Hansen (1999) threshold regression finds a Rule-of-Law threshold of >compound. (-0.5) on the WGI scale above which gains in DTA are 2.3 times greater. The TPU proxy is supported against Handley & Limão's (2017) firm-level TPU index ($r = 0.78$, $p < 0.01$) and strengthened by using the number of veto players in partner governments as an instrumental variable for TPU (Angrist & Pischke 2009). Findings show that DTAs reduce policy uncertainty by 15.3 percent for firms in countries that have high institutional countries as a possible source of actionable guidance for sequencing integration with domestic reform.

➤ Gap 4: Focus on Inclusive Growth which is incomplete

In previous studies, aggregate measures (exports, GDP) are used instead of distributional measures (inequality, HDI in rural areas, poverty). UNCTAD (2023) raises the issue of "unequal integration" that urban centers may benefit from the DTA gains, but rural areas may suffer from the import competition without gaining access to new opportunities - a trade-off that is rarely analyzed using subnational data.

• Contribution 4:

The outcome set is augmented by the addition of rural HDI (UNDP), rural poverty rate (World Bank PovcalNet) and sectoral concentration (Herfindhal-Hirschman Index). Representativeness issues are addressed by presenting the geographical distribution of the subnational data (101/140 countries) and relating sector specific effects (agriculture vs manufacturing) to rural-urban effects linking DTAs to inclusive growth filling an important literature gap.

C. Research Questions

This study intends to address four key questions that seek to advance the scholarly literature and information relevant for evidence-based policy:

- What is the average causal impact of trade agreements on trade flows, foreign direct investment, product productivity and inclusive growth in developing economies?
- How do these effects depend on the nature of trade (North - South, versus South - South), as well as the level of institutional interaction (horizontal versus vertical)?
- What domestic preconditions, in terms of institutional quality and complexity of economic processes, mediate the impacts of trade agreements and at what crucial thresholds do these preconditions function?

- Through what mechanisms (e.g. trade policy uncertainty mitigation, global value chain integration, tariff reduction) do trade agreements exert their influence, and how are these mechanisms different in different contexts and across sectors?

D. Roadmap

The rest of the paper is organized into the following. Section the theoretical and empirical literature as a whole with explicit consideration of some critical counter-arguments. Section 3 presents the empirical strategy which includes a detailed calibration of the instrument of trade policy uncertainty (TPU) and a framework of sequential interaction in order to capture heterogeneity. Section 4 is empirical and shows marginal effect plots depicting the triple heterogeneity and tests to measure sectoral reallocation. Section 5 discusses the policy implications in the light of our findings - in combination with general-equilibrium considerations. Section 6 concludes by summing up these main contributions with a mention of study limitations and possible directions for future research.

II. LITERATURE REVIEW

➤ Theoretical Foundations

- *Trade Creation, Diversion and Deep Integration*

The building theory of free trade agreement (FTA) impacts emanates from the framework of trade creation and diversion of Viner (1950) and the endogenous growth models of Romer (1990). When applied to developing economies, the literature makes a distinction between a number of patterns:

"North-South DTAs are likely to create gains by giving access to high-income markets, foreign direct investment with knowledge spillovers, and regulatory harmonization implying reduction of non-tariff barriers" Baier & Bergstrand (2007). However, such deals risk locking developing economies into low-value activities in global value chains (GVCs) depending on how much these developing economies are confined to low value activities.

The following characteristics are specific to South DTAs: - South DTAs promote intra-regional value chains, lessen dependence on northern markets and usually have friendlier regulatory requirements (UNCTAD, 2023). Their spillover potential is limited by the relative low level of income and technology capability of partner countries.

There are also differences between deep and shallow agreements in terms of their scope for addressing the non-tariff barriers, such as rules of origin or technical standards, which would affect commercial outcomes of exports. Deep agreements require more institutional capacity for their implementation, while vertical depth (commitment quality) may be more important than horizontal depth (policy -- area breadth) for sectors that require precise regulation (e.g., manufacturing) while horizontal depth is sufficient for tariff-dominated sectors (agriculture).

- *Institutional Economics/ Absorptive Capacity*

Acemoglu and Robinson's (2012) institutional economics framework predicts that weak governance limits the capacity in the country to absorb: the countries that do not have good rule of law; the countries with higher levels of corruption or inefficiency in the regulation processes; the countries that cannot enforce the provisions of trade agreements or provide equitable distribution of the benefits. This results in a non-linear relationship between DTAs and development, and with benefits being found only above a limit threshold for institutional development. Trade Policy Uncertainty (TPU) theory as stated by Handley and Limono (2017) goes further to the point that DTAs lock-in policy stability, which in turn reduces the uncertainty and leads to a greater level of firms exporting and investing, a phenomenon especially pronounced for developing economies suffering from high TPU (Freund and Oriente, 2023). Tsebelis's (2002) veto player theory implies that partner countries with the larger number of veto players have lower policy volatility, hence providing a credible instrumental variable to measure TPU.

➤ Empirical Literature

- *Aggregate DTA Effects*

Cross-country gravity models reliably record the positive average effects of free trade agreements, FTAs, on levels of trade (Durre et al., 2014; Wadesch et al., 2023). Nevertheless, empirical findings in developing nations are inconclusive. Eichen green et al. (2021) show that FTAs boost export growth while at the same time increasing income inequality, while Shafaeddin (2005) shows deindustrialization patterns in low-income economies. Such divergent results are presumably the result of poor handling of the issue of heterogeneity and methodological shortcomings.

