

A Network Approach to Understanding Global Trade Vulnerabilities and Resilience

Gino Franco Fazzi

Advisor: Michele Coscia
Submitted: June 2025

IT UNIVERSITY OF COPENHAGEN

Acknowledgements

This thesis marks the end of a long (and somewhat unexpected) journey, and I couldn't have made it without help along the way. First and foremost, I would like to express my deepest gratitude to my supervisor, Michele Coscia, for his invaluable guidance and encouragement throughout this research. His expertise in network analysis and his insights into economics have been greatly appreciated. I'm thankful for our many conversations, his careful reading of countless drafts, and the constructive, thoughtful feedback he consistently offered. I am also profoundly grateful to my family and to my partner, Cookie, for your unconditional love, patience, and unwavering belief in me. I truly appreciate your feedback, insights, and perspective. I know reading such a long work on something outside your domain has not been easy (nor fun). Thank you for standing by me every step of the way.

Abstract

In an era characterized by hyper-globalization, where economic systems are highly interdependent, renewed geopolitical tensions and rising protectionist policies have underscored the fragility of interconnected economic structures. A recurring and significant challenge within this context is the disruption of global supply chains and their cascading impacts on international trade. Traditional econometric approaches have typically examined international trade through bilateral relationships and theoretical frameworks, inadequately capturing the systemic complexity and structural vulnerabilities inherent in the global trade network. More recent research efforts have leveraged network analysis tools to unveil emergent properties of international trade relationships. Recent developments in deep learning offer new opportunities for analysing such complex network structures. This dissertation addresses how can systemic vulnerabilities within international trade networks be effectively identified and predicted at the country level. Specifically, it frames country-level vulnerability assessment as a node classification problem, employing several GNN architectures, including Graph Convolutional Networks (GCN), Graph Attention Networks (GATv2), and GraphSAGE. Additionally, this research explores various ways of representing international trade as multilayered graphs, leveraging innovative embedding techniques to integrate commodity-specific information and geographical characteristics into the node representations. The results demonstrate that incorporating network topology improves predictive accuracy compared to purely attribute-driven baselines. However, performance gains remain constrained by the inherent complexity of accurately identifying and predicting economic vulnerabilities. Beyond the methodological and empirical contributions, this work critically discusses limitations related to data quality, classification thresholds, and methodological choices, highlighting the nuanced challenges that must be addressed in future research. Overall, this research provides a comprehensive methodological framework and practical insights into identifying systemic vulnerabilities in global trade, offering policymakers valuable tools for enhancing economic resilience. It also outlines clear avenues for further exploration, particularly in the integration of explainability and uncertainty quantification, to strengthen the interpretability and practical applicability of GNN-based predictive models for this problem.

Contents

Acknowledgements	iii
Abstract	v
Contents	vii
Nomenclature	1
1 Introduction	5
2 Related Work	9
2.1 Traditional Approaches	10
2.2 Network Analysis	11
2.3 Graph Neural Networks	13
2.4 Summary	15
3 Data	17
3.1 UN Comtrade Database	18
3.2 Atlas of Economic Complexity	21
3.3 Database for International Trade Analysis (BACI)	22
3.4 Additional Data Sources	22
4 Methods	25
4.1 International trade as a network	25
4.2 Multilayered network	26
4.3 Node labels	44
4.4 Node Classification	47
4.5 Experimental setup	56
4.6 Evaluation metrics	58

5	Results	59
6	Discussion	65
6.1	Data Limitations	65
6.2	Design choices	67
6.3	Results Analysis	80
6.4	From Node-Level to Network-Level Vulnerability	84
6.5	Uncertainty	85
6.6	Explainability	86
7	Conclusions	89
	Appendix	92
A	UN Comtrade Data	93
A.1	HS2012 Commodity Nomenclature	93
A.2	Data tables: Reporters	95
A.3	UN Comtrade API	101
A.4	UN Comtrade Tariff Line Raw Data	101
A.5	Countries with no bilateral data	106
B	Atlas of Economic Complexity dataset	107
B.1	Data tables: Trade Data	108
B.2	Data tables: Countries	108
B.3	Data tables: Products	113
B.4	Smoothing discrepancies:	113
B.5	Economic Complexity Index (ECI)	113
B.6	Product Complexity Index (PCI)	115
B.7	Complexity Outlook Index	115
C	BACI dataset	117
D	Additional Data Sources	119
D.1	Global Preferential Trade Agreements Database	119
D.2	Discrepancy Index	120
D.3	Corruption Perception Index	120
D.4	Country Borders	124
D.5	CERDI Maritime Distance Database	129
E	Data Source Integration Diagram	131

Contents	ix
F Optuna	133
F.1 Hyperparameter tuning results	135
G Computational Resources	137
H Use of AI Assistance	139
Bibliography	141

Nomenclature

Acronyms & Abbreviations

AINET	Artificial Intelligence Network Explanation of Trade
BACI	International Trade Analysis Database (FR: Base pour l'Analyse du Commerce International)
CEPII	French center for research and expertise on the world economy (FR: Centre d'Etudes Prospectives et d'Informations Internationales)
CIF	Cost, Insurance and Freight
CNA	Complex Network Analysis
CNN	Convolutional Neural Network
COI	Complexity Outlook Index
CPI	Corruption Perceptions Index
ECI	Economic Complexity Index
FOB	Free on Board
FTA	Free Trade Agreement
GAE	Graph Autoencoder
GAT	Graph Attention Network
GC-LSTM	Graph Convolutional Long Short-Term Memory
GCN	Graph Convolutional Network

GDP	Gross Domestic Product
GGAE	Gravity-informed Graph Autoencoder
GNN	Graph Neural Network
HHI	Herfindahl–Hirschman Index
HS	Harmonized System
ITN	International Trade Network
MLP	Multi-Layer Perceptron
NLP	Natural Language Processing
PCI	Product Complexity Index
RCA	Revealed Comparative Advantage
ReLU	Rectified Linear Unit
RNN	Recurrent Neural Network
SITC	Standard International Trade Classification
SRCA	Smoothed Revealed Comparative Advantage
TI	Trustworthiness Index
UMAP	Uniform Manifold Approximation and Projection
UN	United Nations
VGAE	Variational Graph Autoencoder
WTO	World Trade Organization

Glossary

Affected Importer	A country deemed as affected by a <i>lost exporter</i> .
Autarky	Economic independence or self-sufficiency.
Commodity	In this work, any tradeable good classified under the HS2012 system, including both raw materials and manufactured products. Used interchangeably with “product” for simplicity.

Commodity class	First 2 digits of the HS2012 system.
Edge or Link	A connection between nodes of a network. In this work, a trading flow (exports, imports, or both).
Graph	Representation of a network. In this work, the representation of the international trade flows (i.e., <i>links, edges</i>) between countries (i.e., <i>nodes, vertices</i>).
Lost Exporter	A country deemed exporter at year $n - 1$, which lost its "exporter" status at year n .
Node or Vertex	An <i>actor</i> or <i>agent</i> in a network. In this work, countries.
Trade Flows	Movement of goods and services between countries.

Mathematical & Graph Notations

G	Graph
V	Set of nodes or vertices
$ V $	Number of nodes or vertices
u, v	Single node or vertex
E	Set of edges or links
$ E $	Number of edges
e	Single edge
$\mathcal{N}(u)$	Neighbourhood of u , i.e., the nodes directly connected to u
L	Set of commodity layers
l	Single commodity layer
\mathbf{A}	Adjacency matrix
$\tilde{\mathbf{A}}$	Adjacency matrix (\mathbf{A}) with self-connections

D	Diagonal degree matrix
\tilde{D}	Diagonal degree matrix (D) with self-loops
k	Superscript denoting the k^{th} layer of a GNN architecture
\mathbf{h}^k	Hidden state \mathbf{h} in layer k of a GNN, where \mathbf{h}^0 correspond to the input layer (node features)
\mathbf{W}^k	Layer-specific trainable weight matrix for layer k of a GNN
$\sigma(\cdot)$	Non-linear activation function

Chapter 1

Introduction

“In a world deluged by irrelevant information, clarity is power.”

Yuval Noah Harari

The past century has witnessed an unprecedented globalization of the world economy, binding nations together through intricate commercial relationships. Economies have become interwoven to such a degree that isolated analyses are increasingly insufficient and potentially misleading. As Yuval Noah Harari aptly states, *“The whole of humankind now constitutes a single civilization, with all people sharing common challenges and opportunities.”* [1]

Recent decades have revealed that global trade is not distributed evenly but concentrated within a few central hub countries that dominate the international trade network (ITN) [2]. Economic interconnections have not only intensified in volume but also evolved structurally, making global trade networks denser and more complex [3]. However, while traditional macroeconomic tools remain useful for broad economic assessments, they struggle significantly when tasked with capturing the intricate dependencies and systemic vulnerabilities inherent in modern global trade [2].

This analytical gap has become particularly critical in an era marked by renewed geopolitical tensions and protectionist trade policies. After decades of relative stability and trade liberalization post-World War II [4], the resurgence of nationalist economic policies threatens to fragment global markets into isolated economic blocs [1]. The recent escala-

tion of trade conflicts underscores this vulnerability starkly. For example, the aggressive tariffs imposed by the newly elected President of the United States against major trading partners, such as Mexico, Canada, and China, and subsequent retaliatory measures¹ vividly illustrate how quickly disruptions in trade relationships can escalate, affecting not just directly involved countries but reverberating through the entire trade network.

In this volatile environment, assessing the structural robustness of global trade is paramount. Understanding precisely how trade disruptions spread and which nations stand most vulnerable can provide valuable insights for policymakers aiming to strengthen economic resilience. As emphasized by Nassim Nicholas Taleb in *Antifragile: Things That Gain from Disorder* [5], beyond mere resilience—withstanding shocks—lies antifragility, the property of becoming stronger in response to shocks. Achieving antifragility or even robust resilience within the global trade network necessitates a deep, systemic understanding of network vulnerabilities and interdependencies.

Complex systems, such as the global trade network, exhibit behaviours that cannot be explained solely by analysing their individual parts in isolation. Network analysis offers a powerful alternative that complements and extends traditional macroeconomic approaches, uncovering systemic characteristics often invisible through conventional economic models [2]. By representing international trade as a network, we can systematically identify pathways through which economic disruptions propagate, thus enhancing predictive capabilities regarding economic vulnerabilities and resilience.

This dissertation proposes a novel methodological approach leveraging recent advancements in graph neural networks (GNNs) to bridge the gap between traditional economic analyses and contemporary network-based techniques. Specifically, it aims to provide a nuanced understanding of international trade vulnerabilities through an innovative node classification framework that identifies and predicts economically vulnerable countries in response to supply disruptions. By integrating network theory, economic complexity measures, and advanced deep learning methods, this study contributes a robust analytical toolset capable of addressing and anticipating economic shocks more effectively.

¹See, for example, Canadian Prime Minister Justin Trudeau announcing reciprocal tariffs on American products and urging Canadian businesses to seek alternative suppliers: https://www.youtube.com/watch?v=q1qEaCQa6_w (Last accessed: 20/05/2025).

The primary contributions of this dissertation include:

- A comprehensive assessment of data limitations and methodological choices inherent to analysing international trade networks.
- The introduction of a novel embedding technique for commodities, integrating semantic textual descriptions and economic complexity metrics to enrich node representations.
- Empirical evidence demonstrating the advantages of using hybrid or ensemble GNN models to effectively capture and predict systemic vulnerabilities across various network structures and commodity contexts.
- Concrete ideas for future research.

The remainder of this dissertation is structured as follows: Chapter 2 provides an exhaustive review of previous research on international trade networks; Chapter 3 introduces and describes the data sources utilized in this study; Chapter 4 details the methodological framework employed throughout the research; Chapter 6 offers an in-depth discussion on key results, their limitations, and avenues for future research; and finally, Chapter 7 summarizes the main findings and presents concluding remarks. The code supporting this research is publicly available in the accompanying GitHub repository².

²GitHub repository : <https://github.com/ginofazzi/A-Network-Approach-to-Understanding-Global-Trade-Vulnerabilities-and-Resilience>

Chapter 2

Related Work

“History doesn’t repeat itself,
but it does rhyme.”

Mark Twain

The ITN has been an object of study for several decades, gaining special attention with the acceleration of globalization as a dominant force shaping international commerce. Historically, traditional econometrics provided a theoretical foundation for understanding trade patterns, but it was limited by the availability and granularity of empirical data. In recent decades, however, the emergence of comprehensive, reliable, and publicly available datasets, particularly those compiled by international institutions such as the United Nations, has enabled more robust quantitative and data-driven analyses of global trade.

In this chapter, I review existing literature relevant to understanding the problem of vulnerability and resilience in the ITN. Given that the scope of this dissertation intersects economics, network science, and machine learning, I organize this review around three central themes. First, I summarize traditional economic approaches to modelling and analysing international trade. Second, I discuss how complex network theory has enhanced our understanding of the structural properties, dynamics, and vulnerabilities inherent in global trade relationships. Finally, I present recent advances leveraging geometric deep learning and GNNs, highlighting their application to modelling, analysing, and predicting vulnerabilities in complex interconnected economic systems.

This comprehensive review aims not only to contextualize the methodologies adopted in this research, but also to clearly identify existing knowledge gaps that the current study addresses.

2.1 Traditional Approaches

The theory of international trade dates back over two centuries. Ricardo introduced the principle of comparative advantage [6], suggesting that countries benefit from trade by specializing in goods for which they have a relative efficiency advantage. According to Ricardo, countries engaged in international trade enhance their overall consumption by exporting goods in which they possess comparative advantages and importing those they cannot efficiently produce themselves. This represented a significant evolution from Adam Smith's earlier theory of *absolute advantage* [7], which argued that countries should import commodities that foreign countries could supply more cheaply than domestic production costs. Despite their differences, both theories fundamentally aim to elucidate the rationale and benefits underlying international trade.

Initially, these classical theories were largely theoretical due to limitations in empirical data. However, as data collection improved and econometric methodologies advanced, economists increasingly validated these theoretical foundations empirically. For instance, Daniel Bernhofen and John Brown [8] utilized Japan's sudden transition from economic isolation (autarky) to open trade in the mid-19th century as a natural experiment. Authors showed that opening Japan's markets significantly increased the real prices of its primary exports while imported goods experienced substantial price reductions. These results empirically supported Ricardo's theory.

Neoclassical economists have significantly generalized Ricardo's model. Prominent among these generalizations are the Ricardo–Viner specific factors model [9], which incorporates additional production factors beyond labour, and the Heckscher–Ohlin factor-proportions model [10], which explains trade patterns through differences in countries' factor endowments. Recent empirical studies, such as [11], further reinforced the relationship between relative productivity and observed international trade patterns, providing robust modern statistical evidence for classical trade theories.

The strategic importance of international trade continues to be highlighted in contemporary theoretical studies. For example, [12] em-

phasizes trade's crucial role in facilitating market expansion, securing favourable export prices, and accessing commodities, inputs, and technologies otherwise unavailable domestically. Moreover, recognizing inherent vulnerabilities associated with international trade, recent econometric analyses by [13] and [14] have specifically assessed risks associated with critical commodities vital to global food security, thereby highlighting trade's vulnerability to external shocks and disruptions.

Another important theory for explaining trade between countries is that introduced by Jan Tinbergen, "*The Gravity Model of Trade*" [15]. Inspired by Newton's law of gravitation, the gravity model traditionally predicts bilateral trade flows based on the gross domestic product (GDP) of countries and geographical proximity [16]. The intuitive premise of the gravity model is that trade volume between two nations increases with their economic size and proximity.

Despite their foundational contributions, these traditional approaches commonly exhibit two significant limitations: first, they remain predominantly theoretical with constrained empirical validation; second, when empirical analyses are included, they often adopt a narrow focus by analysing bilateral trade relationships independently, neglecting broader systemic effects stemming from the interconnected nature of global trade. In the following section, I explore how network analysis addresses these limitations by introducing a systemic and structural perspective, offering deeper insights into the vulnerabilities and dynamics of international trade networks.

2.2 Network Analysis

The ITN has emerged as a prominent subject within economic research, encompassing analyses of both financial flows and the physical trade of goods and services. Network-based approaches are driven by two primary motivations: firstly, to gain deeper insights into international trade dynamics, and secondly, to better understand, explain, and anticipate disruptions within the global economy.

One foundational contribution comes from [17], who identified that economic systems inherently exhibit interdependencies through complex and evolving transnational flows. In contrast to traditional economic analyses, they argue that complex-systems methods emphasize emergent properties and systemic behaviours rather than purely individual economic incentives. This study further highlights systemic risks,

recommending policies aimed at fostering network structures resilient to economic shocks.

Building on such theoretical foundations, [18] mapped the ITN as an interdependent complex network using international trade data. They investigated the connectivity of countries within the global economic system, demonstrating that well-connected countries simultaneously present vulnerabilities and strengths: crises originating in highly interconnected countries propagate extensively, yet well-connected target countries can better absorb and mitigate these shocks.

Empirical studies exploring shock propagation through the ITN have reinforced these insights. For instance, [19] examined downstream and upstream effects of network shocks within global supply chains. Their findings empirically validated established economic theories, demonstrating that supply-side shocks propagate downstream, significantly impacting customer industries, whereas demand-side shocks propagate upstream, affecting supplier industries more profoundly.

Similarly, [20] utilized complex network analysis (CNA) tools to assess vulnerabilities associated with commodity-specific trade dependencies, such as wheat. They employed standard network metrics (e.g., PageRank) and shock simulations to evaluate vulnerabilities, defined as susceptibility to scarcity resulting from unexpected disruptions. Their work underscored the strategic relevance of network structures in trade vulnerability assessments, quantifying potential impacts on importer countries.

Earlier foundational work by [21] represented the ITN through directed monetary flows between countries. They uncovered disassortative properties, indicating that countries with extensive trade connections often link with less-connected countries, and characterized hierarchical structures using clustering coefficients. Building on this initial characterization, [22] empirically analysed the ITN, demonstrating typical complex network properties alongside unique attributes, such as a strong association between a country's connectivity and its economic prosperity.

Further refining this characterization, [23] constructed the ITN as a weighted graph, departing from binary connections. They observed distinct statistical properties, notably that most trade links were weak, the weighted ITN was only weakly disassortative, and countries engaged in intense trade were more clustered.

Recent scholarship emphasizes the strategic significance of network-based analyses, particularly in assessing global shocks and informing governance and policy. [2] underscored how global trade concentrates within a limited number of critical hub countries. Dynamic and multilayer network approaches—similar to those employed in this research—facilitate comprehensive analyses across interconnected economic subsystems, highlighting intensified interdependencies caused by globalization and the rapid transmission of local disruptions globally.

Several empirical studies have concentrated specifically on food-related commodities within the ITN, assessing vulnerability using network analysis and simulation techniques: [24] (general food items), [25] (spices), [13] (maize, rice, wheat), [20] (wheat), [26] (rice, maize, wheat, soybeans), and [27] (seafood). Utilizing network metrics—such as node strength, import concentration, local clustering, and centrality—these studies simulate shocks or contamination scenarios, demonstrating the susceptibility of low-income and food-insecure countries to external shocks compared to wealthier countries, which are better equipped to manage disruptions.

In summary, network analysis significantly extends traditional economic approaches by explicitly incorporating systemic complexity and structural interdependencies into analyses of international trade. The next section narrows the methodological focus further, examining recent applications and innovations of GNNs in analysing trade networks.

2.3 Graph Neural Networks

The previous section highlighted how CNA techniques have deepened our understanding of international trade by revealing critical topological features. Extending beyond classical network analyses, in this section I review research employing advanced deep learning architectures, such as GNNs, that integrate explicit node and edge characteristics into their analytical frameworks.

[28] offers a comprehensive introduction to GNNs tailored for graph-structured data and proposes a useful taxonomy structured around three key dimensions: graph structure (structural/non-structural), graph types (directed/undirected, static/dynamic, homogeneous/heterogeneous), and graph learning tasks (node-level/edge-level/graph-level). According to this taxonomy, the present work falls within the category of *structural graphs*, incorporating both directed

and undirected edges, dynamic temporal evolution, homogeneous node types, and a node-level learning task.

Most applications of GNNs to international trade have been oriented towards refining and extending the classical gravity model. An innovative adaptation by [29] introduced the Gravity-informed Graph Autoencoder (GGAE), a specialized GNN-based architecture designed explicitly for link prediction tasks in the ITN. Their model combined elements of classical gravity modelling with graph autoencoder techniques to predict trade volumes effectively. Results indicated substantial performance improvements over traditional gravity-based predictions. However, this study did not directly address assessing vulnerabilities or systemic risks within trade networks.

Similarly, [30] leveraged graph representation learning to model bilateral trade relationships. They utilized various GNN architectures (GCN, ChebNet, and GAT) to classify countries by income level and subsequently predict potential trade partnerships using Graph Autoencoders (GAE) and Variational Graph Autoencoders (VGAE). Notably, their approach emphasized partner prediction rather than quantitative trade volume forecasting.

Focusing on spatio-temporal prediction, [31] developed the AINET (Artificial Intelligence Network Explanation of Trade) framework, integrating Recurrent Neural Networks (RNNs) with GNNs into graph convolutional Long Short-Term Memory (GC-LSTM) models. Their approach demonstrated superior capacity for predicting shifts in trade patterns using historical data and learned graph representations, underscoring the potential for advanced GNN architectures to capture dynamic trade network behaviour.

Building upon this, [32] applied GCN and GAT models to predict international trade flows. Their findings indicated that GNNs significantly outperformed traditional machine learning models (such as random forests), suggesting that GNNs' inherent capability to capture complex, non-linear interdependencies allows for more accurate predictions of trade volumes. Such approaches emphasize the potential of GNNs to identify critical trade links and vulnerabilities within global supply networks effectively.

Beyond international trade, GNN methodologies have proven valuable in assessing network vulnerabilities across various domains. For example, transportation networks [33] and financial systems [34] utilized GNN-based ranking methods to evaluate the criticality of edges

(e.g., road segments) and nodes (e.g., financial institutions), respectively. However, these studies primarily focused on relative rankings rather than explicitly quantifying vulnerability levels.

Despite the promising results achieved by GNN-based models in various network-related tasks, existing literature has not explicitly explored the assessment of vulnerabilities of individual nodes within the ITN. This gap motivates the current research, aiming to directly evaluate country-level vulnerability using node-classification frameworks based on GNN architectures.

2.4 Summary

The previous sections discussed various analytical frameworks utilized in studying the ITN, each with distinct strengths and limitations. Table 2.1 summarizes the main characteristics and contributions of these frameworks, clearly differentiating their approaches and the aspects they address or overlook.

Framework	Key Characteristics and Contributions
Traditional Approaches	<ul style="list-style-type: none"> Primarily theoretical in nature Focuses predominantly on bilateral trade relations Not explicitly modelling interdependencies or structural complexities Uses data for empirical validation of theories (e.g., comparative advantage, gravity model)
Network Analysis	<ul style="list-style-type: none"> Incorporates the structural topology of the trade network Emphasizes the systemic interactions between countries Employs standard complex network metrics (e.g., centrality, clustering, modularity, etc.) Does not explicitly model detailed characteristics of nodes (countries) and edges (trade relationships)
Graph Neural Networks	<ul style="list-style-type: none"> Integrates network topology with detailed node and edge attributes through advanced deep learning techniques Enables nuanced modelling of inter-country trade relationships Explicitly captures both structural and feature-level complexities

Table 2.1: Summary of analytical frameworks for studying the International Trade Network.

As illustrated, each analytical framework progressively incorporates greater complexity and detail, with GNNs representing the most comprehensive approach discussed, offering substantial improvements in

predictive capabilities and vulnerability assessment. To the best of my knowledge, this dissertation represents the first application of GNN-based node classification explicitly aimed at identifying vulnerable countries within the ITN. By directly modelling country-level vulnerability, this work provides novel insights into the systemic risks inherent in global trade structures.

Chapter 3

Data

“Information is the oil of the 21st century, and analytics is the combustion engine.”

Peter Sondergaard

The primary data source for this project is the UN Comtrade Database, which provides detailed annual global trade statistics by product and trading partner. Covering over 200 countries and accounting for more than 99% of global commodity trade, it is the largest and most widely used resource in both commercial and academic research [35][36][37][38][22].

The dataset includes all recorded trade transactions classified under the Harmonized System (HS)¹ and the Standard International Trade Classification (SITC)², both regularly updated by the World Customs Organization.

In particular, the HS comprises more than 5,000 commodity groups, each identified by a six-digit code and organized in a hierarchical structure. The first two digits represent broad commodity categories, the next two digits indicate subcategories, and the final two digits provide finer-grained distinctions.

For example, the code ‘01’ corresponds to ‘*Live animals*’. Under this category, ‘0101’ refers to ‘*Live horses, asses, mules and hinnies*’, and

¹Source: <https://www.wcoomd.org/en/topics/nomenclature/overview/what-is-the-harmonized-system.aspx> (Last accessed: 24/04/2024).

²Source: <https://unstats.un.org/unsd/trade/sitcrev4.htm> (Last accessed: 24/04/2024).

'010121' designates '*Pure-bred breeding animals*'. In this work, I categorize commodities using the first two digits of the HS2012 classification, which I refer to throughout as "Commodity classes." The full list of commodity classes can be found in Appendix A.1. For further details on the classification system visit UN Comtrade website³.

This dataset has been enhanced in various ways by previous research efforts. In this study, I leverage several of these enriched datasets. Most notably, the work of [36] (see Section 3.2) contributes product- and country-level features, such as the Economic Complexity Index (ECI) and the Complexity Outlook Index (COI), both of which are employed in this analysis. These features are described in detail in Section 4.2.4.

Additionally, the *Base pour l'Analyse du Commerce International* (BACI) database [39] (see Section 3.3), from the *Centre d'Etudes Prospectives et d'Informations Internationales* (CEPII⁴) builds on the raw Comtrade data by applying corrections and homogenizing commodity quantities, enabling consistent comparisons of trade flows across countries, products, and years.

The following sections describe the composition of the datasets in greater detail and explain how they are used in this study. For a comprehensive visualization of how these data sources integrate, please refer to Appendix E.

3.1 UN Comtrade Database

The UN collects and maintains international trade data for all its member countries. The data is freely available and accessible via API upon subscription, though it comes with some limitations, such as daily request caps and a maximum of 100,000 records per call (see Appendix A.3). Trade records are reported unilaterally by individual countries on either a monthly or annual basis. The dataset includes the following key attributes:

1. **Trade flow:** Classification of the trade direction (e.g., Import, Export).
2. **Reporter:** The country submitting the trade report.

³[⁴CEPII is the leading French research center on global economic analysis.](https://comtradeplus.un.org>List of References (Last accessed: 12/05/2025).</p></div><div data-bbox=)

3. **Partner:** The trading counterpart as reported by the reporting country.
4. **Product:** Commodity classification using either the HS or the SITC system.
5. **Trade value:** Reported in CIF (Cost, Insurance, and Freight) and/or FOB (Free on Board) terms⁵. The dataset includes a `primaryValue` column, which prioritizes CIF values when available, followed by FOB values otherwise.

An example of bilateral trade data is shown in Table 3.1. A full description of the retrieved dataset –including all available features and API limitations– can be found in Appendix A.4.

Trade flow	Reporter code	Partner code	Classification system	Commodity code	Trade value
Export	12	380	HS	03	168,635
Import	12	380	HS	03	452,266
Export	380	12	HS	03	417,683
Import	380	12	HS	03	696,920

Table 3.1: Trade flow between Algeria (code 12) and Italy (code 380) for a specific commodity class (code 03). The reporter is the country submitting the data to the UN Comtrade system.

Because trade data is self-reported by individual countries, certain inconsistencies and reporting discrepancies arise. For example, note in Table 3.1 that the reported exports from Algeria to Italy do not match Italy’s reported imports from Algeria. These issues are discussed in detail in the following section.

3.1.1 Data Discrepancies

Despite efforts to collect reliable data on international trade, an increasing number of studies have documented discrepancies in reported trade flows. These inconsistencies are a source of concern because they can

⁵Cost, Insurance, and Freight (CIF) and Free On Board (FOB) are international shipping agreements used in the transportation of goods between a buyer and a seller. CIF values include delivery to the frontier of the importing country, whereas FOB values include only the cost at the point of shipment.

distort empirical analyses and compromise the validity of policy recommendations [35].

The UN Statistics Division acknowledges this issue:

"Bilateral asymmetries in trade data (both in goods and services) are a well-known phenomenon in official statistics. Bilateral trade asymmetries occur when the reported exports from country A to country B do not match the reported imports to country B from country A."

According to [35], trade flow discrepancies fall into two broad categories:

- **Logistic discrepancies:** Structural differences related to valuation methods and reporting standards. UN Comtrade identifies three primary sources of asymmetry: differing partner attribution criteria for imports versus exports, use of CIF-type values in imports and FOB-type values in exports, and the application of different trade system definitions. Other contributing factors include time lags (e.g., goods leaving country A in 2012 may only reach country B in 2013), re-exports through third countries, and divergent product classification standards.
- **Non-logistic discrepancies:** Often intentional misreporting—such as misvaluation or misclassification—to avoid import tariffs [40], evade export taxes [41], or minimize corporate profit and turnover tax liabilities [42]. In some cases, the scale of misreporting has been shown to be substantial.

Missing records, inconsistencies, and potentially deliberate misstatements may serve as a proxy for assessing a country's institutional reliability. Under the hypothesis that poor reporting correlates with weak institutions, these discrepancies may offer useful insights into the quality of governance. Accordingly, while I adopt the conventional approach of smoothing these mismatches (as done in prior work), I also leverage a trustworthiness metric from previous research, as a potential signal of reporting quality—detailed in Section 4.2.4.

3.2 Atlas of Economic Complexity

As discussed in Section 3, this work leverages the dataset from [43], which refines and enhances the UN Comtrade data. Their improvements include smoothing trade discrepancies and introducing key economic complexity metrics.

Smoothing discrepancies: The UN Comtrade dataset often contains inconsistencies and gaps due to discrepancies in country reporting (see Section 3.1.1). Authors in [43] mitigate these issues by using mirrored trade flows to fill gaps while accounting for FOB-CIF variations. To standardize these values, they apply a discount rate of 8%, based on prior research (detailed formulas in Appendix B).

After these corrections, the data from [43] is symmetrical in the trading flows: Export of product p from country C to country C' is equal to the Import of product p from country C' to country C .

Economic Complexity Index (ECI): A measure on how diversified and complex a country’s export basket is. Countries with a great diversity of complex specialized commodities correspond to higher (positive) values. It is based on the idea that commodity exports are a good proxy for a country’s ability to produce this commodity. If a country cannot produce a given product (e.g., cars), it is impossible to export that product. On the other hand, while a country may produce a product without exporting it, this often suggests limited production capacity or insufficient quality to compete in global markets. Once we can determine all products each country can produce (i.e., export), we can measure how diversified the product portfolio is (diversity), and how many of these products are hard to find elsewhere (ubiquity). To conclude the measure of the number of capabilities available in a country, or required by a product, authors correct the information that diversity and ubiquity carry by using each one to correct the other. This recursion converges into a single value, that can be obtained using a closed formula (see Appendix B.5 for more details).

Complexity Outlook Index (COI): A measure of how many complex products are near a country’s current set of productive capabilities. We can think of this as a potential extension of the ECI, in the sense that a country’s ECI determines how complex its economy is now, while the

COI gives us an idea of how easily a country could grow its ECI in the near future. As an example, a country that already manufactures cars may find it easier to begin producing trains, as many of the required skills and infrastructure are already in place.

Product Complexity Index (PCI): A measure of how complex a product is. It is conceptually similar to the ECI but applied to products rather than countries. The PCI is computed based on:

- How many countries export the product.
- How diversified those exporters are.

In practical terms, products with high PCI values are more complex and less ubiquitous, meaning they require advanced capabilities to produce.

For detailed mathematical formulations of these indexes, see [36].

3.3 Database for International Trade Analysis (BACI)

Similar to the methodology used in [36], BACI reconciles trade flows to smooth discrepancies by consolidating CIF and FOB values. This is achieved through a regression-based estimation of product-specific CIF/FOB ratios, weighted by measures of data reliability when reconciling bilateral trade flows.

By leveraging bilateral trade reports, BACI ensures broader country coverage and enhances data reliability, particularly for unit-value estimates. Since BACI originates from the same UN Comtrade dataset, its country codes and commodity classifications align with the other data sources used in this work. I incorporate BACI primarily for its standardized quantity units, where all commodities are converted to metric tons. This standardization facilitates direct commodity comparisons and enables additional analyses, such as calculating price per unit –an aspect that was not feasible when considering only gross trade flows.

3.4 Additional Data Sources

In addition to the primary databases described earlier, I use several supplementary datasets to enrich the analysis or provide contextual comparison. These include:

1. **Global Preferential Trade Agreements Database⁶**— Published by the World Bank, this database provides a comprehensive list of regional, plurilateral, and bilateral Free Trade Agreements (FTAs) between countries. It is used in Section 4.2.5 to enhance the representation of trading relationships through enriched link features.
2. **Discrepancy Index[35]⁷**— From the World Bank’s Reproducible Research Repository, this dataset quantifies bilateral asymmetries in trade reporting. It provides a country-pair level index that supports the evaluation of data reliability and reporting quality across countries.
3. **Corruption Perceptions Index (CPI)⁸**— Produced by Transparency International, a non-governmental organization focused on global advocacy, campaigning, and research to promote transparency and integrity in public institutions. The CPI ranks countries by perceived levels of public sector corruption, and it’s referenced in Section 6.2.6 to compare and contextualize the Trustworthiness Index and the Discrepancy Index.
4. **Country Bordering Countries Dataset⁹**— Maintained by the GeoNames geographical database, this dataset provides a list of countries and their bordering neighbours. It is used in Section 4.2.4 to compute geo-positional embeddings.
5. **CERDI-Sea Distance Database [44]¹⁰** — This database contains bilateral maritime distances between 227 countries and territories. It is used in Section 4.2.4 when computing geo-positional embeddings.

⁶Source: <https://wits.worldbank.org/gptad/library.aspx> (Last accessed: 24/04/2025)

⁷Source: <https://reproducibility.worldbank.org/index.php/catalog/145> (Last accessed: 24/04/2025)

⁸Source: <https://www.transparency.org/en/cpi/2023/index/dnk> (Last accessed: 12/05/2025)

⁹Source: <http://download.geonames.org/export/dump/countryInfo.txt> (Last accessed: 24/04/2025)

¹⁰Source: <https://zenodo.org/records/46822#.VvFcNWMvyjp> (Last accessed: 24/04/2025)

Chapter 4

Methods

“All models are wrong, but some are useful.”

George Box

4.1 International trade as a network

The ITN can be easily understood and modelled as a graph, where nodes represent countries, and there is a link between nodes if two countries trade a commodity (Figure 4.1a). Furthermore, we can disentangle these bidirectional edges into a directed link, from country A to B , if country A exports to country B (or grounded on imports, by changing the direction of the edge) (Figure 4.1b). This separation not only makes the relationships asymmetrical, but also adds a more nuanced understanding of the trading relationships –some countries might be net exporters for a certain commodity, but exclusively importers in another set of commodities. This approach is not new, and has been used extensively in research [21][2][45][19][36][18]. According to [2], traditional macroeconomic approaches often fail short to effectively capture the complex relationships raised by international trade. One key pitfall in most traditional macroeconomic tools is that analysis of cause and effect of trading phenomena is done pairwise between related countries, but never fully grasping the potential cascading effect only understandable by looking at the network topology.

While representing international trade as a network provides valuable insights into trade relationships and their directional nature, there

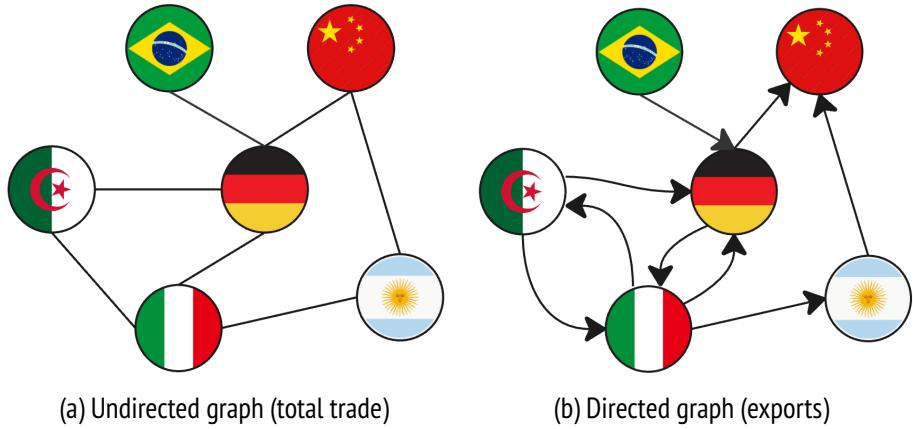


Figure 4.1: ITN as a graph between countries (nodes).

are multiple ways to structure this representation. The next section introduces a multilayered network approach designed to explicitly capture commodity-specific interactions within the ITN.

4.2 Multilayered network

International trade between countries cannot be fully captured by examining only total trade volumes between nodes. One reason is that a country can be considered as a net exporter for one category of commodities, and still be exclusively importer for another category. By collapsing these relationships into a gross trading link, we lose valuable information from the network. One way to account for this complexity, is to think of the ITN as a multilayered dynamic network, where each commodity group is a layer¹, and each year is a snapshot of the period. This is analogous to defining separate edges between countries according to commodity type and trade direction. So in practice, every edge type represents a different layer of the network (called here *commodity layer*), and each year is a single snapshot of the graph. One important note is that each layer and snapshot constitutes a simple graph, and thus we can treat these graphs as individual graphs. Mathematically:

$$G_{\text{year}} = (V, E, L_c)$$

¹Here I mean a layer of a graph. Not to be confused with layers of a Graph Neural Network model. In this work I will refer to the former with the symbol l , and the latter with the superscript k .

where $(u, v, l) \in E$, with $u, v \in V$ as nodes (countries), $l \in L$ as the layer (commodity group), and $G_{\text{year}} = \{G_{2012}, G_{2013}, \dots, G_{2022}\}$. A graphical description of this setup can be seen in Figure 4.2.

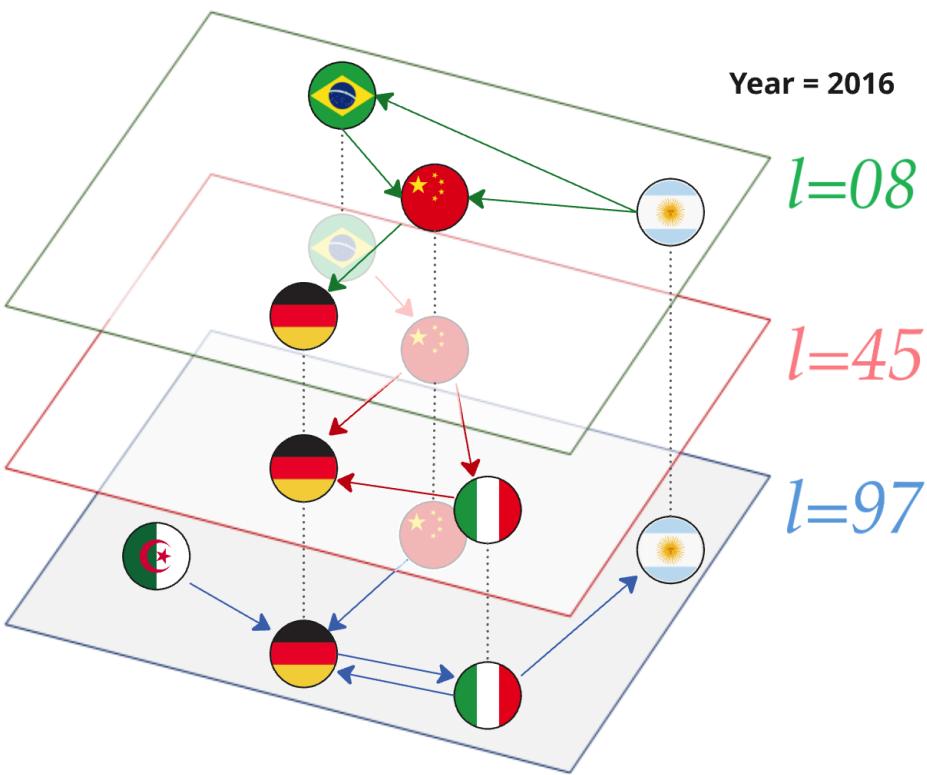


Figure 4.2: Graphical description of multilayered network. For a specific year, each layer (l) represents the trading of a different commodity class.

Notice that there is a one-to-one node mapping between the layers of the network. In other words, the country entities in each layer are always the same.