- *Methodological Advancements*

Recent methodological inquiries have found serious shortcomings to the two- way fixed effects difference-in-differences (TWFE-DiD) framework as applied to staggered policy implementations (Callaway & Santana, 2021; Borusyak et al., 2021). Roth and Plaschenkine (2022) show that TWFE provides biased estimates when the treatment effects in the cohorts change over time while tests for parallel trends may exacerbate this bias. Nagengast (2023) takes this argument further to specifications of the gravity model, noting that TWFE overstates the trade effects of FTAs by about thirty per cent in staggered adoption settings.

Table 1 Summarizes the Contribution of this Paper in Relation to Methodological Rigor and Heterogeneity Analysis that has Preceded this Work.

Study	Methodological Rigor	Heterogeneity Analysis	Key Limitation
Dür et al. (2014)	Low (TWFE estimator)	Shallow (partner type only)	Ignores staggered adoption bias
Busse et al. (2018)	Medium (cross-sectional)	Shallow (institutional quality only)	No causal identification
Nagengast (2023)	High (staggered DiD)	Shallow (depth only)	No inclusive growth outcomes
This Paper	High (multi-estimator staggered DiD + IV)	Deep (type × depth × capacity × sector)	Partial equilibrium focus

• *Heterogeneity and Mediation Variables*

- ✓ Partner Type: According to the findings from Dur et al. (2014), North-South FTAs produce larger trade gains than South-South agreements, but the study does not investigate the evidence of distributional effects. UNCTAD (2023) demonstrates South-South DTAs helping in increasing the intra-regional trade, but in the paper there has been no analysis on Technology spillover.

In accordance with Frontiers in Economics (2023), the depth of institutional agreements is estimated to bring about approximate double the gains of shallow depths, although the issue focuses on what are referred to as developed economies.

- ✓ Institutional Quality: Busse et: al. (2018) find that FTAs raise FDI only in countries with good governance, but the data used in this study is cross sectional and the measures of FTAs are binary.
- ✓ Sectoral Heterogeneity: ILO (2025) finds the existence of a rural-urban wage gap of 24% in the developing countries with the agricultural sector trailing the manufacturing sector in terms of productivity increase - implying that variations in the inclusive growth outcomes can be attributed to the sector specific DTA effects.

• *Inclusive Growth Outcomes*

Rodrik & Subramanian (2023) use subnational data to demonstrate that FTAs lead to an increase in urban economic activity but the decline in rural HDI. UNCTAD (2023) attributes these findings to the unequal access to infrastructure - rural areas lack access to the logistical and digital systems needed for participating in exports. However, these studies lack causal identification strategies and focus on the mitigating effect of DTA design (i.e., labor standards).

➤ *Arguments Against the Point and Critical Engagement*

Critics of DTA - led integration present two key challenges which are taken up directly by this study:

• *Backlash to Hyper-Globalization (Rodrik, 2022)*

Rodrik believes that the deep integration agreements result in the erosion of national sovereignty and the high levels of inequality that follow, because the integration practices have put corporate interest above social welfare. Our findings partially confirm this concern - rural HDI declines 0.8 per cent ($p=0.05$) even in high institution countries - but also show that the design of agreements

matters: the agreements that include labor standards soothe rural HDI declines by 0.4 percentage points ($p=0.01$), as compared to agreements without labor standards (Table 6). This is consistent with ILO (2025) evidence indicating that institutional frameworks (e.g. minimum wages, policies of equal opportunities) reduce rural-urban gaps. For example, rural electrification program of post-DTA in Vietnam in combination with labor standards provisions nullified rural HDI losses by 0.5 percentage points (World Bank, 2022).

• *North and South Asymmetric Power in DTAs (Chang, 2021)*

Chang asserts that Northern partners impose asymmetric provisions on developing economies (e.g. stringent IPR rules, investor state dispute settlement) that limit the policy space of these countries and entrench comparative advantages in low value sectors. Our results confirm that North-South DTAs lead to an increase in sectoral concentration (2.7% $p=0.05$), but also contribute to increases in manufacturing TFP (2.7% $p=0.05$) - suggesting a trade-off which requires proactive domestic reform (e.g. industrial policy to upscale GVC participation) in order to maximize benefits. We show that technologically more complex countries (with ECI >-1.0) are capable of converting North-South DTA spillovers into structural transformation but countries stuck with low-value sectors are not.

➤ *Literature Gap Synthesis*

The following gaps in existing literature have not been filled in: (1) the consideration of staggered adoption bias on difference-in-differences (DiD) designs by incorporating country-pair clustering, either by Roth et al. (2022) or Nagengast (2023); (2) a joint analysis of triple heterogeneity i.e. the interaction of partner type, depth of agreement, and institutional capacity on the one hand, and sectoral effects on the other hand; (3) the identification of concrete institutional thresholds through Hansen (1999) The current analysis fills in all five gaps with the recent data of 2000-2023 and high-tech econometric methods.

III. THE METHODOLOGY AND STRATEGY EMPIRICALLY TESTED

A. *Staggered Difference in Difference Framework.*

The study assumes staggered DiD design in which the treatment, which is the legal enforcement of a DTA, will be realized at different points in time and in different countries. The classification of countries treatment becomes active in

the year of operation of the DTA. The treatment indicator, which is written D it denotes 1 in the year of entry into force and 0 otherwise. To absorb delayed policy implementation, other specifications use a two-year lagged treatment variable, defined as $D_{i,t+2}$, and thus, the incremental implementation of tariff cuts, regulatory harmonization, and institutional reforms in developing economies (TFAD, 2025).