4.2.1 One Graph, Many Edges

The multilayer approach described above provides a clean, interpretable representation of commodity-specific trading patterns. Each individual graph captures the flow of a specific commodity between countries. However, this representation may overlook potential interactions and dependencies that occur across commodity layers. Prior research [46] has shown that changes in the trade behaviour of certain product categories can trigger ripple effects across interconnected sectors.

To capture such interdependencies, I construct an alternative representation: a graph in which a pair of countries may be connected by multiple directed edges, each corresponding to a distinct traded commodity (Figure 4.3), sometimes called Heterogeneous Multi-Graph. In this setup, multiple edges between the same node pair are distinguished by their edge attributes (see Section 4.2.5), reflecting the commodity-specific nature of each trade connection.

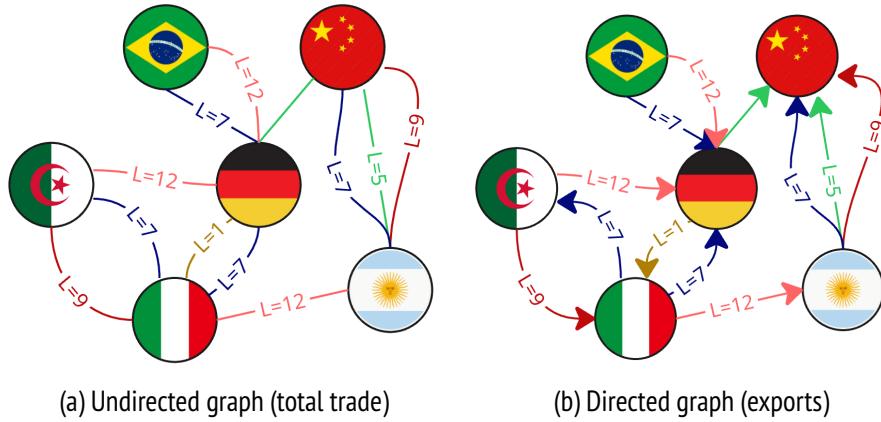


Figure 4.3: ITN as a graph between countries (nodes), with multiple edges. each edge representing a different traded commodity.

4.2.2 Graph Types

In this work, I proceed to construct five different graphs:

4.2.2.1 Commodity-Layered: Following Section 4.2, where each layer is an individual graph, in three different modalities:

- a) *Total Trade*: Graphs constructed using total trade volume (imports + exports) as edge weights, resulting in an undirected graph. This setup captures the overall strength of trade relations per layer, without directional information (Figure 4.1a).
- b) *Exports*: Graphs built using only export flows as directed edges from exporters to importers. This representation explicitly captures the directionality inherent in trade relations, allowing the model to distinguish between importing and exporting roles (Figure 4.1b).

- c) *Export + Layer*: Extends the *Exports* setup by enriching node features (see Section 4.2.4) with an additional pre-learned layer embedding (see Section 4.2.3). This setup enables the model to differentiate between commodity layers explicitly, thus incorporating information about the commodity type involved in each trade relationship. Formally, each node's extended feature representation is defined as:

$$\mathbf{h}_l^0 = [\mathbf{h}^0 \parallel \mathbf{l}]$$

where \mathbf{h}_l^0 denotes the enhanced node feature representation, $\mathbf{l} \in \mathbb{R}^{16}$ is a learned layer embedding vector, and \parallel signifies vector concatenation. Further details regarding the generation of these embeddings are provided in the following subsection.

4.2.2.2 Multi-Graph: Following Section 4.2.1, where one graph allows many connections, in two different modalities:

- a) *Total Trade*: Same as (4.2.2.1a), but in the Multi-Graph setting (Figure 4.3a).
- b) *Exports*: Same as (4.2.2.1b), but in the Multi-Graph setting (Figure 4.3b).

A summary of the distinct graphs referred in this work can be seen in Table 4.1

Graphs					
	Multilayer		Multi-Graph		
Total	Export	Export + Layer	Multi-Graph-Total	Multi-Graph-Export	
Undirected	Directed	Directed	Undirected	Directed	

Table 4.1: A summary of the 5 different graphs constructed in this work.

4.2.3 Layer Embedding

To explicitly encode commodity-layer information, I pre-compute layer embeddings from their commodity descriptions and average PCI of its products, using a transformer-based architecture. Specifically, I utilize the pre-trained model `all-MiniLM-L6-v2`² from the Sentence-Transformers family, which is designed to capture semantic information

²Source: <https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2> (Last accessed: 15/04/2025).

from sentences and short texts—making it particularly suitable for this task. The following pipeline summarizes the generation of semantically and economically enriched layer embeddings.

1. **Description Generation:** For each commodity layer (l), I first create a descriptive text by aggregating individual product descriptions provided by the HS2012 commodity nomenclature. This aggregation involves careful text pre-processing, including the removal of digits, punctuation, and redundant characters. Unique keywords are retained to construct a concise, non-redundant textual description that effectively captures the semantic characteristics of each commodity class.
2. **Embedding Computation:** The curated textual descriptions are fed into the Sentence-Transformer model, which produces 384-dimensional embeddings using mean pooling over the final hidden layer’s token representations. A linear projection layer is then applied to reduce the dimensionality of each embedding to 8.
3. **Fine-Tuning:** To better align the embeddings with the specific semantics of commodity classifications, I further fine-tune the model. This is done by appending a second linear projection layer that outputs 96 logits corresponding to the number of *commodity classes*. All original parameters of the transformer model are frozen during this step, and only the new classification head is trained using a cross-entropy loss function over 20 epochs. After training, the intermediate 8-dimensional embeddings (prior to the classification layer) are retained as the final representation of each commodity layer.
4. **PCI-Based Scaling:** Finally, I scale the learned layer embeddings by the mean PCI of the products included in each commodity class for a given year. This step integrates economic sophistication with the semantic content of each layer, enhancing the informativeness of the embeddings.

A visual overview of this process is provided in Figure 4.4.

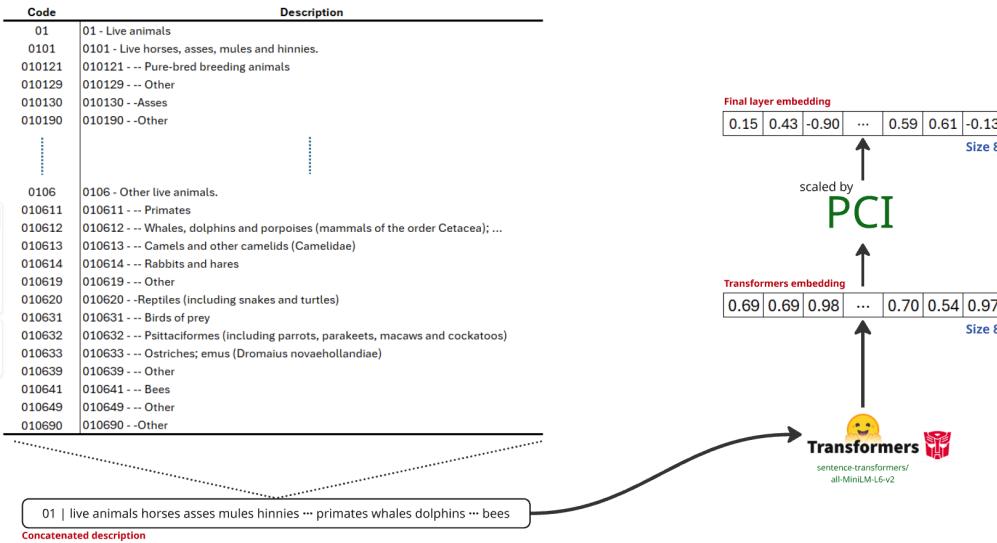


Figure 4.4: Visual depiction of the process to compute layer embeddings. Textual descriptions are preprocessed and concatenated, passed through a transformer model to obtain semantically meaningful embeddings, and then scaled by PCI values to incorporate product sophistication.

To evaluate the effectiveness of the resulting layer embeddings in differentiating commodity classes, I compute pairwise cosine similarities and visualize them in a heatmap (Figure 4.5). This visualisation reveals clear distinctions among most classes, while also highlighting meaningful clusters—such as codes 01 to 20 (Food and Derivatives), 27 to 40 (Minerals and Chemicals), and 50 to 58 (Textiles). These clusters suggest that the embeddings effectively capture underlying semantic and economic similarities among commodity groups. After scaling for PCI, further distinctions can be made among classes, as a result of their inherent complexity differences.

4.2.4 Countries as nodes

In the ITN representation, each country is assigned a node, enriched with country-specific attributes. Formally, each node v is represented as a vector of attributes:

$$v = (a_1, a_2, a_3, \dots, a_n)$$

where a_n is the value for v of the n^{th} attribute.

The attributes considered for each node are presented below.

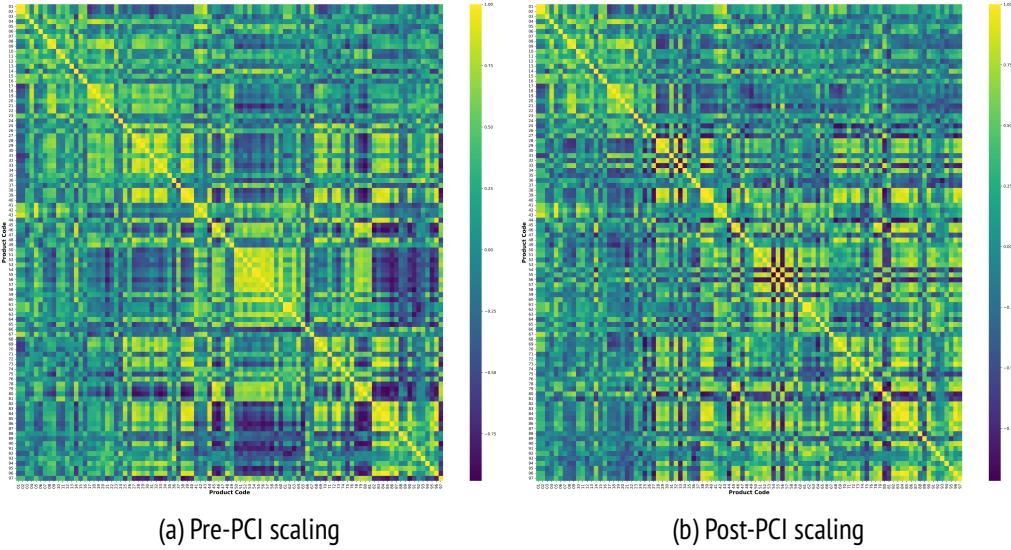


Figure 4.5: Cosine similarity matrix of commodity class embeddings, pre- and post-scaling by PCI.

Economic Complexity Index (ECI) & Complexity Outlook Index (COI) [36]

ECI measures how diversified and complex a country's export basket is. Countries exporting a wide range of specialized commodities have higher ECI values (see Section 3.2 for details).

COI Measures how many complex products are near a country's current set of productive capabilities. It captures the ease of diversification for a country, where a high COI reflects an economy's existing capabilities to drive easy diversification into its production. A low COI thus reflects that a country will find it difficult to acquire new know-how and increase their economic complexity (for more details, see Section 3.2).

The reason to include these two attributes resides in the assumption that the level of current and future specialization of a country can be determining in how resilient a node is to shocks in their neighbourhood. Both these attributes are taken from [43].

Number of Distinct Products & Smoothed Revealed Comparative Advantage (SRCA)

These two attributes aim to capture diversification and importance of a country's export basket for a commodity class. The former (number of distinct products) reflects the country's trade diversification in that sector; while the latter (SRCA) measures a country's relative advantage in trading a product compared to global trade patterns. A country is

considered an effective trader of a product if its trade share for that product is equal or higher than the global average ($\text{SRCA} \geq 1$).

The number of distinct products is simply the unique number of commodities using the first two digits of the HS classification. As an example, when looking at commodity class '10' (i.e., *Cereals*), a country exporting 'Rye' (Code 1002), 'Oats' (Code 1004) and 'Buckwheat, millet and canary seeds; other cereals' (Code 1008) counts as three products within this class, since although these products share the same first two digits, they are different subcategories; whereas a country exporting only 'Oats' (Code 1004) would count as one product, since all its export depends on one specific subcategory of the commodity class.

The SRCA requires a more careful calculation. Following in the footsteps of [36], the first step is to construct the Country-Product Matrix for each year, where each row represents a country (c), each column represents a product (p), and the entries correspond to trade volumes (*exports* [X], *imports* [M], or *total trade* [$T = X + M$]). For example, the product-matrix for exports³ is given by:

$$X = \begin{vmatrix} x_{c_1,p_1} & x_{c_1,p_2} & \cdots & x_{c_1,p_m} \\ x_{c_2,p_1} & x_{c_2,p_2} & \cdots & x_{c_2,p_m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{c_n,p_1} & x_{c_n,p_2} & \cdots & x_{c_n,p_m} \end{vmatrix}$$

Then, I can calculate the Revealed Comparative Advantage (RCA) in my data. For a country c , and a product p , with trading volume $x_{c,p}$, the RCA is given by:

$$\text{RCA}_{cp} = \frac{x_{cp}}{\sum_c x_{cp}} / \frac{\sum_p x_{cp}}{\sum_{c,p} x_{cp}} \quad (4.1)$$

Following [36], I smooth changes in trading volumes induced by the price fluctuation of commodities by using a modification of Equation 4.1, in which $x_{c,p}$ for a given year t is averaged along with the previous two years by using weights⁴:

$$\tilde{x}_{c,p}^{(t)} = \frac{2x_{c,p}^{(t)} + x_{c,p}^{(t-1)} + x_{c,p}^{(t-2)}}{4}$$

³This can be done analogously for either Imports or Total Trade.

⁴For the first year of data (2012), the Smoothed RCA is identical to the RCA from Equation 4.1, as no prior years exist for averaging. For the following year (2013), the smoothing process incorporates only one previous year (2012) instead of two.

The final Smoothed RCA (*SRCA*) is then computed as:

$$SRCA_{c,p}^{(t)} = \frac{\tilde{x}_{c,p}^{(t)}}{\sum_c \tilde{x}_{c,p}^{(t)}} / \frac{\sum_p \tilde{x}_{c,p}^{(t)}}{\sum_c \sum_p \tilde{x}_{c,p}^{(t)}} \quad (4.2)$$

SRCA values range from zero (no trade) to potentially infinite (if a country exclusively trades product p and no other country does). *SRCA* is undefined when a country has no trade in a specific product category in a given year, so I manually set that to zero.

Geo-Positional Embeddings

A country's geographic position plays a fundamental role in shaping its trade relations, as physical distance is a major determinant of transport costs and trade feasibility [3][27][15]. To capture this geographic influence, I include a learned geo-positional embedding for each country, computed based on spatial proximity and transport connectivity (see Section 3.4).

The first step involves constructing a proximity graph where nodes represent countries. An edge is added between two countries u and v if they share a land border. These border edges are assigned a fixed weight of 1, reflecting the notion of near-zero logistical barriers at land borders. This step captures immediate geopolitical adjacency.

To account for maritime trade, additional edges are included between countries with possible commercial sea routes, based on the work of [44]. These maritime edges are weighted proportionally to the inverse of the maritime distance between the countries, scaled to reflect relative proximity. The edge weight is computed using the following formula:

$$\text{MaritimeWeight}_{u,v} = \epsilon + (1 - \epsilon) \cdot \left(1 - \frac{d_{u,v}}{0.5 \cdot C_E} \right) \quad (4.3)$$

where:

- $d_{u,v}$ is the maritime distance between countries u and v (in km),
- C_E is the Earth's circumference (approximately 40,075 km),
- ϵ is a smoothing factor (set to 0.1) to prevent edge weights from vanishing at maximum distances.

According to this formula, a zero-distance connection (hypothetical in practice) results in a maximum weight of 1, matching the weight of a shared land border. Conversely, a maximum maritime distance of 20,037.5 km yields a weight of $\epsilon = 0.1$, representing minimal but non-zero connectivity.

The total edge weight between two countries is the sum of their land border weight (if applicable) and their maritime connection weight (if applicable). For example, Denmark and Germany share both a land border and a maritime connection via their main ports, Aarhus and Hamburg, respectively. The land border contributes a weight of 1, and the maritime connection (approximately 815 km) contributes an additional weight of 0.96. On the other hand, Bolivia and Germany do not share a land border but can trade via maritime routes. Notice that Bolivia is a landlocked country, but following [44], I allow Bolivia to have a maritime connection via the use of the nearest relevant port (Port of Santos, Brazil). This yields a maritime distance of approximately 12,512 km, corresponding to a maritime edge weight of about 0.38.

To convert this proximity graph into geo-positional embeddings, I apply the *Node2Vec* algorithm [47]. *Node2Vec* generates continuous vector representations for nodes by simulating biased random walks over the graph and optimizing a skip-gram objective. In this implementation:

- Embeddings are generated in an 8-dimensional space ($d = 8$),
- 200 random walks of length 10 are performed for each node,
- Edge weights (as defined above) guide the random walk probabilities.

Finally, I use Uniform Manifold Approximation and Projection (UMAP) to reduce and visualize the embeddings in a 2-dimensional space. Figure 4.6 shows the resulting layout, where geographically close countries are generally embedded near each other.

Herfindahl–Hirschman index (HHI) [48]

The HHI is a commonly used measure to determine the amount of competition in an industry. It is calculated by squaring the market share of each competing firm and summing the results. The outcome is proportional to the average market share, weighted by market share.

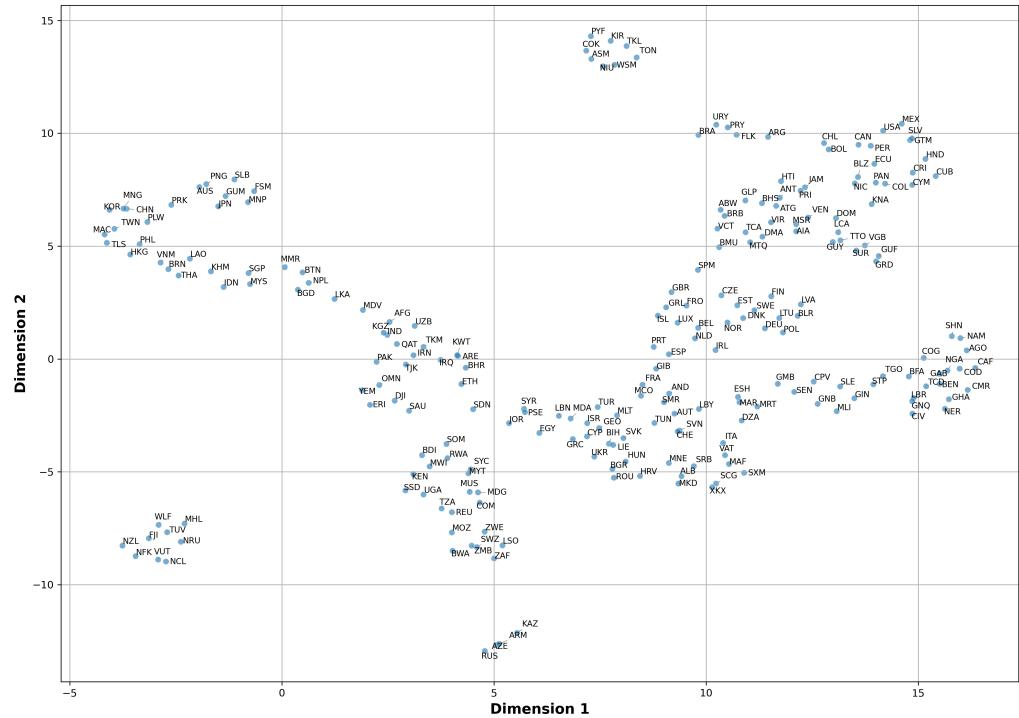


Figure 4.6: Geo-Positional Node Embeddings for countries, visualized in 2-Dimensions using UMAP.

With a reinterpretation of the original concept, HHI is used here to quantify trade diversity in terms of commercial partners. By replacing firms' market shares with countries' trade shares, the formula becomes:

$$HHI_c = \sum_{i=1}^N (TS_i)^2$$

where N is the number of trade partners of country c , and $TS_i = \frac{\text{trade volume with partner } i}{\text{total trade volume of country } c}$ is the trade share of partner i .