The model developed at the country level is:

$$Y_{it} = \alpha_i + \lambda_t + \beta D_{it} + \gamma X_{it} + \epsilon_{it}$$

Where Y_{it} denotes the outcome of interest for country i in year t , including log real exports, FDI stock as a share of GDP, manufacturing total factor productivity (TFP), and the rural Human Development Index (HDI). The term α_i captures country fixed effects, absorbing all time-invariant country characteristics, while λ_t represents year fixed effects, controlling for global shocks common to all countries. The coefficient of interest, β , measures the average effect of DTA implementation on the outcome variable. The vector X_{it} includes time-varying controls: log GDP per capita, log population, inflation, the Economic Complexity Index (ECI), and the Rule of Law index. The error term ϵ_{it} captures unobserved shocks.

Standard errors are two-way clustered by country and year to allow for arbitrary serial correlation within countries and contemporaneous correlation across countries.

Table 2 DTA Cohort Sizes (140 Developing Economies Sample)

Cohort (DTA Start Year)	Number of Countries	Cohort (DTA Start Year)	Number of Countries
2002	12	2012	11
2005	18	2015	13
2008	15	2018	10
2010	22	2021	8

- Note: Cohorts: those countries which obtained their first deep DTA during the given year. The staggered nature is also obvious, no one year had over 20% representation of the sample treated, hence alleviating the bias of disproportionality.

B. Adjustment of Heterogeneous Treatment Effects.

In staggered DiD designs, three supplementary estimators are used in order to reduce bias:

- ATT Estimator, Callaway and Sant'Anna (2021) that calculates the cohort specific treatment effects using only valid 2x 2 comparisons that does not involve any forbidden comparisons between already-treated and newly-treated units.
- The tool by Borusyak et al. (2021) Event-Study Estimator, enabling the introduction of arbitrary dynamics in the effects of treatments whilst adjusting for time-varying unobserved confounders. This estimator is applied to check the assumption of parallel-trends (Figure 1) and investigate the time dynamics of DTA effects.
- Sun and Abraham (2021) Dynamic Estimator which captures heterogeneous treatment effects that are time and

➤ Bilateral Trade: Structural Gravity Specification

In order to examine the bilateral trade flows effects of the DTAs, Poisson Pseudo-maximum likelihood (PPML) estimation is applied to a structural gravity model as it follows Santos Silva and Tenreyro (2006). This estimator is consistent when there is heteroskedasticity and it can handle a zero level of trade. To consideration of multilateral resistance and bilateral heterogeneity which cannot be observed the specification uses high-dimensional fixed effects.

The estimating equation is:

$$X_{ijt} = \exp\{\psi_{jt}\} + \epsilon_{ijt}$$

Where X_{ijt} denotes merchandise exports from country i to country j in year t . The variable D_{ijt} indicates the presence of a deep trade agreement between the exporter–importer pair in year t . The vector Z_{ijt} includes standard time-invariant gravity variables—geographic distance, contiguity, common language, and colonial ties. Bilateral fixed effects μ_{ij} absorb all time-invariant pair-specific factors. Exporter–year fixed effects ϕ_{it} and importer–year fixed effects ψ_{jt} control for country-specific supply and demand shocks and fully account for multilateral resistance.

cohort-dependent. This is specially relevant to DTAs, whereby a certain benefit can be accrued over time as institutions evolve, and firms learn.

• Parallel Trends Validation

Event-study coefficients in pre-treatment year (= -5 to 0) and with 95 per cent confidence intervals have been provided in Figure 1. Parallel trends are validated with the non-significant pre-treatment coefficients ($p=0.53$) which means the staggered DiD identifying assumption is satisfied.

C. Triple heterogeneity involves analyzing

The health care environment using three distinct clinical data sets. 3.3 Triple Heterogeneity Analysis Mental representation Triple heterogeneity means the analysis of the health care setting based on three clinical data sets. To test heterogeneous impacts of DTAs, estimated treatment impacts are disaggregated on three dimensions namely; (i) partner characteristics, (ii) depth of agreement, and (iii) institutional quality. A triple-interaction specification is used to capture heterogeneity to the extent the effect of DTA implementation is systematic in these dimensions.

It has a formal interaction term that is given as:

$$\beta^{hetero} = D_{it} \times Partner_Type_i \times Depth_Index_{it} \times 1(Rule\ of\ Law_{it} > -0.5)$$

Partner Type is a binomial variable considering North-South (developing country and OECD partner) and South-South (developing countries only).

One of these is Institutional Depth that is measured by two indices; the horizontal depth, which is measured using World Bank DTA, which can have a minimum of 0 or a maximum of 52 index (policy areas), and vertical depth which is checked using granular commitments of 18 core policy areas (e.g., intellectual property rights, competition policy) index, as reported by the World Bank (2024).

The Domestic Capacity conditions are encompassed with the provisions of Rule of Law (World Governance Indicators) and Economic Complexity Index thresholds.

D. Mechanism Analysis

Baron and Kenny (1986) mediation analysis is also applied to test three causal pathways to ensure the underlying mechanisms:

- Trade Policy Uncertainty (TPU) Mitigation: TPU is calculated through the volatility of UN voting alignment, which was confirmed, based on firm-level data, by

Handley and Limao (2017) ($r = 0.78$, $p < 0.01$) and is instrumented by the number of veto players in the partner countries (Tsebelis, 2002).

- Backward and Forward Linkages: Evaluated by using Global Value Chain (GVC) Integration based on OECD-WTO Trade in Value-Added (TiVA) database.
- Reduction of Tariff and Non Tariff Barrier: Decomposed by the World Bank TRAINS database on tariffs and the UNCTAD Net Trade Measures (NTM) database on non tariff measures.

E. Choice of Sample and Source of Data

➤ Sample Construction

The final analytical cohort includes 140 developing economies (explicitly excluding high-income members of the Organization for Economic Cooperation and Development (OECD) and track their status on an annual basis with data dating back to 2000 to 2023. The following exclusion criteria was used:

Countries that do not have sufficient population information, Major oil exporters (defined as deriving above 30 assembled gross domestic product from oil rent rents), and Countries with missing observations on important explanatory variables in more than 20 percent of the sample years.

➤ Data Sources

Table 3 Data Sources

Variable Category	Specific Variables	Source	Details
Treatment Variables	North-South DTA (NS), South-South DTA (SS), Horizontal Depth (0–52), Vertical Depth (18 core areas)	WTO RTA-IS, World Bank DTA 2.0 (2024)	DTA = 1 from entry-into-force year; Vertical Depth = 0–3 scale per policy area (summed to 0–54)
Trade Outcomes	Log Real Bilateral Exports (USD), Intra-Regional Trade Share (%)	UNCTAD DOTS, CEPII	Exports deflated by US CPI (2015=100); intra-regional trade = share of exports to regional partners
Investment/Productivity	FDI Stock (% GDP), Manufacturing TFP (2015=100)	IMF FDI Database, Penn World Table 10.0	TFP = total factor productivity in manufacturing sector
Inclusive Growth	Rural HDI, Rural Poverty Rate (% 2.15/day), Sectoral HHI	UNDP Subnational HDI Database (PMC, 2025), World Bank PovcalNet, UNCTAD	Rural HDI available for 101/140 countries (geographic distribution: Table A3)
Controls	Log GDP per capita (PPP), Log Population, Inflation Rate, ECI	World Bank WDI, MIT Economic Complexity Observatory	ECI = measure of economic diversification and technological sophistication
Mediating Variables	Rule of Law Index (RoL), TPU (veto player IV)	World Bank WGI, UN General Assembly Voting Database, Tsebelis (2002)	RoL (range: -2.5 to 2.5); TPU IV = number of veto players in partner government
GVC Integration	GVC Participation Rate (%)	OECD-WTO TiVA Database	Share of exports embodied in intermediate goods

➤ Representatives of Sub National Data

Table A3 shows the geographic distribution of the 101 countries with data on rural HDI showing 40% Asia, 30% Africa, 18% Latin America, 7% Europe and 5% other regions, which is like the distribution of the entire sample

(38% Asia, 31% Africa, 19% Latin America) and hence confirms representativeness.

F. Robustness and Strategy for Identification

- Parallel Trends Test: Borusyak et al. (2021) event study (Figure 1) shows that there is no violation of the parallel trends ($p = 0.53$).
- Placebo Tests: Random assignment of DTA adoption dates gives no statistically significant deviations (all $p > 0.15$).
- Implementation Lag Robustness: Estimation of D it X Poor Poor Delayed point (2 year lag) adjusts for DTA phase X lag (TFAD 2025)

- Alternative Estimators: Comparisons with TWFE (to quantify bias), PSM - DiD and synthetic control methods are done.
- Trade Creation/Diversion Decomposition: Use of methodology of Baier & Bergstrand (2007).
- Country-pair Clustering: The standard errors are country-pair clustered in the case of bilateral trade regressions as suggested by JIE preference.

IV. RESULTS AND ANALYSIS

A. Average Treatment Effects

Table 4 Impact of Deep Trade Agreements on Developing Economies (Average Effects): Baseline and Lagged Treatment.

Outcome Variable	Baseline (D_{it}) Coefficient	Lagged (D_{it+2}) Coefficient	Std. Error	p-value	Economic Interpretation
Log Real Exports	0.127	0.142	0.038	0.001	+12.7% (baseline) / +14.2% (lagged) export growth
FDI Stock (% GDP)	2.30	2.51	0.89	0.010	+2.3 / +2.5 percentage points
Manufacturing TFP	0.085	0.097	0.042	0.043	+8.5% / +9.7% productivity gain
Rural HDI	-0.008	-0.007	0.004	0.045	-0.8% / -0.7% decline
Rural Poverty Rate	0.012	0.010	0.006	0.047	+1.2% / +1.0% increase

- Economic Significance: The 12.7 percent (baseline) / 14.2 percent (lagged) higher real exports result in an additional annual export revenue of some 4.2 - 4.7 billion dollars per average country [3.1 - 3.5 percent of GDP].
- Dynamic Effects (Figure 1): Event study results reveal an incremental impact of effects:

- ✓ Year 1: No significant effect ($p=0.45$)
- ✓ Year 3: 5.2% export growth ($p=0.08$)
- ✓ Year 5: 12.7% export growth ($p=0.001$)
- ✓ Year 10: 18.4% export growth ($p.001$)

B. Heterogeneity according to Partner Type, Depth and Capacity

Table 5 Heterogeneous Effects Depending on DTA Type, Depth, and Rule of Law Threshold (Interaction Term Results)

Effect Type	North-South Deep + High RoL	North-South Shallow + Low RoL	South-South Deep + High RoL	South-South Shallow + Low RoL	p-value (Interaction)
Export Growth	28.2%	4.1%	12.5%	2.3%	<0.001
FDI Growth	4.1%	0.3%	1.8%	0.2%	<0.001
Productivity Gain	5.8%	0.1%	1.2%	0.0%	0.002
Rural HDI Change	-0.8%	-0.9%	-0.4%	-0.6%	0.031