The index ranges from 0 (highly diverse trade) to 1 (highly concentrated trade), capturing both the number of trading partners and the distribution of trade volumes. As an example, consider countries A , B , and C in Table 4.2:

- Country A 's HHI is $0.29^2 + 0.71^2 \approx 0.59$
- Country B 's HHI is $0.02^2 + 0.98^2 \approx 0.96$
- Country C 's HHI is $0.5^2 + 0.5^2 = 0.50$

Country A			Country B			Country C		
Partner	Volume	Share	Partner	Volume	Share	Partner	Volume	Share
a_1	1000	0.29	b_1	50	0.02	c_1	500	0.50
a_2	2500	0.71	b_2	3000	0.98	c_2	500	0.50

Table 4.2: Example calculation of the HHI for three countries trading with two partners each. Although all countries have the same number of partners ($n = 2$), trade concentration varies, resulting in different HHI values. Countries A and B face higher concentration risk due to imbalanced trade.

The underlying assumption is that a country's systemic risk is reduced by this trading diversification, in agreement with previous research [27][49]. This assigns a higher risk to countries that concentrate trade in a few partners. The more partners, and the more evenly distributed the trade, the lower the index.

Other risk measures for evaluating trade partner diversification were also considered, and are discussed in Section 6.2.5. HHI was ultimately chosen due to its ease of interpretation and computation, while still effectively capturing both partner count and trade volume distribution.

Trustworthiness index (TI) [50]

The TI measures a country's reporting reliability based on discrepancies in trade data for a given year-product pair. I compute this index following [50], using the raw tariff line data reported to UN Comtrade (Section 3).

Constructing Trade Discrepancy Data

To assess discrepancies, I retrieve, for all years and commodities:

- Reported exports from each country to each of its partners.
- Reported imports by the partner from the exporting country.

Following the example from Section 3.1, in 2016, Algeria reported exports of 168,635 USD in commodity class 03 ('*Fish and crustaceans, molluscs, and other aquatic invertebrates*') to Italy. Meanwhile, Algeria reported imports from Italy for 452,266 USD in the same year and commodity class.

On the other hand, Italy reported imports from Algeria of 696,920 USD and exports to Algeria of 417,683 USD. These values create a trade discrepancy, visualized as a graph with countries as nodes and reported trade flows as edges (Figure 4.7).

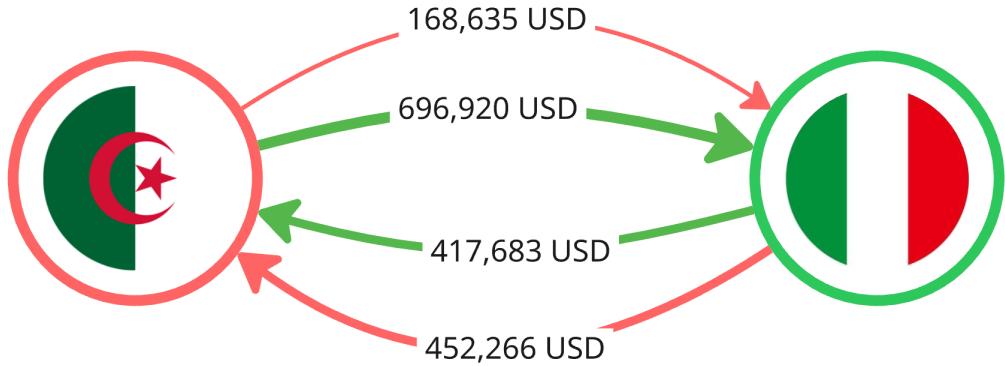


Figure 4.7: The constructed graph of reported trading flows. Algeria and Italy have both reported exports and imports of commodity class 03 with each other. The direction of the edge goes from exporter to importer. The colour of the edge matches the reporter node colour. The amount reported is represented by the edge's label and thickness.

A noticeable discrepancy exists, particularly in Algeria's reported exports to Italy versus Italy's reported imports from Algeria.

Defining the Trustworthiness Index

To quantify these discrepancies, the authors of [50] define the mismatch function for a country pair (a, b) as:

$$M(a, b) = |\alpha_a(a \rightarrow b) - \alpha_b(a \rightarrow b)| + |\alpha_a(b \rightarrow a) - \alpha_b(b \rightarrow a)| \quad (4.4)$$

where $\alpha_a(a \rightarrow b)$ denotes the value of the flow reported by a to b . Next, the total operation volume for the country pair is:

$$\Psi(a, b) = \alpha_a(a \rightarrow b) + \alpha_b(a \rightarrow b) + \alpha_a(b \rightarrow a) + \alpha_b(b \rightarrow a) \quad (4.5)$$

The initial trustworthiness at step 0 is then computed as:

$$T_0(a, b) = 1 - \frac{M(a, b)}{\Psi(a, b)} \quad (4.6)$$

Example Calculation for Algeria and Italy

Following our example, the mismatch function (Eq. 4.4) for Algeria (A) and Italy (I) equals:

$$\begin{aligned}
M(A, I) &= |\alpha_A(A \rightarrow I) - \alpha_I(A \rightarrow I)| + |\alpha_A(I \rightarrow A) - \alpha_I(I \rightarrow A)| \\
&= |168,635 - 696,920| + |452,266 - 417,683| \\
&= 562,868
\end{aligned}$$

And the operation volume (Eq. 4.5):

$$\begin{aligned}
\Psi(A, I) &= \alpha_A(A \rightarrow I) + \alpha_I(A \rightarrow I) + \alpha_A(I \rightarrow A) + \alpha_I(I \rightarrow A) \\
&= 168,635 + 696,920 + 452,266 + 417,683 \\
&= 1,735,504
\end{aligned}$$

Finally, the trustworthiness of the flow is:

$$T_0(A, I) = 1 - \frac{M(A, I)}{\Psi(A, I)} = 1 - \frac{562,868}{1,735,504} = 0.676$$

Then, the overall trustworthiness of a country can be calculated simply by averaging T_0 with respect to all its partners ($T_0(a, \cdot)$). So for country a , its trustworthiness is given by:

$$T_0(a) = \frac{1}{|N_p(a)|} \sum_{b \in N_p(a)} T_0(a, b) \quad (4.7)$$

where $N_p(a)$ is the set of all country partners (neighbours) of a in the constructed graph of trades. In our example we get:

$$\begin{aligned}
T_0(A) &= 0.78 \\
T_0(I) &= 0.77
\end{aligned}$$

Iterative Trustworthiness Refinement

I then compute T_n , a recursive refinement of T_0 , which accounts for discrepancies in the trustworthiness of a country's neighbours. This process resembles *message passing* in GNNs, where a node progressively integrates information from its neighbours. The update formula relies on the trustworthiness balance (B):

$$B_n(a, b) = \frac{T_{n-1}(a)}{T_{n-1}(b)}$$

where $B_n(a, b) > 1$ if country a was previously judged more trustworthy than b . The trust update rule is:

$$T_n(a, b) = \begin{cases} \frac{B_n(a, b)}{B_n(a, b) + (1 - T_{n-1}(a))} & \text{if } T_{n-1}(b) \neq 0 \\ 1 & \text{if } T_{n-1}(b) = 0 \end{cases}$$

This equation also defines the stopping condition for the recursion.

Following [50], I run the iterative update for 200 iterations and monitor the average trustworthiness convergence (Figure 4.8). To assess how the trustworthiness changes over the iterations, I apply the knee detection algorithm [51], allowing me to find the elbow of the curve –i.e., the moment at which the rate of convergence is maximum. As shown in Figure 4.8, after 25 iterations the change rate slows down, and after 100 iterations, the average trustworthiness difference per iteration becomes negligible, allowing to halt the process. This aligns with findings in [50].

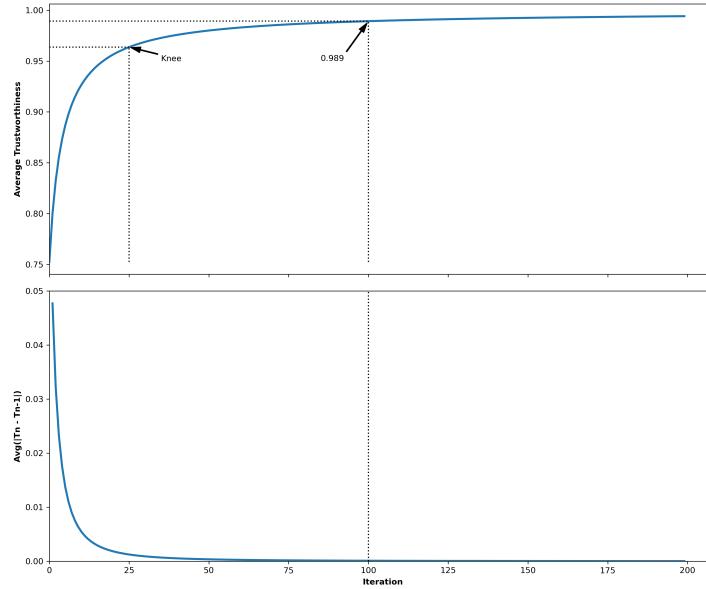


Figure 4.8: Trustworthiness algorithm convergence: Average country trustworthiness at each iteration (top) and average difference between iterations (bottom).

Handling Missing Data

Some countries present in [43] lack corresponding raw trade data in UN Comtrade (see Appendix A.5 for the full list). To solve this, I proceed as follows:

- If a country appears in Atlas (Section 3.2) but not in UN Comtrade (Section 3.1), I assign it the average trustworthiness of all countries for that year-product (17 such cases).
- If a country is in UN Comtrade but has no reported data, I assign it a trustworthiness score of zero, penalizing missing reporting (30 such cases).

Attribute normalization

Attributes used from [36] (ECI & COI) are normalized within the country-year data set, to have a mean of 0 and standard deviation of 1. To ensure comparability across features, I apply normalization to the rest of the attributes:

- The *Number of Distinct Products (# Prod.)* and the *Smoothed Revealed Comparative Advantage (SRCA)* are thus normalized within the year-product dataset to have a mean of 0 and a standard deviation of 1, using the standard transformation:

$$z = \frac{(x - \mu)}{\sigma} \quad (4.8)$$

- The *Herfindahl-Hirschman Index (HHI)* remains unchanged, as it is already bounded between 0 and 1.
- The *Trustworthiness Index (TI)* is log-transformed to amplify small differences in values, making variations more distinguishable.

An example of the final attributes for each country used in the single layer graphs, can be seen in Table 4.3.

Multilayer Graphs										
Country ID	COI	ECI	# Prod.	SRCA	Geo-Positional Embedding			HHI	TI	
12	-1.253	-0.989	0.307	-0.186	-0.562	0.547	...	-0.689	0.302	-0.016
380	0.686	1.363	0.958	-1.000	-0.542	1.230	...	1.823	0.095	-0.003

Table 4.3: Country attributes (Multilayer Graphs) for Algeria (ID 12) and Italy (ID 380). Geo-Positional Embeddings are shortened due to lack of space.

In the case of the Heterogeneous Multigraph (Section 4.2.1), I exclude the ‘Number of Products’ and ‘SRCA’ attributes from the node features,

as these are commodity-specific. Additionally, the ‘Trustworthiness’ attribute is redefined as an edge attribute to better reflect the reliability of trade reporting between specific country pairs. An example of the resulting features can be seen in Table 4.4.

Multi-Graphs								
Country ID	COI	ECI	Geo-Positional Embedding				HHI	
12	-1.253	-0.989	-0.562	0.547	...	-0.102	-0.689	0.302
380	0.686	1.363	-0.542	1.230	...	0.140	1.823	0.095

Table 4.4: Country attributes (Multi-Graphs) for Algeria (ID 12) and Italy (ID 380). Geo-Positional Embeddings are shortened due to lack of space.

At the end of this process, I construct a node-feature matrix (denoted as h^0)⁵, for each graph. This matrix has dimensions $|V| \times D$, where D is the total number of node attributes obtained through concatenation. Formally, it is represented as:

$$h^0 = \begin{bmatrix} v_1, a_1 & v_1, a_2 & \cdots & v_1, a_m \\ v_2, a_1 & v_2, a_2 & \cdots & v_2, a_m \\ \vdots & \vdots & \ddots & \vdots \\ v_n, a_1 & v_n, a_2 & \cdots & v_n, a_m \end{bmatrix}$$

This representation ensures that all relevant attributes are incorporated into the graph structure. Notice that the final shape for each graph will have a variable number of rows, since not all countries are present for each year-commodity trading graph, but the number of features will remain invariant ($D = 14$ for single graphs, and $D = 11$ for multigraphs). These node-level attributes collectively encode economic, geographic, and relational factors that contribute to each country’s vulnerability and role in the trade network.

4.2.5 Trading as Edges

In addition to node attributes, edges can also carry properties that differentiate trade connections. These attributes serve as an extension or

⁵I use as notation h to differentiate this representation from the traditional plain list of nodes V . The superscript zero indicates the initial state of this embedding, which will be subsequently updated during the training phase of the GNN (see Section 4.4).

generalization of edge weights, providing additional contextual information about trade relationships. Formally, an edge e is represented as a tuple of attributes:

$$e = (a_1, a_2, a_3, \dots, a_n)$$

The following edge attributes are used in this work:

1. **Transaction volume** – The total trade value (USD) between two countries, used as the primary edge weight. This value is log-transformed to ensure a more balanced representation. Previous research [45] has observed that the distribution of flux between countries (i.e., links) is heavy-tailed, meaning that only a small percentage of all the connections in the network carry most of its total flow. For example, major economies such as the United States or China engage in disproportionately large trade flows compared to smaller economies. Log-scaling preserves relative differences across several orders of magnitude, aligning with economic intuition (e.g., a tenfold increase in trade is often more meaningful than an absolute increase). From a technical perspective, this transformation also contributes to more stable training in GNNs, as large raw values may cause exploding gradients.
2. **Product Complexity Index (PCI)** – The average PCI of the traded products within the commodity class (Section 3.2). This captures the complexity of exchanged goods. This index is already normalized by [36].
3. **Number of unique products** – The count of distinct products traded within the commodity class. This reflects the diversity of the trading relationship and is also scaled using Equation 4.8.
4. **Trade Agreements** – The number of active trade agreements between countries, obtained from the Global Preferential Trade Agreements Database (Section 3.4). This attribute reflects the presence of institutional arrangements that may facilitate trade, reduce barriers, or improve trust between partners (more details in Appendix D.1). This value is scaled using Equation 4.8.
5. **Trustworthiness** (Multigraph only) - The trustworthiness (Section 4.2.4) computed for the connecting nodes.

6. **Layer Embedding** (Multigraph only) - The learned layer embedding (Section 4.2.3) for the specific commodity represented in the edge. This allows to better characterize the edge.

An example of the final edge attributes for each flow can be seen in Table 4.5 (multilayer) and Table 4.6 (multigraph).

Source	Target	Export Value	Avg. PCI	# Prod.	Trade Agreements
12	380	12.183	-1.558	-0.025	-0.168
380	12	12.923	-1.042	-0.938	-0.168

Table 4.5: Edge attributes (Multilayer Graphs) for edge Algeria (12) → Italy (380) and Italy → Algeria. Export value serves as an edge weight, while the Avg. PCI of the transacted products, the number of distinct products, and the number of trade agreements serve as edge attributes.

Source	Target	Export Value	Avg. PCI	# Prod.	Trade Agreements	Trust-worthiness	Layer Embedding
12	380	12.183	-1.558	-0.025	-0.168	-0.010	0.125 ... 0.951
380	12	12.923	-1.042	-0.938	-0.168	-0.010	0.125 ... 0.951

Table 4.6: Edge attributes (Multi-Graphs) for edge Algeria (12) → Italy (380) and Italy → Algeria. Export value serves as an edge weight, while the Avg. PCI of the transacted products, the number of distinct products, the number of trade agreements, the trustworthiness between countries and the layer embedding serve as edge attributes.

Throughout this study, whenever the graph architecture supports edge attributes, all available features will be utilized. In cases where a model does not accommodate edge attributes (e.g., MLPs or certain GNN variants), only the edge weight (transaction volume) will be used. These limitations will be stated explicitly in the relevant sections.

4.3 Node labels

A key approach to assessing the resilience of a trade network is to examine the impact of supply disruptions on individual countries. To do this, I analyse historical trade data from 2012 to 2022 to identify nodes that can be classified as "*affected importers*" resulting from "*lost exporters*." Following [36], and grounded in Ricardo's theory [6], I consider a country u as an exporter of commodity c if its **SRCA** (Section 4.2.4) in exporting that commodity is greater than or equal to 1. In this work, a country is labelled as a *lost exporter* for a specific commodity from year n to year

$n + 1$ if the country was deemed an exporter at year n ($SRCA \geq 1$), but stopped being an exporter at year $n + 1$ ($SRCA < 1$), and it experiences a SRCA decline of at least 20%. This threshold aligns with prior studies: for instance, [13] used a 10% reduction in supply, while [14] applied a 20% decrease in production by the initiating country. In a related study, [52] uses three different scenarios with shocks of the magnitude of 10%, 30% and 50% in firm suppliers. Thus, when a country ceases to be a relatively significant exporter of a given commodity, it may trigger downstream disruptions in global supply chains, as documented by [19] and [52].

The first step in this process is to compute the change in SRCA for each country-commodity pair from year n to year $n + 1$. Formally, for each year n in the range 2012 to 2021, the SRCA change for a country node u in commodity layer l_c is defined as:

$$\Delta SRCA(u, l_c, y_{n \rightarrow n+1}) = \frac{SRCA(u, l_c, y_{n+1})}{SRCA(u, l_c, y_n)} - 1$$

Next, for each country node u where the net change in SRCA is less than -20%, I identify its importers, denoted as $I(u)$, as follows:

$$I(u, l_c, y_{n+1}) = \{v \in V \mid (u, v) \in E_{l_c, y_{n+1}}\}$$

where $E_{l_c, y_{n+1}}$ represents the trade edges for commodity layer l_c in year y_{n+1} . I then assess whether these importers have been significantly impacted by the *lost exporter*. A country node v is classified as an *affected importer* if at least one of the following conditions holds:

1. The total import value of the commodity for the importer has decreased by at least 20%.
2. The price paid for the commodity by the importer has increased by at least 20% relative to the overall commodity price change⁶

If either of these conditions is met, the country is labelled as an *affected importer* of the *lost exporter*.

To illustrate the labelling process, consider Bulgaria as an exporter of ‘Live Animals’ (commodity class ‘01’). In 2012, its SRCA was 1.141, classifying it as an exporter of this commodity. By 2013, its SRCA had

⁶The overall price change is calculated as the average price change paid by all importers in a given year, mitigating effects due to inflation or external factors.

dropped to 0.875, a decline of approximately 23%, causing Bulgaria to lose its *exporter status*.

Now, let's examine the impact on some of its importers:

- Bosnia-Herzegovina was one of Bulgaria's trading partners for '*Live Animals*'. In 2012, Bosnia imported 59,809,131 USD of this commodity, but in 2013, this amount dropped to 47,301,110 USD, a decrease of nearly 21%. Since this satisfies our first condition (a decline in import value $\geq 20\%$), Bosnia is labelled as an *affected importer* of Bulgaria's *lost exporter* status.
- Denmark, another importer of '*Live Animals*' from Bulgaria, saw its import value increase slightly from 48,412,095 USD (2012) to 50,454,862 USD (2013) (+4% change). While this does not meet the first criterion, we examine the second condition: the price paid per cubic meter. In 2012, Denmark paid 6.72 USD/ m^3 , which surged to 10.72 USD/ m^3 in 2013, a +59% increase. By comparison, the global average price increased from 3.03 USD/ m^3 to 3.10 USD/ m^3 , a modest +2.44% rise. If Denmark's price increase had followed the global trend, we would expect a price of ~ 6.90 USD/ m^3 rather than 10.72 USD/ m^3 . Since Denmark's price increase is significantly above the global average, it meets our second condition and is also labelled as an *affected importer*.
- Albania, on the other hand, imported 26,669,729 USD in 2012 and 26,211,547 USD in 2013, a -1.7% change. Its price per cubic meter shifted only slightly from 1.90 USD/ m^3 in 2012 to 1.92 USD/ m^3 in 2013 (1.3% below the overall commodity price increase of 2.3%). Since neither the import value nor the price change meets our conditions, Albania is not classified as an *affected importer* of Bulgaria's *lost exporter* status.

A graphical depiction of this situation can be seen in Figure 4.9.

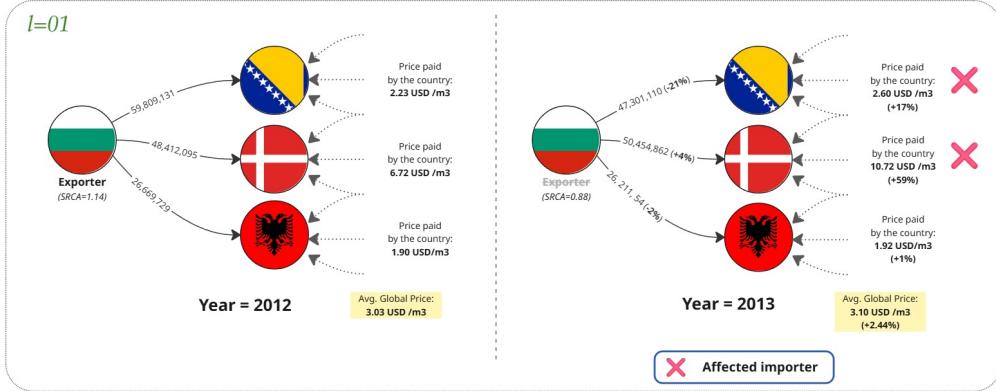


Figure 4.9: Node labelling process. Countries that meet at least one of the conditions from Section 4.3 are deemed “affected importers” of a given “lost exporter.”

4.4 Node Classification

In this work, I use node classification to identify when a country (node) is an *affected importer* within the global trade network. To perform this classification, I test two deep learning architectures from the family of “*message-passing graph neural networks*”. They can be understood as generalizations of previous architectures used in different domains. The main generalization is that these models can make use of a graph’s topology when aggregating information, and they are not tied to a specific structure.

Message-passing Graph Neural Networks (GNN) are particularly suited for analysing international trade networks because they naturally incorporate network structure, allowing each country to dynamically aggregate economic information from trade partners.

Specifically, I consider three variants: **GraphSAGE** [53], a general inductive framework that uses node features to learn a function that generates embeddings by sampling and aggregating features from a node’s local neighbourhood; **Graph Convolutional Networks (GCN)** [54], which aggregate neighbour information through a normalized graph Laplacian; and **Graph Attention Networks (GATv2)** [55], which uses attention mechanisms to dynamically weigh neighbour importance.

Each node begins with a feature vector, constructed as outlined in Section 4.2.4. Recall the initial feature matrix defined as h^0 is a $|V| \times D$ matrix, associating each node with a vector of length D of features. With

$\mathbf{h}_u^k \in \mathbb{R}^D$ denoting the node features of node u in layer k and $\mathbf{e}_{v,u} \in \mathbb{R}^D$ denoting edge features from node v to node u , a message passing GNN can be described as:

$$\mathbf{h}_u^{(k+1)} = \gamma^{(k+1)} \left(\mathbf{h}_u^k, \oplus_{v \in \mathcal{N}(u)} \phi^{(k+1)}(\mathbf{h}_u^{(k)}, \mathbf{h}_v^{(k)}, \mathbf{e}_{v,u}) \right)$$

where \oplus denotes a differentiable, permutation invariant function (e.g., sum, mean, max), and γ and ϕ are differentiable transformations (typically neural networks such as Multi-Layer Perceptrons). The neighbourhood set $\mathcal{N}(u)$ refers to the nodes directly connected to node u in the graph.

In simpler words, at the end of each layer k of the GNN, the node embedding $\mathbf{h}_u^{(k)}$ will get updated with information coming from its neighbours. Different architectures specify particular forms of aggregation and transformation functions, as well as the number of layers. A high-level diagram of this framework can be seen in Figure 4.10. While additional layers allow embeddings to incorporate broader graph-contextual information, too many layers may lead to over-smoothing, causing embeddings to lose node-specific distinctions. I empirically explore this trade-off by performing hyperparameter tuning using Optuna [56]⁷, an open source optimization framework for efficient hyperparameter search (see Appendix F for more details).

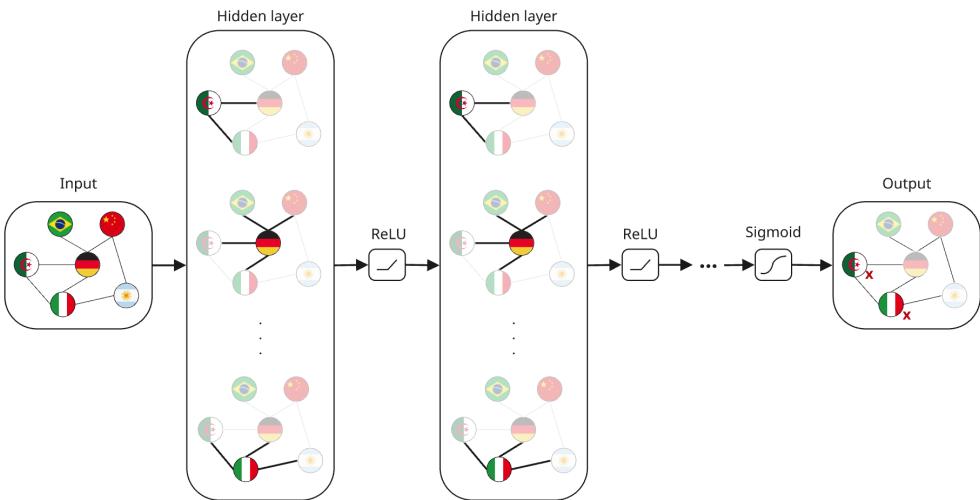


Figure 4.10: Multi-layer Graph Neural Network (GNN) for binary node classification of the ITN.

In the following sections, I expand further on all architectures.

⁷<https://optuna.org/>

4.4.1 Model Architectures

Graph Convolutional Network (GCN)

Convolutional Neural Networks (CNN) have been extensively studied and widely used in computer vision tasks for several decades. Recently, [54] introduced the **Graph Convolutional Network (GCN)**, a variant of CNN that extends convolution operations to graph structures. The key innovation of GCNs is replacing the strictly Euclidean relationships observed in images—where pixels have explicit spatial neighbours—with relationships defined through a graph’s adjacency matrix, thereby accommodating the non-Euclidean, spectral geometry inherent in graph structures.

Following [54], a multi-layer GCN is defined by the following layer-wise propagation rule for layer k :

$$\mathbf{H}^{(k+1)} = \sigma(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{H}^{(k)} \mathbf{W}^{(k)})$$

where:

- $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}$ (i.e., the adjacency matrix with self-connections),
- $\tilde{\mathbf{D}}$ is the diagonal degree matrix (with self loops),
- $\mathbf{W}^{(k)}$ is a layer-specific trainable weight matrix,
- $\sigma(\cdot)$ is a non-linear activation function,
- $\mathbf{H}^{(k)} \in \mathbb{R}^{N \times D}$ is the matrix of activations in the k^{th} layer, where $\mathbf{H}^{(0)} = \mathbf{X}$ (e.g., the initial node features).

This structure allows for stacking convolutional layers, whereby the output of one layer becomes the input to the next. For node classification, the final layer’s ($k = n$) activations are passed through a sigmoid function to generate probabilities for each node belonging to the target class:

$$Z = \text{sigmoid}(H^n) \tag{4.9}$$

An important aspect of the GCN model is its ability to incorporate edge weights by directly introducing them into the adjacency matrix; however, it cannot utilize additional edge attributes beyond weights. Therefore, in this study, GCNs exclusively use the "Export Value" from Section 4.2.5 as the edge weight in the adjacency matrix.

To determine the optimal model parameters, I employ hyperparameter tuning using Optuna. The hyperparameter search space, inspired by [54], includes:

- Number of Layers (how many convolutional layers to stack in the model): 1, 2, and 3;
- Hidden layer size (the size of the embedding): 8, 16, 32, 64, and 128;
- Learning rate (controls how much the weights are updated each time the model is trained): from 1e-4 to 1e-2;
- Weight decay (regularization technique applied to the weights): from 1e-6 to 1e-2;
- Dropout rate (the dropped-out neurons do not contribute to the computation of activations): from 0.1 to 0.5;
- Optimizer (responsible for updating the weights of the neural network during the training process): Adam [57], AdamW [58], SGD [59], RMSprop [60], and Adagrad [61].

GraphSAGE

GraphSAGE is a graph neural network algorithm specifically designed for inductive node embedding problems—i.e., generalizing to unseen nodes. In their work, [53] extend the traditional GCN to inductive unsupervised learning and propose a framework that learns aggregation functions capable of generalizing across different graph structures.

Unlike approaches based on matrix factorization, GraphSAGE leverages node features to learn a function that maps input features to embeddings. Rather than training a unique embedding for each node, GraphSAGE trains a set of aggregator functions that learn to aggregate information from a node’s local neighbourhood. Each aggregator corresponds to a different hop distance from the central node, effectively expanding the receptive field as layers increase. The core intuition is that nodes iteratively aggregate features from increasingly distant neighbours, enabling representations that incorporate multi-hop relational context.

The GraphSAGE operator is defined as:

$$H_u^{(k+1)} = W_1^{(k)} H_u^{(k)} + W_2^{(k)} \cdot \text{AGG}_{v \in \mathcal{N}(u)} H_v^{(k)}$$

where:

- $H_u^{(k)} \in \mathbb{R}^D$ is the feature representation of node u at layer k , with $H_u^{(0)} = x_u$ (the initial node features),
- $W_1^{(k)}$ and $W_2^{(k)}$ are layer-specific learnable weight matrices,
- AGG is an aggregation function (e.g., *mean*, *max*, or *LSTM*),
- $\mathcal{N}(u)$ denotes the set of neighbours of node u .

In the original formulation, [53] propose subsampling a fixed-size neighbourhood for efficiency. However, in this work, I use the complete neighbourhood of each node to ensure fair comparisons across models and leverage the relatively modest graph sizes.

GraphSAGE supports several aggregation strategies, including `mean`, `max`, and `LSTM`. Importantly, it applies separate learnable transformations to the root node and its neighbours, a property known as "root weighting." Since the `mean` aggregator (without root weighting) behaves similarly to a 1-layer GCN, I limit my experiments to the `max` and `LSTM` aggregators.

Additionally, GraphSAGE allows optional feature normalization using ℓ_2 -normalization:

$$h_u = \frac{h_u}{\|h_u\|_2}$$

An additional architectural variation involves applying a linear transformation and non-linear activation to the node features before aggregation:

$$\text{AGG} \left(\sigma \left(W H_u^{(k)} + b \right) \right)$$

Unlike GAT and GCN, this architecture does not incorporate edge attributes, making it structurally distinct from the other models considered in this work.

To identify the optimal model configuration, I employ hyperparameter tuning with Optuna. The search space, inspired by [53], includes:

- Number of layers: 1, 2, and 3;

- Hidden layer size: 8, 16, 32, 64, and 128;
- Learning rate: from 1e-4 to 1e-2;
- Weight decay: from 1e-6 to 1e-2;
- Dropout rate: from 0.1 to 0.5;
- Optimizer: Adam [57], AdamW [58], SGD [59], RMSprop [60], and Adagrad [61];
- Aggregator: max and LSTM
- Normalize output: True and False
- Project before aggregation: True and False
- Bias term: True, and False.

Graph Attention Networks (GATv2)

Graph Attention Networks (GATs) are among the most popular graph neural network architectures, widely considered state-of-the-art for representation learning on graphs. Inspired by the attention mechanism introduced in Natural Language Processing (NLP) in the famous paper "*Attention is all you need*", by [62], GATs dynamically compute the relative importance of nodes in a neighbourhood rather than aggregating all neighbours just as in GCNs. Specifically, attention allows the model to assign higher weights to neighbours that are most relevant or similar to the node being updated.

First introduced by [55], this architecture leverages masked self-attentional layers to efficiently weight the features of neighbour nodes. By stacking multiple attention layers, GATs specify different attention weights to neighbours without expensive matrix operations or a priori knowledge of the entire graph structure.

This work employs a variant known as GATv2 [63], which addresses a limitation of the original GAT: attention scores in the original GAT were independent of the query node during ranking. GATv2 resolves this by changing the order of operations—applying the non-linearity (LeakyReLU) before the linear projection with vector \mathbf{a} , and performing the feature transformation \mathbf{W} after concatenation (see Equation 4.4.1 below). This adjustment allows the model to produce more expressive, query-conditioned attention weights.

Formally, GATv2 computes an attention scoring function s_e for each edge (v, u) , indicating the importance of node v to node u :

$$s_e(\mathbf{h}_u, \mathbf{h}_v, \mathbf{e}_{uv}) = \mathbf{a}^\top \text{LeakyReLU}(\mathbf{W} \cdot [\mathbf{h}_u || \mathbf{h}_v || \mathbf{e}_{uv}])$$

where $\mathbf{W} \in \mathbb{R}^{d' \times d}$ is a learnable feature transformation, $\mathbf{a} \in \mathbb{R}^{2d'}$ computes scalar attention scores, and $||$ denotes vector concatenation. Optionally, a bias vector \mathbf{b} can be added before applying the non-linearity. These attention scores are normalized across all neighbours $v \in \mathcal{N}(u)$ using the *Softmax* function, resulting in attention coefficients $\alpha_{u,v}$ that represent how much node u should attend to node v :

$$\alpha_{u,v} = \text{softmax}_v(s_e(\mathbf{h}_u, \mathbf{h}_v, \mathbf{e}_{uv})) = \frac{\exp(s_e(\mathbf{h}_u, \mathbf{h}_v, \mathbf{e}_{uv}))}{\sum_{v' \in \mathcal{N}(u)} \exp(s_e(\mathbf{h}_u, \mathbf{h}_{v'}, \mathbf{e}_{uv'}))}$$

Finally, the new embedding of node u is computed as a weighted average of its neighbours' transformed features, followed by a non-linear activation σ :

$$\mathbf{h}_u^{k+1} = \sigma \left(\sum_{v \in \mathcal{N}(u)} \alpha_{uv} \cdot \mathbf{W} \mathbf{h}_v^k \right)$$

As with GCN, node classification is done by applying a sigmoid function to the last embedding, as in Equation (4.9).

Note that GATs allow for a richer representation of edges by incorporating full edge attributes (i.e., the embeddings denoted as \mathbf{e}_{ij}). In practice, the complete attribute vector described in Section 4.2.5 is used.

Similarly to GCN, I employ hyperparameter tuning (using Optuna) to identify optimal parameters for the GATv2 model. The hyperparameter search space, inspired by [63], includes:

- Number of Layers (how many attention layers to stack in the model): 1, 2, 3, and 6;
- Hidden layer size (the size of the embedding): 8, 32, 64, and 128;
- Heads: (number of multi-head-attentions): 1, 2, 4, and 8;
- Learning rate (controls how much the weights are updated each time the model is trained): 5e-4, 1e-3, 5e-3, and 1e-2;

- Weight decay (regularization technique applied to the weights): from 1e-6 to 1e-2;
- Dropout rate (the dropped-out neurons do not contribute to the computation of activations): 0.4, 0.6, and 0.8;
- Bias (whether to add an additive bias term \mathbf{b} or not): True and False;
- Residual connections (whether to add a learnable skip-connection to the layers): True and False;
- Optimizer: Adam [57], AdamW [58], SGD [59], RMSprop [60], and Adagrad [61].

Model Ensemble

In addition to evaluating each GNN model independently, I also implement a simple ensemble strategy to combine the strengths of the individual architectures: GCN, GATv2, and GraphSAGE. Specifically, I adopt a majority-voting scheme in which each model casts a hard classification vote for every node. A node is labelled as positive (“*Affected Importer*”) if at least two of the three models predict it as such.

This approach aims to capture potential complementarities among models: where one model may misclassify a node due to limitations in structure sensitivity or feature aggregation, others may succeed. Ensemble methods are known to improve generalization by reducing individual model variance, especially when model errors are not strongly correlated.

An added benefit of this ensemble setup is the ability to systematically examine disagreement cases which can offer insights into failure modes or edge cases that challenge the predictive capacity of GNN architectures.

4.4.2 Baseline

I compare the performance of the proposed GNN models with two baseline models: a Random Predictor and a simple Multi-Layer Perceptron (MLP). The random predictor serves as an absolute lower bound, while the MLP baseline provides a meaningful lower bound by relying exclusively on node attributes, disregarding any relational information from the trade network.

- The **Random Predictor** predicts node labels at random, following a uniform distribution (i.e., both labels are equally probable).
- The **simple MLP** baseline uses only the node feature matrix described in Section 4.2.4, ignoring the trade network's relational structure. Thus, the performance of the MLP baseline can reveal whether node features alone contain sufficient predictive information or if relational dependencies captured by GNNs significantly enhance prediction.

The baseline MLP consists of one hidden layer, with ReLU activation [64] applied after the first layer:

$$\mathbf{h} = \text{ReLU}(\mathbf{W}_1 \mathbf{h}^0 + \mathbf{b}_1)$$

$$\mathbf{y} = \sigma(\mathbf{W}_2 \mathbf{h} + \mathbf{b}_2)$$

where:

- $\mathbf{h}^0 \in \mathbb{R}^{14}$ is the node feature matrix,
- $\mathbf{W} \in \mathbb{R}^{16 \times 14}$ and $\mathbf{b}_1 \in \mathbb{R}^{16}$ are the weights and biases for the layers,
- $\mathbf{h} \in \mathbb{R}^{16}$ is the hidden state after the first layer,
- $y \in \mathbb{R}^1$ is the final classification score, passed through sigmoid activation (σ) for binary classification.

A graphical depiction of this architecture is shown in Figure 4.11.

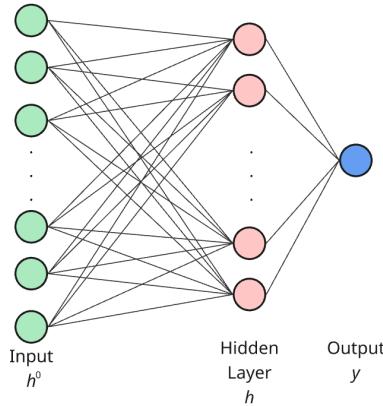


Figure 4.11: The baseline MLP architecture consists of two linear layers: the first maps input features to a hidden representation of size 16, followed by a second layer outputting a single classification score.

The baseline MLP is optimized (using Optuna) with the following hyperparameter search space:

- Hidden layer size (the size of the embedding): 8, 32, 64, and 128;
- Learning rate (controls how much the weights are updated each time the model is trained): from 1e-4 to 1e-2;
- Dropout rate (the dropped-out neurons do not contribute to the computation of activations): 0.1, 0.2, 0.3, 0.4, and 0.5.

4.5 Experimental setup

Since the node labels defined in Section 4.3 are transitional between years—meaning the label of a node in year $n + 1$ depends on changes occurring from year n to year $n + 1$ —node features from each year can be matched with labels from the subsequent year. Specifically, the goal is to predict whether a country will be affected by a supply disruption in the next year, based on the trade network structure and country attributes of the current year.

Given 11 years of data spanning 2012 to 2022, there are 10 transitional years with labels available for model training and evaluation. This is because node attributes and graph topology from 2022 would predict disruptions occurring in 2023, which are unavailable in this work. Thus, I partition the dataset into two sets:

- **Training and Validation Set:** The first nine transitional periods (2012→2013 through 2020→2021; 96 different commodity layers; 864 individual graphs for the commodity-layered approach).
- **Test Set:** The final transitional period (2021→2022; 96 graph/layers), reserved exclusively for evaluating model performance.

4.5.1 Data Splitting Strategy

To properly train and validate the models, I split the training and validation graphs with an 80/20 ratio. Given the class imbalance—where positively labelled nodes (*affected importers*) are relatively rare—I implement a stratified splitting strategy:

1. First, separate graphs into two groups:
 - a) Graphs containing at least one positively labelled node.

- b) Graphs with no positively labelled nodes.
- 2. Then apply an 80/20 split independently within each group.
- 3. Finally, combine these splits to form balanced and representative training and validation sets, ensuring both sets contain positive examples.

4.5.2 Model Training

All GNN models (GCN, GATv2 and GraphSAGE) are trained using the optimal hyperparameters determined by the Optuna optimization framework (see Appendix F.1 for details). Models are trained for up to 500 epochs, with an early stopping strategy implemented based on validation loss. If the validation loss does not improve after 50 epochs (patience = 50), training is halted. During training, the model weights corresponding to the lowest observed validation loss are saved and used for evaluation.