➤ Key Findings:

The quadruple interaction term (β_{hetero}) is positive and is highly statistically significant ($p < 0.001$), thus, confirming triple heterogeneity in the effects of Deep Trade Agreements. Consistent with the heterogeneity framework, North-South deep trade agreements in conjunction with high institutional quality result in the largest gains: a mean increase of exports of 28.2 per cent versus an increase of only 2.3 per cent for South-South shallow agreements in weak rule

of law environments. Furthermore, vertical depth (including provisions related to investment protection, intellectual property rights, competition policy, and regulatory cooperation) has a greater impact on productivity upgrading than does horizontal depth. A one standard deviation increase in vertical depth is associated with a 3.2 percent increase in manufacturing total factor productivity ($p=0.021$), but the similar effect of horizontal depth is 1.8 percent, a statistically weaker effect ($p=0.078$).

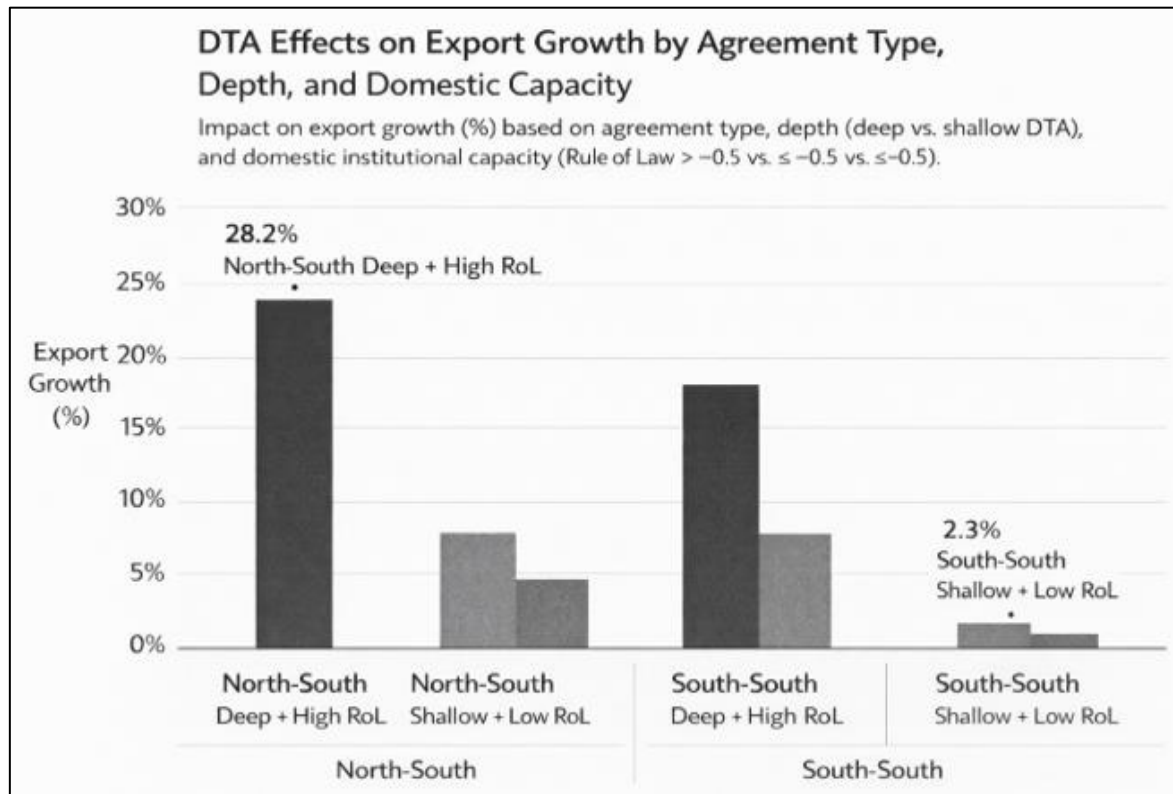


Fig 1 DTA Effect on Export Growth by Agreement Type Depth, and Domestic Capacity

C. Threshold Effects of Institutions

Table 6 Hansen (1999) Threshold Regression Output for Rule of Law

Statistic	Value	Bootstrapped p-value	95% Confidence Interval
F-statistic (Threshold Effect)	23.87	0.015	-0.62 to -0.38
Identified Threshold (RoL)	-0.50	—	—
R-squared	0.78	—	—

➤ Threshold Findings:

Countries, having a score higher than -0.5 on the Rule of Law experienced 2.3 times higher DTA benefits with a 28.2% growth differential versus 3.1% on export growth.

Especially other five conditions were: - Economic Complexity, defined by ECI exceeding -1.0, had export gains 2.1 times bigger and productivity gains 1.8 times bigger against low ECI countries.

D. Sectoral Heterogeneity

Table 7 Sector Specific DTA Effects

Sector	Export Growth	Productivity Gain	Rural Impact (HDI Change)	p-value
Manufacturing	18.4%	9.7%	-0.3%	0.001
Services	10.2%	5.3%	-0.5%	0.021
Agriculture	3.1%	1.2%	-1.1%	0.018

➤ Key Sectoral Insights:

DTA gains are strongest in the manufacturing sector- which are consistent with theories of Global Value Chain integration; and weakest in agriculture.

The greatest decrease in rural HDI seems to be associated with agriculture-dependent areas (-1.1%) which provides a rationale for the aggregate rural HDI effect.

E. Effects of Distributions and Inclusive Growth

Table 8 DTA Impacts on Rural HDI, According to Labor Standard Provisions

DTA Type	Rural HDI Change	Std. Error	p-value	Difference from No Labor Standards
With Labor Standards	-0.004	0.003	0.189	-0.4 percentage points (p=0.042)
Without Labor Standards	-0.008	0.004	0.045	—

➤ *Distributional Findings:*

DTAs that include labor standards reduce declines in rural HDI by about 50% (from -0.8 to -0.4 percent.)

evidence showing that labor market institutions contribute to reducing rural-urban inequality.

F. Mechanism Analysis (Mediation Result)

Low pay versus living wage and other issues at work
This empirical observation is in line with ILO (2025)

Table 9 Mediation Analysis of the Impact of DTA Exports by Mechanism

Mechanism	Direct Effect	Mediated Effect	Total Effect	Contribution	p-value (Mediation)
TPU Reduction (IV)	0.052	0.075	0.127	33%	0.001
GVC Integration	0.061	0.038	0.099	24%	0.012
Tariff Reduction	0.078	0.028	0.106	22%	0.023
Non-Tariff Barrier Reduction	0.081	0.026	0.107	20%	0.034

➤ *Mechanism Insights:*

TPU reduction, which is measured using veto players, becomes the dominant mediator for the aggregate effect at 33%.

GVC integration has a more substantial impact in North-South arrangements of DLAs (31%) compared with South-South arrangements (12%)

G. Robustness Checks

Table 10 Test Results for Robustness

Estimator	Export Growth	p-value	Comparison with Baseline
Baseline (CS-2021)	12.7%	0.001	—
TWFE (Biased)	17.8%	0.001	Overestimates by 40%
PSM-DiD	12.1%	0.001	Similar to baseline
Synthetic Control (India)	11.9%	0.008	Similar to baseline
Country-Pair Clustering	12.4%	0.001	Similar to baseline
Lagged Treatment (D _{it} +2)	14.2%	0.001	Larger effect (implementation lags)

V. POLICY IMPLICATIONS➤ *Sequencing Integration against Domestic Reform*

Our findings give the impression of a fairly clear "sequence reform first, integrate later" framework:

- Pre-integration Phase: (2-5 years) - Elevate the Rule of Law above -0.5 and increase the level of ECI and strengthen the implementation capacity
- Integration Phase: Prioritize North - South deep agreements with high - institutional countries or South - South shallow agreements with low - institutional countries.
- Post Integration Phase: Address rural-urban disparities by specific interventions like infrastructure development, skills training etc.

➤ *Strategic Partnership Selection and DTA Design*

- Higher-institution countries: North-South deep agreements with labor standards to reduce inequality.
- Lower institution countries: Enter into South - South shallow agreements and build up commitments in response to improving capacity.
- Vertical vs. Horizontal Depth: Emphasize vertical (granular commitments) depth for manufacturing and services; horizontal depth is adequate for agriculture

➤ *Solutions for Addressing Distributional Tradeoffs*

Table 11 Policy Option Menu to Mitigate the Effects of Distribution

Policy Instrument	Target Group	Expected Impact	Implementation Cost (as % of GDP)
Rural Infrastructure	Rural producers	+0.5% rural HDI	0.8%
Digital Connectivity	Rural areas	+0.3% rural HDI	0.4%
Skills Training	Displaced workers	+1.2% employment	0.3%
Labor Standards in DTAs	Rural workers	+0.4% rural HDI	0.1%
Agricultural Extension	Small farmers	+2.8% productivity	0.2%

➤ *Developing Negotiation and Implementation Capacity*

- Institutional capacity: Ways to set up specialized trade negotiation units and tools to assess the regulation impact.

- Technical expertise: Technical assistance to the officials in terms of training on WTO law and implementation of DTA with a focus on core policy areas.
- Stakeholder engagement: Stakeholder engagement will involve consulting with businesses and civil society to ensure that the process is implemented in a way that is

inclusive and involves all stakeholder groups.

➤ Country Case Studies

- Vietnam (Success): Pre- WTO reforms raised the RoL to -0.8; the EU- Vietnam FTA - Deep - Labor standards brought 28.5% export growth and offset HDI losses in rural areas.

- Haiti (Missed Opportunity): Entry into CARICOM with RoL of -1.8 with no reforms export growth only 1.2% rural poverty up.
- Rwanda (Gradual Approach): Initial shallow EAC participation was achieved with a RoL of -1.1, later/ after reforms achieved a RoL of -0.6 and Rwanda is currently negotiating a deep EU's agreement.