Loss Function

Due to the class imbalance and the specific interest in correctly identifying countries vulnerable to disruptions, I employ a customized loss function based on the **F₁-score** and inspired by [65], termed *Soft F1 Loss*. This loss function emphasizes the correct prediction of the positive class, penalizing errors proportionally to the ratio of positives to negatives, thus addressing imbalance explicitly. Mathematically, the *Soft F1 Loss* is defined as:

$$\text{SoftF1Loss} = (1 - F_1(\text{positives})) \times \frac{\#\text{positives}}{\#\text{negatives}}$$

where:

$$F_1(\text{positives}) = \frac{2 \times \text{Precision}_{\text{positives}} \times \text{Recall}_{\text{positives}}}{\text{Precision}_{\text{positives}} + \text{Recall}_{\text{positives}}} \quad (4.10)$$

Precision is the number of true positive results divided by the number of all samples predicted to be positive, including those not identified correctly; while recall is the number of true positive results divided by the number of all samples that should have been identified as positive:

$$\text{Precision}_{\text{class}=1} = \frac{TP_{\text{class}=1}}{TP_{\text{class}=1} + FP_{\text{class}=1}}$$

$$\text{Recall}_{\text{class}=1} = \frac{TP_{\text{class}=1}}{TP_{\text{class}=1} + FN_{\text{class}=1}}$$

In these definitions:

- **TP (True Positives)**: Countries correctly predicted as affected by disruptions.
- **FP (False Positives)**: Countries incorrectly predicted as affected.
- **FN (False Negatives)**: Countries incorrectly predicted as unaffected.

Finally, the factor $\frac{\# \text{positives}}{\# \text{negatives}}$ scales the loss based on class imbalance.

4.6 Evaluation metrics

Given the practical objective of identifying countries vulnerable to disruptions, the model evaluation focuses primarily on capturing positive labels accurately (high recall) while avoiding incorrect positive predictions (high precision). Therefore, the main evaluation metric is the F_1 -score computed specifically for the positive class (Equation 4.10). The F_1 -Score is the harmonic mean of the precision and recall, and represents both precision and recall jointly. Its highest possible value of score is 1, indicating perfect precision and recall, and the lowest possible value is 0, if the precision or the recall is zero.

Maximizing this F_1 -score ensures a balanced performance that prioritizes accurate detection of genuinely vulnerable countries while minimizing false alarms.

Chapter 5

Results

“If we knew what we were doing, it wouldn’t be called research.”

Albert Einstein

This chapter reports the main experimental findings obtained with the modelling framework described in Section 4.5. All results are produced with the best hyperparameter configuration selected by Optuna for each model/graph combination. Table 5.1 summarises those key settings. For more details on these parameters, please refer to Appendix F.

The evaluation of model performance includes comparisons across the MLP baseline (Section 4.4.2), three GNN architectures: GCN (Section 4.4.1), GATv2 (Section 4.4.1), GraphSAGE (Section 4.4.1), and an Ensemble model (Section 4.4.1). Results are categorized by different graph configurations:

- **Total (T)**: Commodity-layered, Total Trade graph (Section 4.2.2.1a)
- **Exports (E)**: Commodity-layered, Exports graph (Section 4.2.2.1b)
- **Exports + Layer Embedding (E+L)**: Exports graph augmented with commodity-specific node embeddings (Section 4.2.2.1c)
- **Multi-Graph Total (MG T)**: Multi-Graph, Total Trade graph (Section 4.2.2.2a)
- **Multi-Graph Export (MG E)**: Multi-Graph, Exports graph (Section 4.2.2.2b)

Graph	Model	Layers	Channels	Heads	Learning Rate	Weight Decay	Dropout
E	MLP	1	64	—	0.0024	—	0.1817
	GCN	2	64	—	0.0023	0.0014	0.1158
	SAGE	2	128	—	0.0016	0.0058	0.1192
	GATv2	2	128	4	0.0010	0.0074	0.4000
T	MLP	1	64	—	0.0037	—	0.3337
	GCN	2	64	—	0.0027	0.0029	0.1778
	SAGE	3	128	—	0.0003	0.0083	0.1005
	GATv2	2	128	8	0.0005	0.0089	0.6000
E+L	MLP	1	64	—	0.0033	—	0.2283
	GCN	2	32	—	0.0063	0.0003	0.1108
	SAGE	2	128	—	0.0095	0.0001	0.1372
	GATv2	3	128	8	0.0100	0.00003	0.4000
MG T	MLP	1	128	—	0.0079	—	0.3845
	GCN	2	128	—	0.0090	0.0060	0.1826
	SAGE	2	8	—	0.0097	0.0093	0.1196
	GATv2	3	32	8	0.0010	0.0065	0.4000
MG E	MLP	1	128	—	0.0069	—	0.2920
	GCN	2	64	—	0.0087	0.0023	0.2789
	SAGE	1	64	—	0.0096	0.0024	0.4650
	GATv2	2	32	4	0.0050	0.0033	0.8000

Table 5.1: Best-performing hyperparameters selected by Optuna for each model/graph variant: (E) Export, (T) Total, (E+L) Export + Layer Embedding, (MG T) Multi-Graph Total, (MG E) Multi-Graph Export

For each configuration, I report the mean F_1 -Score for the positive class (“*Affected Importer*”) across ten runs with different random model initializations, alongside the standard deviations (in parentheses). Superscripts indicate statistically significant differences ($\alpha = 0.05$) between models; for example, a superscript ^{1,4} would denote a result significantly different from models 1 and 4. Results are summarized in Table 5.2,

Graph	F_1 (positives): Avg. (Std)				
	MLP ¹	GCN ²	GATv2 ³	SAGE ⁴	ENS ⁵
T	0.098 (0.008)	0.117 (0.008) ^[1]	0.131 (0.013) ^[1,2]	0.126 (0.009) ^[1]	0.142 (0.010) ^[1,2,3,4]
E	0.096 (0.008)	0.106 (0.006) ^[1]	0.124 (0.012) ^[1,2]	0.125 (0.005) ^[1,2]	0.130 (0.005) ^[1,2,4]
E+L	0.101 (0.004)	0.103 (0.004)	0.117 (0.007) ^[1,2]	0.128 (0.007) ^[1,2,3]	0.126 (0.007) ^[1,2,3]
MG T	0.905 (0.004)	0.922 (0.005) ^[1,4]	0.920 (0.011) ^[1]	0.909 (0.012)	0.924 (0.006) ^[1,4]
MG E	0.905 (0.006)	0.910 (0.007)	0.910 (0.012)	0.911 (0.006) ^[1]	0.918 (0.006) ^[1,2,3,4]

Table 5.2: Results across all models and graph configurations. Reported values are the mean and standard deviation F_1 -score over 10 runs with different random initializations. Bold indicates highest value for the graph type (row). Superscripts denote stat. significant differences ($p < 0.05$) compared to the referenced model indicated by the superscript number. All results are significantly better than the Random Predictor baseline.

All evaluated models performed significantly better than the random predictor baseline (mean F_1 of 0.06 for layer-specific graphs, and 0.60 for Multi-Graphs); hence, it is excluded from the table. Figure 5.1 shows a summary of the improvement of the models over the random baseline. In some cases, using a GNN can double on the performance of a difficult task such as the one presented here.

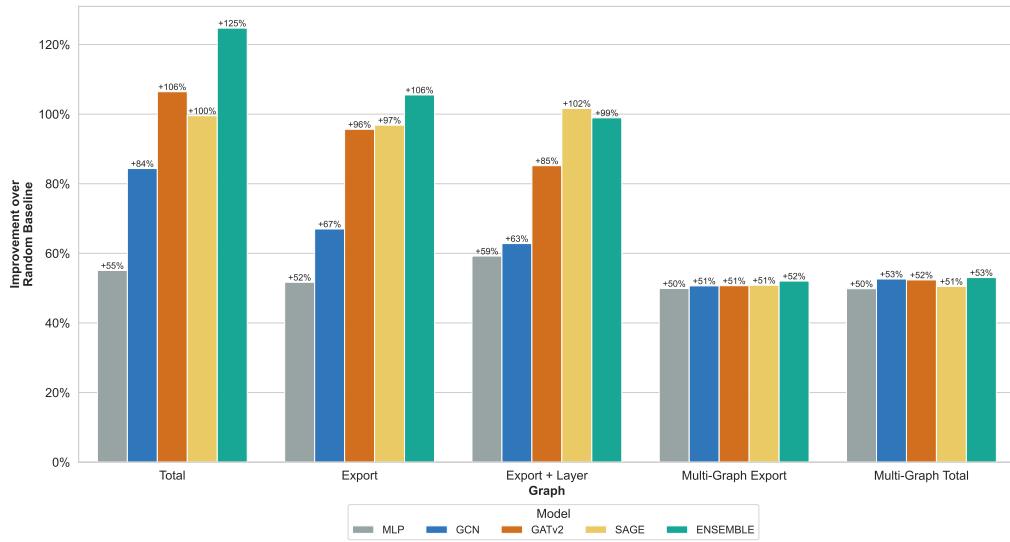


Figure 5.1: Average improvement in $\mathbf{F}_1\text{-score}$ for the positive class over a random baseline. For the Multi-Graphs setup, the baseline $\mathbf{F}_1\text{-score}$ is approximately 0.60, which places an upper bound on the possible relative improvement at around 67% (i.e., $1.00/0.60 \approx 1.67$).

The Ensemble model consistently outperformed individual models across most configurations, underscoring the advantages of combining multiple GNN architectures to capture diverse predictive signals. Among the individual models, GATv2 and GraphSAGE frequently demonstrated superior performance, surpassing GCN and the baseline MLP, particularly in commodity-layered graphs. However, GraphSAGE and GATv2 also exhibited greater variability in performance across initializations, suggesting some instability compared to other methods.

In the E+L configuration, incorporating commodity-specific layer embeddings impacted model performance differently across architectures. Notably, GCN did not outperform the MLP baseline, reflecting minor performance deterioration. Conversely, GraphSAGE maintained strong performance, demonstrating robustness in effectively utilizing the additional embedding information, thus matching the Ensemble's performance. GATv2, while still outperforming MLP and GCN, showed

a slight reduction in performance relative to its results without embeddings.

For Multi-Graph setups (MG T and MG E), performance differences among models narrowed considerably, reflecting the increased baseline performance provided by richer data representations. Nonetheless, in the undirected Multi-Graph (MG T), GCN notably outperformed both MLP and GraphSAGE, achieving comparable performance to GATv2 and the Ensemble. In the directed Multi-Graph (MG E), the Ensemble model was uniquely able to achieve significantly higher performance, underscoring the effectiveness of model combination strategies in complex, directed trade scenarios.

Figure 5.2 presents box-plots illustrating the distribution of F_1 -Scores across models and graph setups. GCN demonstrates notably lower variability, indicating robust but generally lower performance compared to other models. In contrast, Multi-Graph setups show more uniform performance among models, consistently outperforming baselines.

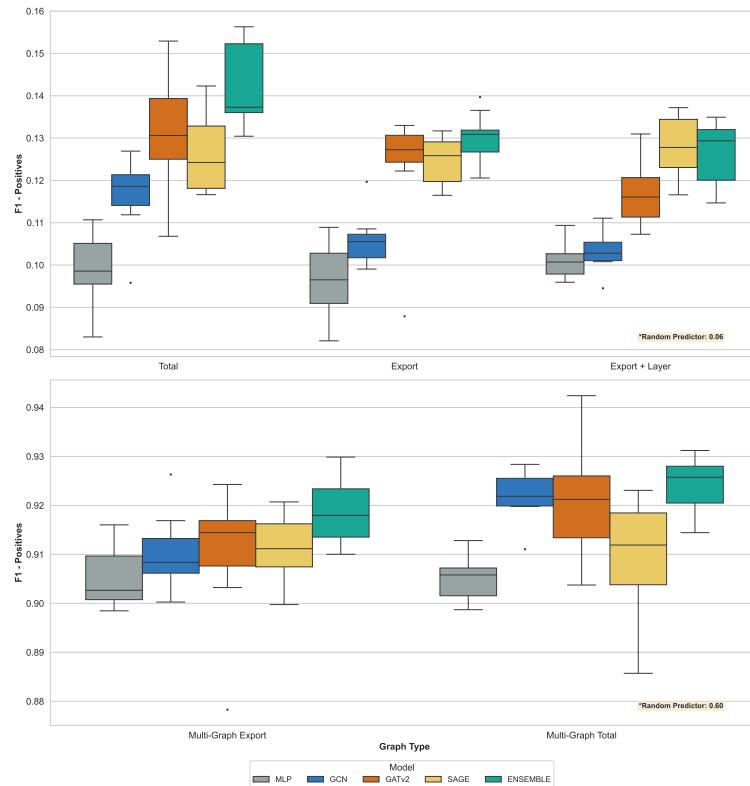


Figure 5.2: Results for models in each graph type. The Random Predictor's value of reference is included in the bottom right of each plot.

Figure 5.3 highlights the ratio of countries impacted by commodity disruptions per year, revealing that over the analysed decade, approximately half of the countries faced disruptions in at least one commodity. Such analyses, extended over longer periods, could provide valuable insights into the evolving dynamics of international trade networks influenced by geopolitical factors and economic policy shifts.

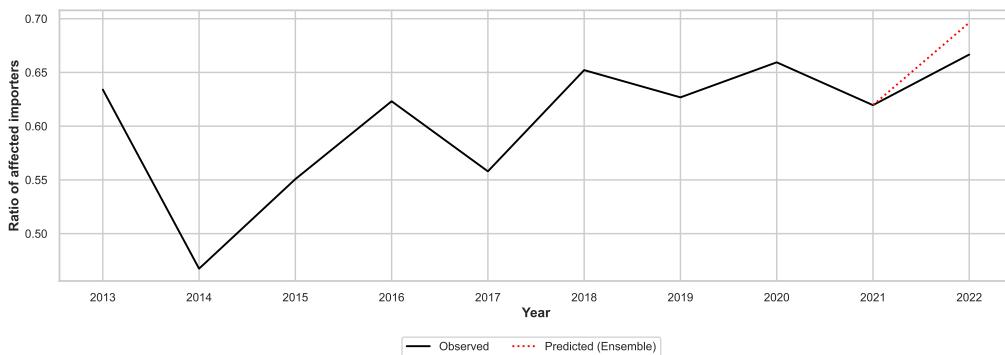


Figure 5.3: Ratio of affected importers per year (black solid line) and predicted ratio using Ensemble of models for last year (red dotted line).

Figure 5.4 further explores vulnerability at the commodity level, illustrating variations in the ratio of affected importers across different commodities and years. Some commodities consistently show higher vulnerability, such as class '71' ('Natural, cultured pearls; precious, semi-precious stones; precious metals, metals clad with precious metal, and articles thereof; imitation jewellery; coin') and '89' ('Ships, boats and floating structures'); while others demonstrate greater resilience, such as '48' ('Paper and paperboard; articles of paper pulp, of paper or paperboard') and '69' ('Ceramic products'). Commodities such as classes '75' ('Nickel and articles thereof'), '93' ('Arms and ammunition; parts and accessories thereof'), and the aforementioned '89' and '71', display notable variability, reflecting dynamic industries subject to rapid shifts in supply and demand conditions driven by factors such as geopolitical conflict, economic fluctuations, or resource availability.

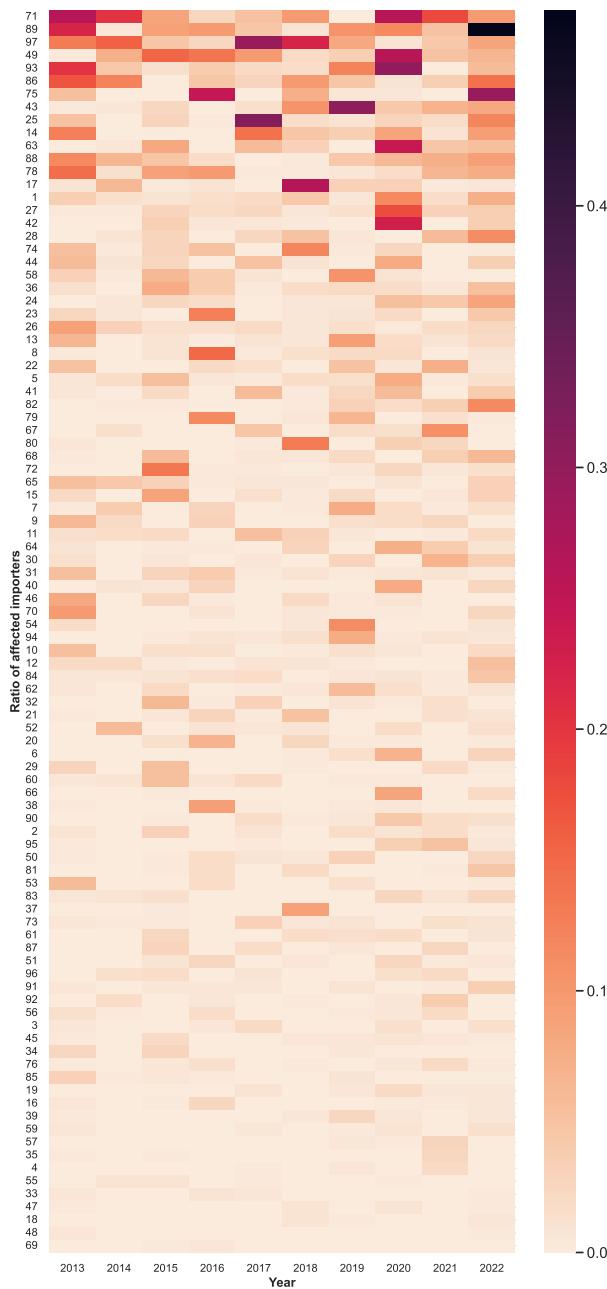


Figure 5.4: Ratio of affected importers per commodity layer and year, sorted by average ratio for the decade.

Chapter 6

Discussion

“We are not what we know but what we are willing to learn.”

Mary Catherine Bateson

In this chapter, I discuss key factors, methodological decisions, and inherent limitations related to the data and methods employed in this research. I critically reflect upon the scope and boundaries of the findings, offering an interpretation of their implications and suggesting potential avenues for future research to further advance our understanding of vulnerabilities within international trade networks.

6.1 Data Limitations

Charles Babbage, often regarded as the “father of the computer,” famously stated, *“Errors using inadequate data are much less than those using no data at all.”* This perspective highlights the critical importance of data quality and comprehensiveness, a point previously addressed in Section 3. Among the inherent limitations acknowledged in this study, the accuracy and completeness of self-reported national trade data stand out prominently. At best, self-reported trade figures by countries exhibit minor inaccuracies due to reporting inconsistencies; at worst, they are entirely absent or severely underreported. Similarly, the World Trade Organization (WTO) acknowledges that some bilateral trade agreements

might not be reported to the organization¹. Such limitations pose substantial challenges, especially when analysing a highly interconnected system like the ITN. Missing or inaccurate data in one country can propagate errors throughout the entire network, thereby skewing system-level insights and interpretations. Although systematic methods, such as data smoothing and imputation, have been applied in this and previous studies to mitigate these discrepancies, the influence of incomplete or inaccurate reporting remains unavoidable.

A second notable limitation concerns the granularity and classification decisions inherent in the data collection process. Choices made during data compilation significantly influence subsequent analytical outcomes. Two prominent examples include commodity classification granularity and temporal categorization. In this research, I utilized commodity data classified at the first two digits of the HS nomenclature, yielding 96 distinct commodity classes. Employing a more granular classification could enhance the analytical precision by capturing more nuanced trade dynamics; however, this would substantially increase computational complexity. For instance, adopting the four-digit HS classification would produce 1,266 distinct commodity classes, whereas employing the full six-digit HS codes results in 6,897 distinct categories. Moreover, commodity classifications inherently involve ambiguity, particularly when similar products fall under different categories due to their intended use or minor differences. For example, lithium batteries intended for mobile phones are typically classified under commodity code '85' ('Electrical machinery and equipment'), whereas similar lithium batteries used for electric vehicles are usually categorized under code '87' ('Vehicles and parts thereof'). Such classification nuances can obscure meaningful economic linkages, complicating the accurate representation and interpretation of commodity-level trade relationships.

Additionally, temporal categorization introduces potential distortions. Exports recorded on December 31st are attributed to one calendar year, while those recorded a day later, on January 1st, fall into the subsequent year. Given the strong reliance of this analysis on transitional period labels (with each year's outcomes depending directly on the previous year's data), these seemingly minor temporal classification choices can significantly affect the accuracy and robustness of the findings, particularly in year-to-year transition analyses.