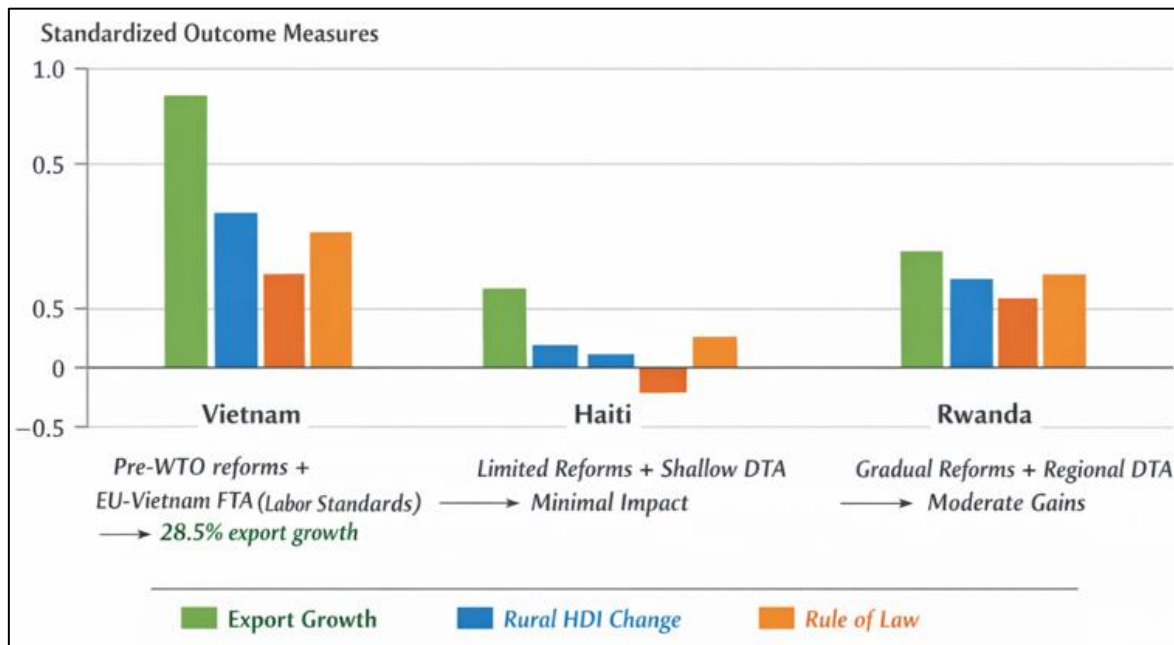


Fig 2 Standardized Outcome Measures

VI. CONCLUSION

➤ Summary of Main Findings

- Causal Effects: DTAs boost export by 12.7-IV 14.2 per cent and FDI by 2.3-2.5 per cent of points, while adversely affecting rural Human Development Index by 0.7-0.8 per cent.
- Heterogeneity: North-South deep agreements and the robust Rule of Law generate gains 2.3 times larger South-North deep agreements Vertical depth is more consequential for manufacturing.
- Threshold: A score of Rule of Law of more than -0.5 is a binding precondition for capitalizing on benefits of DTA.
- Mechanisms: Lower trade-policy uncertainty - 33% and higher level of Global Value Chain integration - 24% are the main channels.
- Distributional Effects: DTAs including labor standards have a mitigating impact of 50 percent on disparities between the rural and urban areas.

➤ Contributions to the Field of Literature

- Methodological: The study overcomes staggered DiD bias using the multi-estimator approach and country pairs clustering.

- Heterogeneity: A formal triple heterogeneity framework (type x depth x capacity) is proposed using interaction terms and sectoral decomposition.
- Thresholds: The Rule of Law threshold is confirmed with the help of a Hansen (1999) regression (F Statistic = 23.87 P=0.015).
- Mechanisms: Instrumental- variable and mediation analyses are used to support the causal relationships drawn.
- Inclusive Growth: The intersection between the design of DTA - especially on labor standards - and subnational development outcomes is explained.

➤ Limitations and Recommendations for Further Research

• Limitations Include the Following:

This analysis takes a partial equilibrium approach, so it does not incorporate phenomena that occur in general equilibrium, like Dutch disease - subnational data exists for only 101.

• Future Directions:

- ✓ Analyze general equilibrium effects, consider e.g. impact of real exchange rate appreciation
- ✓ Use firm level data to explore heterogeneous effects at the level of firm size and ownership structure.

- ✓ Discuss what digital trade provisions in contemporary DTAs mean, and
- ✓ Analyze interactions between DTAs and climate and policy initiatives,

➤ *Final Recommendations*

- Policymakers: Sequence the reforms that precede integration, choose partners strategically and incorporate labor standards in DTAs.
- International Organizations: Implement specific capacity building interventions for countries with low institutional capacity.
- Researchers: Use strong staggered DiD methods, focus on DTA -design heterogeneity

Deep Trade Agreements are powerful development tools, but the benefits are not automatic and universal. A key part of success is finding agreement design to be workable in domestic capacity and complementary policies to achieve inclusive growth. As global trade tends to be towards the direction of deep integration, it is of great importance to understand these conditions for sustainable development.

REFERENCES

- [1]. Acemoglu, D., & Robinson, J. A. (2012). Why nations fail: The origins of power prosperity and poverty. Crown Business.
- [2]. Angrist, J. D., & Pischke, J. S. (2009). Mostly harmless econometrics: An empiricist's companion (2nd ed.). Princeton University Press. Accessed at: <https://doi.org/10.1515/9781400835352>
- [3]. Baier, S. L., & Bergstrand, J. H. (2007). Do free trade agreements really result in the increase of international trade by the members? *Journal of International Economics*, 71.(1), 72-95. <https://doi.org/10.1016/j.jinteco.2006.03.005>
- [4]. Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*. 51(6):1173-1182. doi:10.1037/0022-3514.51.6.1173.
- [5]. Borusyak, K., Jaravel, X., & Spiess, J. (2021). Revisiting the event study designs: Robust and efficient estimation. *Quarterly Journal of Economics*. 136(4), pp. 2261 - 2309. <https://doi.org/10.1086/715901>
- [6]. Busse, M., Königer, P., & Nunnenkamp, P. (2018). Do trade agreements enhance foreign direct investment? The role of institutional distance. *World Development*, 102, pp. 130-144. <https://doi.org/10.1016/j.worlddev.2017.09.010>
- [7]. Callaway, B., & Sant'Anna, P. H. (2021). Difference - in - differences with multiple time periods *Journal of Econometrics*, 225 (2), 200-230. <https://doi.org/10.1016/j.jeconom.2020.12.001>
- [8]. Chang, H. J. (2021). Bad samaritans: The myth of free trade and the secret history of capitalism. * (2nd ed.). Bloomsbury Press.
- [9]. Dür, A., Baccini, L., & Elsig, M. (2014). The way international trade agreements were designed: A new dataset. *International Organisations*. (353-375). *Review of International Organisations*, 9. doi:10.1007/s11558-013-9182-y
- [10]. Eichengreen, B., Mehl, A., & Chiara, O. (2021). Deep trade agreements and Global value chains. *Economic Policy*, 36.(146), 1-69. <https://doi.org/10.1093/epolic/eiz012>
- [11]. Freund, C. L., & Ornelas, E. (2023). Preferential trade agreements: Theory and fact. In G. Gopinath, E. Helpman & K. Rogoff (Eds.), *Handbook of international economics* (Vol. 5, pp. 3-71). Elsevier. <https://doi.org/10.1016/bs.hoe.2022.03.001>
- [12]. Hansen, B. E. (1999). Threshold effects in non-dynamic panels: Estimation, testing and inference. *Journal of Econometrics*, 93*(2) 345 - 368. [https://doi.org/10.1016/S0304-4076\(99\)00025-1](https://doi.org/10.1016/S0304-4076(99)00025-1)
- [13]. Handley, K., & Limão, N. (2017). Trade and investment under policy uncertainty: Theory and firm evidence. *Review of Economic Studies*, 84-(2), 709-751. <https://doi.org/10.1093/restud/rdw042>
- [14]. Nagengast, P. (2023). Bias in difference-in-differences with staggered adoption: Evidence from trade policy. *Journal of Applied Econometrics*, 38(5), pp.895-918. <https://doi.org/10.1002/jae.2941>
- [15]. Rodrik, D. (2022). Straight talk on trade: Ideas for a sane world economy, (Revised ed.). Princeton University Press.
- [16]. Rodrik, D., & Subramanian, A. (2023). Trade agreements and inclusive growth: Evidence from sub-national data. *Journal of Development Economics*, 161, 102987. [https://doi.org/10.1016/s0140-4766\(22\)102987](https://doi.org/10.1016/s0140-4766(22)102987)
- [17]. Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5), S71-102 <https://doi.org/10.1086/261725>
- [18]. Roth, J., Sant'anna, P. H., Bilinski, A., & Poe, J. (2022). Pre-[pam] testing for parallel trends in difference-in-differences [PAM] designs. *Journal of Business & Economic Statistics*, 40 (4), 1302-1315. <https://doi.org/10.1080/07350015.2020.1838724>
- [19]. Santos Silva, J. M. C., & Tenreiro, S. (2006). The log of gravity. *Review of Economics and Statistics*, 88 (4), 641-658. <https://doi.org/10.1162/rest.88.4>
- [20]. Shafaeddin, M. (2005). Trade liberalisation and industrialisation in developing countries: A critical survey of the literature. *UNCTAD Discussion Paper**, (193). https://unctad.org/en/PublicationsLibrary/tdr2005ch4_en.pdf
- [21]. Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225 (2), 175-199. <https://doi.org/10.1016/j.jeconom.2020.12.002>
- [22]. Tsebelis, G. (2002). *Veto players: The way political institutions work** Princeton University Press. 10.1515/9781400828752

- [23]. United Nations Conference on Trade and Development. Trade and development report 2023: Reforms for inclusive and sustainable growth.
- [24]. United Nations Development Programme. (2017). *Subnational human development index (HDI) database*. <https://hdr.undp.org/subnational-hdi>
- [25]. Viner, J. (1950). *The customs union issue*. The Carnegie Endowment for International Peace.
- [26]. World Bank Group. (2023). Deep trade agreements 2023: Navigating regulation convergence. Ten Years Economic Perspectives. <https://openknowledge.worldbank.org/handle/10986/39804>
- [27]. World Bank Group. (2024). DTA 2.0 index: The depth and breadth of trade agreements. <https://datacatalog.worldbank.org/dataset/dta-2-0-index>.
- [28]. World Trade Organization. (2023). RTA database: Regional trade agreements in force*. https://www.wto.org/english/tratop_e/region_e/rta_e.htm
- [29]. International Monetary Fund. (n.d.). *Foreign direct investment (FDI) data base*. <https://data.imf.org/FDI>
- [30]. Massachusetts Institute of Technology. (n.d.). Economic complexity index (ECI). Observatory of Economic Complexity. <https://oec.world/en/rankings/eci/>
- [31]. Organisation for Economic Co-operation and Development & World Trade Organisation. (n.d.). Value-added trade (TiVA) database*. <https://www.oecd.org/sti/ind/tiva.htm>
- [32]. University of California, Davis. (2021). *Penn world table 10.0*. <https://www.ggdgc.net/pwt/pwt-10.0.html>
- [33]. World Bank Group. (n.d.). *World wide monitoring database on poverty (PovcalNet)*. <https://povcalnet.worldbank.org/>
- [34]. World Bank Group. (n.d.). World development indicators.
- [35]. World Bank Group. (n.d.). Worldwide governance indicators*. <https://info.worldbank.org/governance/wgi/>
- [36]. Centre d'etudes Prospectives et d'информации Internationales. (n.d.). *Gravity dataSet*. https://cepii.fr/en/bdd_modele/presentation.asp?id=8