¹Source: https://www.wto.org/english/thewto_e/whatis_e/tif_e/fact2_e.htm (Last accessed: 14/04/2025)

These data-related constraints underscore the necessity for cautious interpretation and emphasize the importance of explicitly acknowledging such limitations when evaluating the conclusions drawn from this research. Recognizing these boundaries clearly is essential for contextualizing the results and ensuring a nuanced understanding of the insights and recommendations presented.

6.2 Design choices

Throughout this research, several critical methodological choices were necessary when translating the complexities of international trade relationships into a computational framework. In this section, I critically evaluate these choices, their inherent limitations, and the implications they hold for interpreting results.

6.2.1 Creating the Network

Throughout this work, I have explored several methodologies to construct a graph representation of the ITN, drawing inspiration from prior studies (Section 2). However, the true nature of commercial interactions among countries is significantly more complex than any single graph model can fully capture. A graph, in essence, serves as a simplified representation or model of a multifaceted and intricate reality. While such simplifications facilitate the analysis of complex systems and enable clearer insights, they inherently involve sacrificing certain nuances and details. This research is no exception to this general principle.

One significant simplification made in this study pertains to the scope of transactions considered. I have focused specifically on the commercial exchange of tangible goods (referred to as commodities), deliberately omitting purely financial transactions and services. The rationale for this choice is twofold. First, goods transactions are generally more consistently reported by countries, making data availability and accuracy more reliable. Conversely, services and financial movements often suffer from less reliable reporting and are challenging to trace comprehensively. Although previous studies (e.g., [18]) have emphasized financial flows between countries, these financial transactions do not always mirror actual goods movements. For instance, countries can incur debts while importing goods, indicating financial flows that are asynchronous with, or opposite to, the direction of physical trade (such

as deferred payments). Second, tangible goods often directly impact domestic production and consumption more significantly. Financial crises may sometimes be alleviated through debt forgiveness, restructuring, or international bailouts, whereas supply disruptions in tangible goods can have immediate and prolonged adverse effects, potentially halting entire industries without quick recovery options².

Another crucial decision in constructing the network was defining its structural characteristics. Some choices are straightforward: for instance, utilizing a weighted network, as transaction volumes between countries provide critical signals about the strength and significance of commercial ties. However, other decisions involve selecting among equally viable alternatives. In this research, I opted for a unipartite network representation, although alternative approaches, such as a bipartite network (where one type of node represents countries and another represents commodities; see Figure 6.1), were also possible. A significant limitation of the bipartite approach would be the loss of direct country-to-country connections, a feature I deemed essential for analysing direct trade relationships and systemic vulnerabilities.

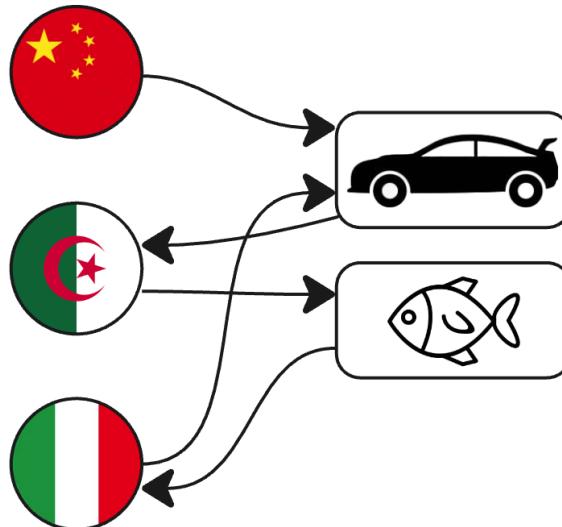


Figure 6.1: ITN as a bipartite network of Countries and Products. This setup allows for edge directionality and weights, but loses the direct country-to-country link.

²A recent example coming from China's restrictions on essential metals: <https://finance-commerce.com/2025/05/critical-minerals-supply-chain-risk-china/> (Last accessed: 22/05/2025).

Additionally, I explored constructing a heterogeneous network characterized by multiple types of edges, each representing a distinct commodity class. While this approach enriches the representation by capturing commodity-specific interactions, it results in an extremely dense network structure, as virtually all countries maintain at least one commodity-specific connection with many other countries. A possible solution could have been employing backbone decomposition techniques to extract a simplified yet informative core network. Nevertheless, given that my networks already incorporate edge weights to reflect trade significance, I chose instead to rely on the capacity of GNNs to implicitly discern the relative importance of these edges during learning.

Regarding the directionality of edges, I explored both directed and undirected graph representations. Previous research by [22] has highlighted a strong correlation ($r=0.91$) between in-degree and out-degree, as well as a considerable link reciprocity correlation ($r=0.61$) in the aggregate ITN. These findings suggest that representing the ITN as undirected could preserve much of its relevant topological information. However, reciprocity and degree correlations are not the full story: there are also edge weights. Moreover, this general observation may not universally apply to individual commodities. Consequently, the directed approach likely captures additional, potentially crucial details regarding commodity-specific trade relationships, thus providing a richer understanding of directional dynamics within the network.

6.2.2 Node Features

This research fundamentally relies on the assumption that there is a causal relationship between exporting countries and importers potentially affected by supply disruptions. This relationship can be influenced by characteristics inherent to both exporters and importers (i.e., node features), as well as by the specifics of their bilateral trade connections (i.e., edge features). Consequently, the manner in which countries and their relationships are characterized holds paramount importance for the validity and interpretability of this study.

Certain features, such as actual trade volume between countries, were intuitive and straightforward to select, serving effectively as weighted edges to indicate relationship strength. However, additional node features necessitated careful deliberation. While previous research often used the Gravity Model framework for predicting trade flows

based on countries' GDP and geographical proximity, I opted for excluding GDP and instead using the more nuanced Economic Complexity Index (ECI), proposed by [36]. Although correlated with GDP per capita, ECI provides deeper insights into economic structure, capturing sophistication and diversity in a country's production capabilities. The choice of ECI is particularly valuable because it inherently relates to a country's economic resilience and substitutability, offering a more comprehensive view of vulnerability beyond mere economic scale.

Nonetheless, the decision to rely predominantly on ECI is not without limitations. Additional economic or institutional indicators, such as governance quality, technological capabilities, or infrastructure robustness, could significantly enhance node characterization. Integrating such indicators could enrich model performance and interpretability, albeit at the potential cost of increased model complexity and multicollinearity. Future research might benefit from systematically evaluating the added value of these alternative attributes through ablation studies or sensitivity analyses, providing insights into the robustness and interpretability of chosen node features.

6.2.3 Geo-Positional embeddings

In this research, I integrated geographical information about countries through geo-positional embeddings (Section 4.2.4). While this approach provides a useful approximation of geographic proximity between countries, it does carry certain inherent limitations. The two most significant ones are:

1. **Uniform border proximity assumption:** The current method treats all bordering countries as equally close. In reality, country size and geographic layout significantly influence the ease and cost of transporting goods to neighbouring nations. For example, Luxembourg, with an area of approximately $2,586 \text{ km}^2$, has a maximum internal distance of around 100 km to its border with France. Conversely, the maximum distance within larger countries such as the United States can be substantial; the distance from Maine (U.S.) to Tijuana (Mexico) is roughly 5,200 km. A more precise approach could involve calculating distances between country centroids or major economic centres to better capture actual terrestrial transport distances.

2. **Maritime port selection:** The CERDI Sea-Distance database employed in this study (Section 3.4) provides maritime distances between the most relevant ports of each country. However, actual international trade frequently involves multiple ports, often choosing those closest or most logically advantageous, rather than strictly the largest or most prominent ports. Thus, the maritime distances calculated may not always represent typical trade routes accurately.

Despite these limitations, I believe the geo-positional embedding method used in this study gives a reasonable and practically effective approximation of geographical distances. Future research could incorporate more detailed geographic data to further refine these distance measures and enhance the precision of trade network modelling.

6.2.4 Layer Embeddings

In the initial experimental setup, where each commodity is modelled as a single-layer graph, the model is trained without explicit awareness of the particular commodity being analysed. While such an agnostic approach might be advantageous in certain contexts, a significant limitation arises: valuable information regarding inherent differences between commodities and their distinct influences on trade relationships is lost. Different commodities may intrinsically impact network robustness or vulnerability in varying degrees, depending on factors such as their substitutability, criticality, or economic complexity.

To mitigate this limitation and explicitly encode commodity-specific information into the model, I developed layer embeddings that encapsulate the nuanced characteristics of each commodity class. Conceptually, these embeddings serve as indicators attached to the nodes, signalling to the model the commodity-specific context, effectively stating: "this node represents a country's role **within the trade network of this particular commodity**."

An immediate question arises concerning the use of the two-digit HS commodity codes themselves as embeddings. Treating these identifiers as continuous variables would erroneously imply an inherent ordinal relationship between classes (e.g., suggesting that commodity class '01' is intrinsically inferior or substantially different from class '90'). Conversely, using these identifiers as categorical variables presupposes that

commodity classes are entirely independent and qualitatively distinct entities without nuanced relationships or similarities. Both approaches are thus inadequate in capturing the true complexity and interrelations inherent within the commodities themselves.

To address this challenge, I adopted an innovative (and possibly controversial) embedding approach inspired by NLP methodologies. Specifically, I generated semantic embeddings by combining textual information derived from the commodity descriptions (using a transformer-based NLP model) with numeric indicators related to their Product Complexity Index (PCI). This method better encapsulates subtle semantic and economic differences between commodity classes.

It is important to acknowledge that by incorporating commodity-specific details into node embeddings, some information typically captured through edge attributes is shifted directly to nodes. However, since this embedding strategy is exclusively applied within single-layer graphs (each dedicated to a single commodity), this shift does not distort or disrupt the graph's fundamental relational structure. Instead, it enriches the node-level representation, potentially enabling the model to more effectively discern nuanced commodity-specific vulnerabilities and interactions within international trade networks.

Future work could explore how these embeddings perform in multilayer or multi-commodity contexts, potentially evaluating their effectiveness in preserving meaningful relationships across commodities and further refining the robustness of this embedding methodology.

6.2.5 Alternative Risk Measures

A key factor incorporated into the countries' features in this work is an indicator capturing systemic risk based on their trading relationships. Previous studies [18] have found that a strongly integrated country can be a double-edged sword: *"a crisis is amplified if the epicentre country is better integrated into the trade network. However, target countries affected by such a shock are in turn better able to dissipate the impact if they are well integrated into the network."*

The underlying assumption is that countries with highly concentrated trade relationships for specific commodities face higher risks of supply disruptions, given their dependency on fewer trading partners. Conversely, a diversified trade landscape—with numerous partners and balanced trade volumes—should theoretically reduce systemic risk. This

intuition draws upon foundational principles of portfolio diversification in finance [66], which emphasize risk mitigation through diversification.

While the Herfindahl-Hirschman Index (HHI) was ultimately chosen for this analysis, several alternative measures were also considered for assessing diversification-related risk. Among these alternatives, the Gini coefficient [67] and an extended Gini-based *Risk Score* were explored in depth. Below, we discuss both alternatives, their strengths, and their limitations relative to HHI.

Gini coefficient

The Gini coefficient is a widely-used measure in economics for assessing statistical dispersion, frequently applied to income, wealth, or consumption inequality. It ranges from 0 to 1, where 0 indicates perfect equality and 1 indicates maximum inequality (e.g., one individual controls all resources). This metric can analogously measure inequality in trade flows, thus quantifying how diversified a country's trade is.

Formally, the Gini coefficient can be computed for trading as:

$$G = \frac{2 \sum_{i=1}^N i \cdot T_i}{N \sum_{i=1}^N T_i} - \frac{N+1}{N} \quad (6.1)$$

where T_i represents the trade volumes sorted in ascending order and $N = |T|$ denotes the total number of trading partners.

To illustrate, using the example from Table 4.2, the Gini coefficients for countries *A*, *B*, and *C* would be approximately 0.214, 0.484, and 0, respectively. This aligns with the ranking provided by the HHI scores discussed in Section 4.2.4.

However, unlike HHI, the Gini coefficient does not differentiate scenarios where trade is equally diversified but involves different numbers of partners. Consider an extended scenario: a country *D* trades equally (volume of 500) with 20 different partners. Here, the Gini coefficient for country *D* would be 0 (implying perfect equality). In this scenario, countries *C* and *D* are considered equally, although country *D* has 10 times more partners. On the other hand, the HHI would yield a score of 0.05 for country *D* (versus the 0.5 of country *C*), reflecting the beneficial effect of increased diversification across multiple partners. The example is summarized in Table 6.1. This nuanced differentiation offered by HHI is valuable, as a greater number of partners inherently mitigates systemic trade risk.

Gini coef. vs. HHI comparison						
Country	Partner	Volume	Share	Gini coef.	HHI	
A	a_1	1000	0.29	0.21	0.59	
	a_2	2500	0.71			
B	b_1	50	0.02	0.48	0.96	
	b_2	3000	0.98			
C	c_1	500	0.50	0	0.50	
	c_2	500	0.50			
D	d_1	500	0.05	0	0.05	
	:	:	:			
	d_{20}	500	0.05			

Table 6.1: Example calculation of the Gini coefficient and the Herfindahl–Hirschman index (HHI) for four countries trading with several partners each. HHI is able to capture differences in trading volumes as well as different number of partners.

Risk Score

To address the limitations of the Gini coefficient regarding the number of trading partners, I developed an extended risk measure (referred to here as *Risk Score*). This measure integrates both inequality (via the Gini coefficient) and the number of trading partners into a unified risk metric. Mathematically, the Risk Score R_c for country c is defined as:

$$R_c = (G + \alpha \cdot (1 - G))^{\beta \cdot \frac{n}{N}}$$

where:

- G is the Gini coefficient (Equation 6.1),
- n is the number of trading partners of country c ,
- N is the total possible number of trading partners (e.g., all countries in the dataset),
- α and β are tunable parameters.

Conceptually, α adjusts the influence of the Gini coefficient within the risk evaluation, ranging from 0 (pure Gini) to 1 (maximally influenced by the partner count). The parameter β controls how sharply the risk decreases with an increasing proportion of partners (i.e., $\frac{n}{N}$), essentially dictating the sensitivity of risk reduction relative to the diversification of trading partners.

Figure 6.2 illustrates the effect of varying α and β values on the Risk Score, demonstrating how the metric adjusts with differing degrees of sensitivity to diversification.

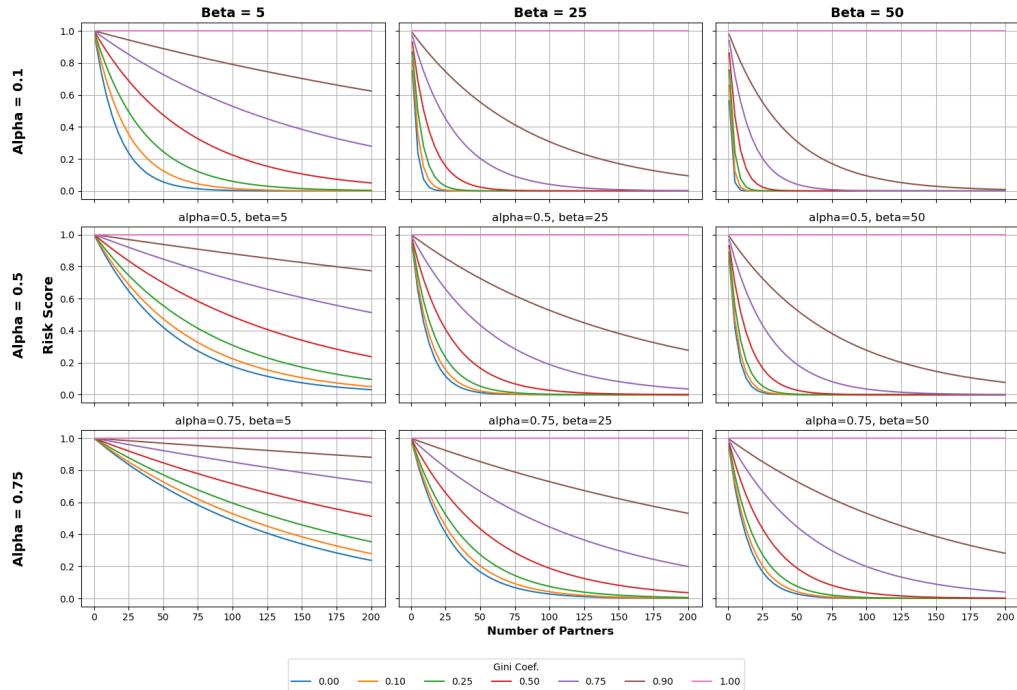


Figure 6.2: Risk Score and the effect of α (rows) and β (columns) in the score value, for varying degrees of Gini coefficients (coloured lines).

Upon comparison with HHI (Figure 6.3), the Risk Score exhibited a high positive non-linear correlation (Spearman correlation = 0.965, p -value: 0.000). Despite the promising alignment, HHI was ultimately preferred due to its intuitive interpretability, straightforward calculation, and established acceptance in risk management literature [45].

6.2.6 Alternative Discrepancy Assessments

In this study, I utilize discrepancies in countries' reporting as a proxy for data reliability, quantified through a proposed *Trustworthiness Index* (Section 4.2.4). Other studies have similarly assessed discrepancies in trade data reported to international databases like UN Comtrade. Notably, the work of [35] introduces a Discrepancy Index (DI) to evaluate the quality of trade reporting. This index measures divergence between

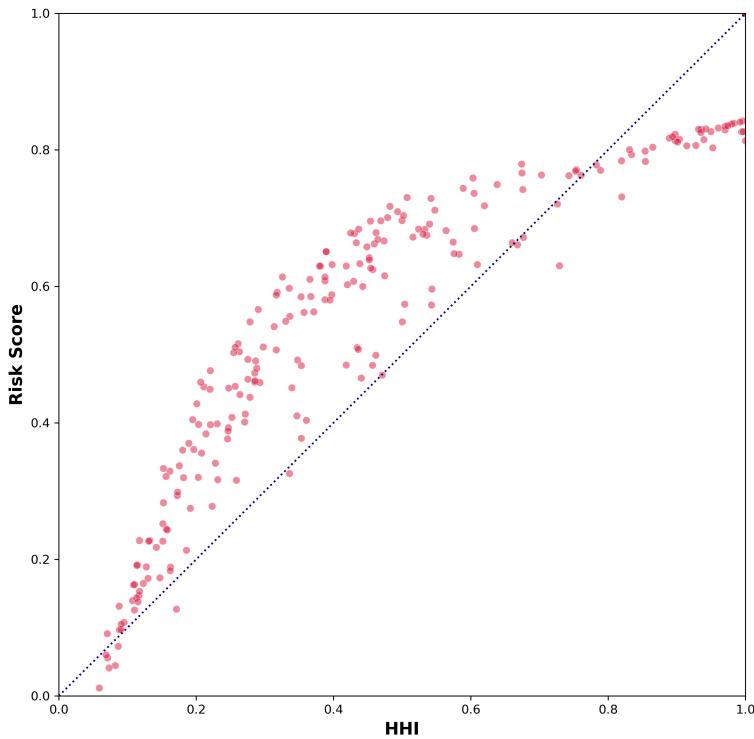


Figure 6.3: Comparison between Risk Score ($\alpha = 0.25, \beta = 50$) and HHI for one commodity layer ($l = 71$). Measures exhibit a strong non-linear correlation.

a country's reported trade values and the corresponding mirror data (i.e., trade values reported by the partner country).

The DI ranges between -1 and 1, where:

- A value of -1 indicates an orphan import (i.e., the reported import lacks any corresponding mirror export report),
- A value of 1 represents a lost export (i.e., the reported export lacks any corresponding mirror import report),
- A value of 0 indicates perfect alignment between a country's reports and its mirror data, implying no discrepancy.

Values below 0 indicate over-reported imports, whereas values above 0 indicate under-reported imports. To facilitate comparison between this Discrepancy Index and the Trustworthiness Index computed in my analysis, I take the absolute value of the DI:

$$DI' = |DI|$$

A Spearman correlation analysis comparing the Trustworthiness Index derived from [50] and the absolute DI from [35] (aggregating data across all available countries, years, and commodities³) yields a correlation of -0.612 (p -value = 0.000). Given the inverse relationship inherent between these two measures (i.e., higher trustworthiness implies lower discrepancy), this correlation underscores the strong consistency and interrelatedness between the two metrics.

A key limitation of the DI, however, is that it is computed equally across every country-pair relationship, thus assigning equal weight to each relationship regardless of trade volume. In contrast, the Trustworthiness Index proposed in this work incorporates trade flow volumes directly into its calculation, thereby weighting discrepancies by their economic significance.

Trustworthiness, Discrepancy, and Corruption

[35] suggest that misreporting in trade data can often be attributed to corruption, notably practices like tax evasion. Supporting this assertion, [68] confirmed previous findings demonstrating that tariffs contribute significantly to import under-reporting and that corruption strongly influences discrepancies for both importers and exporters.

Given this context, I examined the correlation between the computed Trustworthiness Index and the Corruption Perceptions Index (CPI), published by Transparency International (Section 3.4). The CPI, standardized on a scale from 0 to 100, aggregates data from 13 institutions to reflect perceived corruption levels, with 0 representing the highest corruption and 100 representing the lowest (see Appendix D.3 for the complete list). Intuitively, higher levels of corruption (lower CPI scores) are expected to correlate with lower trustworthiness scores (indicating higher discrepancies).

Indeed, a Spearman correlation analysis between these two variables resulted in a statistically significant correlation of 0.57 (p -value = $1.33e^{-88}$). This result strongly supports the hypothesis that increased perceived corruption is significantly associated with increased reporting discrepancies, affirming the validity of using discrepancies in trade data as an indirect indicator of corruption.

³DI data is unavailable for years prior to 2017.

6.2.7 Node Labels

The central objective of this research involves identifying countries that could potentially be vulnerable due to diminished exports from their trading partners (Section 4.3). The underlying logic is straightforward: the more significantly an economy is impacted by disruptions in another country's supply, the more vulnerable that economy is considered. However, operationalizing this logic into explicit criteria for labelling countries as either vulnerable or resilient involves several methodological decisions, each subject to scrutiny.

Firstly, identifying when a country qualifies as a "*significant exporter*" relies heavily on the concept of Revealed Comparative Advantage (RCA). In this study, following the reasoning in [36], a country is designated an exporter of a specific commodity if it exhibits an above 1 RCA, relative to global exports. Under this approach, a country exporting a relatively small volume (e.g., 100 units) could be classified as an exporter if these exports represent a significant share of its total trade, while another country exporting substantially larger absolute quantities (e.g., 10,000 units) might not qualify due to its diverse trade portfolio. Although justified by existing literature, this approach may not fully capture nuances in actual export volumes, suggesting room for complementary criteria (e.g., absolute thresholds or hybrid indicators).

Additionally, I define "loss of exporter status" as a drop in RCA of at least 20%. This threshold is informed by prior research into supply disruptions [14]. Nevertheless, this choice is not without its limitations. The threshold of 20% is somewhat arbitrary and may not universally apply across all commodities or country contexts. Future research could benefit from systematically exploring a broader spectrum of thresholds, potentially improving robustness by conducting sensitivity analyses or scenario-based investigations.

The next critical step involves determining whether an importer is genuinely affected by the diminished supply from a particular exporter. Here, I adopt two primary criteria. Firstly, an importer is considered affected if it experiences a significant reduction in goods imported from the impacted exporter, matching the 20% RCA reduction threshold. Secondly, if an importer replaces its supplier but incurs higher costs, it is also classified as affected. Conversely, if the importer can seamlessly substitute the affected exporter without additional costs or significant disruptions, it is not classified as vulnerable. This scenario commonly

applies to commodities that are ubiquitous or easily substitutable in international markets.

However, this approach implicitly neglects certain complex market behaviours. For instance, consider a scenario in which an importing country reduces its total import of a specific commodity, such as steel, following a drop in exports from a primary supplier. According to the current criteria, the importer would automatically be classified as affected. Yet, this reduction might result from internal economic shifts, changes in domestic demand, or structural economic transformation rather than external supply disruptions. Consequently, the analysis could mistakenly attribute vulnerability to external causes when, in reality, internal economic factors might be responsible.

6.2.8 Inter-layer relationships

International trade involves complex interdependencies among various commodities. It is easy to recognize that many goods directly influence the production of others. To cite two examples, raw materials frequently serve as crucial inputs for manufacturing more complex products, while specialized machinery and equipment are essential components in the production of various consumer and industrial goods. When analysing commodities independently, these intricate interdependencies might be overlooked, potentially leading to incomplete assessments of how supply disruptions propagate through the network.

In reality, a disruption or shortage in the supply of one particular commodity can trigger cascading effects across multiple product classes. [46] underscored this complexity by employing Granger causality analyses to identify commodity groups exhibiting significant interconnectedness within the ITN. Their findings highlight that changes in the availability or production of certain commodities can have widespread, cascading impacts on the availability, pricing, and production of related goods. Such interdependencies suggest that disruptions in key raw materials or intermediate products could rapidly propagate through the global economy, affecting industries seemingly unrelated at first glance.

Inspired by these findings, I incorporated two multi-edge graph variations in this work to better capture the layered and interconnected nature of trade relationships. However, fully understanding and modelling these inter-layer dependencies require more sophisticated and targeted approaches than those employed here. Future research could explore

advanced methodologies—such as dynamic multilayer networks, input-output modelling, or the integration of supply-chain analysis—to more comprehensively capture and quantify the intricate commodity interdependencies. Such endeavours would substantially deepen our understanding of systemic vulnerabilities in global trade networks, potentially providing more robust tools for predicting and mitigating the cascading consequences of trade disruptions.

6.3 Results Analysis

In this research, I evaluated three GNN architectures extensively adopted in prior studies—Graph Convolutional Networks (GCN), Graph Attention Networks (GATv2), and GraphSAGE. Observed performance differences among these architectures can be attributed to their distinct structural properties and how they process network information.

Graph Convolutional Networks (GCN). GCN represents the simplest architecture employed, relying directly on the graph adjacency matrix to compute convolutions. Therefore, in this architecture only edge weights (i.e., transaction volume) are considered in the convolution, and all other edge attributes are ignored. Leveraging network connectivity and trade volumes, already provides increased predictive power over purely attribute-driven models such as the MLP. In this work, the GCN consistently employed two-layer structures, enabling information aggregation from nodes located up to two hops away. This ensures the GCN doesn't fall victim of the *over-smoothing* problem, common to this type of architectures. Interestingly, results indicated that GCN performance deteriorated in directed graphs compared to undirected configurations. This decline likely occurs because convolution operations in directed graphs propagate information asymmetrically (from exporter to importer), thus limiting the bidirectional informational context critical for effectively identifying vulnerabilities. Conversely, GCN showed robust performance even with increasing graph density, as observed in multi-graph scenarios, suggesting resilience in handling complex connectivity patterns.

GraphSAGE. GraphSAGE results are really interesting, as this architecture does not utilize edge attributes or even edge weights. These re-

sults suggest that model effectiveness may not be strongly dependent on transaction volumes, an outcome that is, at first glance, counter-intuitive. Moreover, the model exhibited consistently reliable performance across different configurations, which can be attributed to its distinctive architectural design, including root–neighbour separation and the max-based aggregator. These characteristics enable GraphSAGE to effectively leverage additional node-specific information (such as commodity-layer embeddings) without relying on densely parametrized attention mechanisms. A key distinction between GraphSAGE and GCN lies in the former’s use of concatenation: after aggregating neighbourhood information, GraphSAGE appends the root node’s own representation, potentially preserving more of its intrinsic feature identity.

Graph Attention Networks (GATv2). Among the architectures examined in this study, GATv2 is the most complex, introducing attention mechanisms that dynamically assign weights to node relationships based on learned relevance. Despite this sophistication, GATv2 often could not surpass GraphSAGE. Its performance was particularly compromised in multi-graph settings characterized by exceptionally high node degrees. A plausible explanation lies in the behaviour of the softmax function used to compute attention coefficients: when a node is connected to a large number of neighbours, these coefficients tend to become excessively diffuse. As a result, the gradients used for parameter updates weaken, leading the attention mechanism to converge toward near-uniform weighting. This collapse affects the selective advantage of attention, rendering the architecture less effective than simpler alternatives such as GCN or GraphSAGE.

Key Insights and Practical Implications. Several important insights emerge from the comparative analysis of model performance:

1. **Topology matters:** In almost all cases, including topology information to the task resulted in better predictions.
2. **More attributes are not always needed:** Increasing the number of node or edge attributes (e.g., the addition of layer embeddings) does not always enhance model performance. In fact, when the network becomes highly dense, supplementary attributes can be counterproductive. This underscores the importance of conducting systematic ablation studies to identify and retain only the most

informative features, thereby maintaining a compact and efficient feature space.

3. **Hybrid and ensemble approaches perform best overall:** As expected, ensembles that combine multiple model architectures offer a more balanced and robust predictive framework. In this study, the ensemble approach led to notable gains in both generalizability and predictive stability. The downside is obviously added complexity by training and maintaining several models.
4. **Specific models might not be so important:** The specific choice of GNN architecture appears to be less consequential than factors such as graph representation and feature selection. Additionally, hyperparameter tuning revealed minimal variability in performance, suggesting that architectural differences may not be the primary driver of outcome disparities.

The results from the previous chapter (Section 5) show that, although GNNs bring significant improvements to the task resolution, this performance is still far from satisfactory. Many reasons could explain this lack of excellent performance results, starting with admitting the difficult task, the highly demanding metric (**F₁-score** on the positive class), lack of adequate data, etc. I summarize some of these next.

A difficult task. Predicting which countries will be affected, following an arbitrary classification, can deem the labelling task very difficult. Specifically, the labelling task is sensitive to small differences in thresholds (e.g., distinguishing between countries labelled as *affected importer* with a 21% decrease in import volume versus those not labelled as such, experiencing a 19% decrease). Additionally, countries may undergo changes in their trade patterns driven by external factors that are not directly reflected in the available data or network structure. Subtle shifts such as changes in political parties, or policy changes, can significantly alter trade relationships without immediately noticeable signals within standard economic data, making it difficult for models to capture such nuances effectively⁴.

⁴See the example discussed in Section 1: https://www.youtube.com/watch?v=q1qEaCQa6_w (Accessed: May 20, 2025).

A difficult metric. Many previous work using GNNs [63][53][54][28] have assessed model performance for classification tasks mostly using Accuracy. Although a good metric for summarizing classification, I opted here for a more complicated metric that better captures the goal of the task: to correctly (recall) and safely (precision) determine affected countries. More specifically, instead of using a Macro Averaged **F₁-score** (the harmonic mean of precision and recall), I focus more specifically in the **F₁-score** of the positive label. Since the positive cases are a minority in this work, the sensibility of this metric is much higher (i.e., each mistake made in a prediction of a positive class, either reducing recall or precision, can cause a big drop in **F₁-score**).

A difficult causal connection. Finding the best attributes that can help explain the relationship between diminishing exporters and affected importers is far from trivial. In this work, I have indirectly settled the models to determine which countries are exporters (remember I pass SRCA [Section 4.2.4] but I do not explicitly mark exporters), then identify which countries are potentially going to fail to continue as exporters, and then which countries connected to these exporters are going to suffer from this change in the exporter. A visual depiction of this two parts process can be seen in Figure 6.4. This complexity makes it quite challenging for a model to perform the classification task without any further assistance.

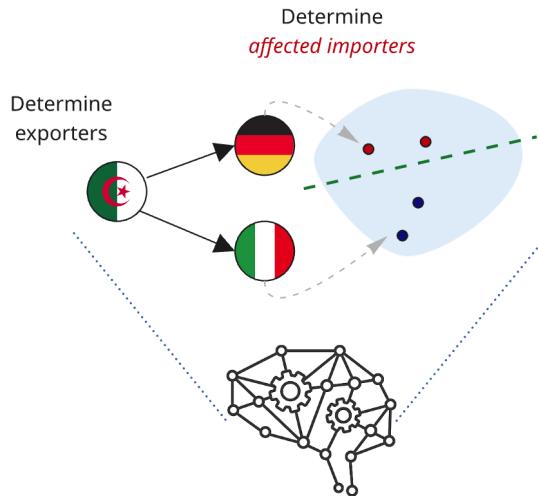


Figure 6.4: A model needs to perform a two-fold process to identify vulnerable nodes from neighbouring nodes that are deemed as ‘exporters’.

6.4 From Node-Level to Network-Level Vulnerability

Throughout this research, the primary focus has been on node-level classification, specifically identifying individual countries within the ITN that are vulnerable to supply disruptions. While understanding vulnerabilities at the individual country level provides valuable insights, another critical dimension involves assessing the vulnerability of the entire trade network itself.

One straightforward approach explored in this study involves analysing the proportion of vulnerable countries across different industries and time periods. This measure offers an initial indication of how various industries respond to disruptions and provides a temporal perspective on global vulnerability trends. However, this simplistic aggregation of node-level vulnerabilities may overlook deeper structural attributes and resilience characteristics inherent within the network as a whole.

Future research could beneficially extend the scope of this analysis by explicitly addressing the vulnerability and robustness of the entire network using graph-level metrics. GNNs naturally lend themselves to this endeavour, as they can effectively summarize and embed the complete network structure into holistic graph-level representations. Shifting the task from node classification to graph classification would allow the explicit modelling of entire networks based on structural properties indicative of resilience or susceptibility to disruptions.

Graph classification refers explicitly to the problem of classifying entire graphs—rather than individual nodes—based on structural characteristics or aggregated node attributes. In this scenario, networks would be represented by comprehensive embeddings designed to capture salient structural features relevant to their overall vulnerability. By creating such graph embeddings, networks can become linearly separable based on their vulnerability characteristics, thus facilitating a straightforward classification task (Figure 6.5).

Employing this graph-level approach could enhance our understanding of global trade vulnerability from a systemic perspective, and also provide policymakers and stakeholders extra tools for assessing the resilience of international trade. Moreover, exploring graph classification approaches could illuminate the relationship between global network properties, such as modularity, centrality, or density—and vulnerability to shocks, thereby guiding strategies aimed at enhancing systemic resilience.

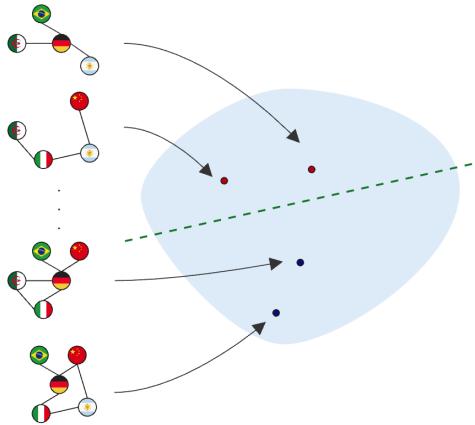


Figure 6.5: Visual depiction of the graph classification task in GNNs. Graphs get embedded in high-dimensional space, where they can become linearly separable.

6.5 Uncertainty

When analysing the results for this type of task, it is very important to include some measure of uncertainty, allowing stakeholders to understand the degree of confidence or potential variability in the findings. The quantification of model confidence in deep learning is a highly complex topic that goes beyond the scope of this work. Nevertheless, I would like to include a few remarks on this topic.

6.5.1 Aleatoric and Epistemic Uncertainty

[69] distinguish two types of uncertainty relevant for machine learning models: aleatoric and epistemic. Roughly speaking, aleatoric (aka statistical) uncertainty refers to the notion of randomness, that is, the variability in the outcome of an experiment which is due to inherently random effects. In this work, we encounter aleatoric uncertainty every time the model fails to predict the correct status of a country, due to a stochastic component that cannot be reduced by any additional source of information. In some way, this irreducible uncertainty marks the cap or ceiling of the provided GNNs in capturing the phenomenon. On the other hand, epistemic (aka systematic) uncertainty refers to uncertainty caused by a lack of knowledge (about the best model). In other words, it refers to the ignorance of the agent or decision maker, and hence to the epistemic state of the agent instead of any underlying random phenomenon. As opposed to uncertainty caused by randomness,

uncertainty caused by ignorance can in principle be reduced on the basis of additional information (e.g., more and better country attributes, or more expanded timeframes). Notice that, in this work, I am unable to differentiate or separate this two types of uncertainty in the models' results, and therefore it is not clear how much better this performance could be in presence of more and better data.

6.5.2 Uncertainty for Decision Making

Providing explicit measures of uncertainty alongside model predictions can significantly assist stakeholders in making informed decisions. One practical approach is to communicate probabilistic outputs rather than binary classifications alone. For instance, instead of merely predicting whether a country will be affected next year, the model could return an associated probability, thereby conveying a level of confidence in each prediction. This approach can be extended to the ensemble framework, where instead of using simple majority voting, aggregating probabilities across multiple models can yield more robust and calibrated predictions.

Another valuable strategy to manage uncertainty is to implement confidence thresholds. In scenarios where the model's output probability falls below a specified level, it may deliberately abstain from issuing a decision. This approach, known as selective prediction, can enhance overall performance by reducing erroneous classifications. However, this advantage must be weighed against the reduced expressiveness of the model (i.e., its reluctance to produce outputs in borderline cases), which might lead to missed opportunities for valuable insights. Therefore, threshold selection and trade-off calibration should be carefully adapted to the specific application, balancing the cost of incorrect outcomes against the benefits of broader coverage.

6.6 Explainability

In Section 1, I emphasize the critical need for improved understanding of systemic vulnerabilities within the ITN, and how GNNs can serve as valuable tools for decision-makers, policymakers, and other stakeholders. Besides uncertainty, discussed in the previous section, another essential factor in the practical deployment of these models is explainability.

Explainability, particularly within Deep Learning and broadly within Machine Learning, is a complex and actively researched domain, extensively studied yet continuously evolving. Although a thorough exploration of this field is beyond the scope of this project, its significance cannot be overstated. As highlighted by [70], explanations play a vital role in interpreting the behaviour and predictions of GNNs. Evaluating the quality and reliability of these explanations is crucial, as they directly influence stakeholder trust and decision-making effectiveness.

Thus, future research should prioritize integrating explainability into GNN-based predictions. Providing meaningful explanations alongside predictions would not only enhance stakeholder trust but also offer practical insights into mitigating identified vulnerabilities. Incorporating established explainability frameworks, such as those proposed by [71], could substantially improve the interpretability and practical utility of the methods developed in this thesis, enabling informed, actionable decisions and ultimately contributing to enhanced economic resilience.

Chapter 7

Conclusions

“Knowing is not enough; we must apply. Willing is not enough; we must do.”

Johann Wolfgang von Goethe

This research explores the feasibility and effectiveness of using Graph Neural Networks (GNNs) to assess the vulnerabilities of individual countries within the context of international trade networks. By leveraging publicly available trade data, this work examines how network topology and node-specific characteristics can be integrated into deep learning models to enhance our understanding of systemic vulnerabilities across various industries and commodities.

The results clearly demonstrate the benefits of incorporating topological attributes into trade vulnerability analyses, moving beyond traditional bilateral approaches. Specifically, the GNN architectures evaluated here —Graph Convolutional Networks (GCN), Graph Attention Networks (GATv2), and GraphSAGE— each exhibit unique strengths and limitations dependent on the network’s density and structural characteristics. Nevertheless, differences in architectural complexity do not consistently emerge as the primary drivers of performance. In fact, architectures with greater complexity and a higher number of parameters, such as GATv2, do not always translate directly into improved performance in every scenario.

A significant finding of this research is the gained performance in using ensemble methods. Combining predictions from multiple GNN

architectures effectively balances bias and variance, leading to enhanced predictive accuracy and robustness across diverse analytical contexts. These results underscore the value of hybrid modelling strategies for practical applications, particularly when dealing with complex, real-world networks such as the International Trade Network (ITN). Given the relatively moderate size and complexity of the ITN graphs, employing ensemble approaches remains computationally feasible, enabling the integration of multiple models within reasonable computational constraints.

Throughout this study, I proposed and evaluated several innovative representations of the ITN, including both single-layer and multi-layer graph structures. These approaches allowed for capturing different facets of international trade interactions, highlighting how specific design choices can influence analytical outcomes. A notable methodological contribution was the introduction of layer embeddings, which combined Natural Language Processing (NLP) techniques with economic complexity metrics. This approach provided nuanced commodity-specific contextualization, enriching node-level features, although not resulting in performance gains.

Despite these advancements, the research highlighted several critical challenges. Feature engineering proved particularly crucial; accurately characterizing countries and commodities required careful consideration of both economic theory and empirical validity. Additionally, limitations inherent to publicly available trade data, such as reporting inaccuracies and classification ambiguities, necessitate cautious interpretation of the findings and underscore the importance of improving data quality.

Moving forward, several promising avenues exist for future exploration. Comprehensive ablation studies should be conducted to systematically identify optimal feature sets and refine node labelling strategies. Incorporating deeper domain expertise in economics and international trade could greatly enhance the quality and relevance of both feature construction and labelling methodologies. Furthermore, longitudinal analyses covering extended temporal periods could offer richer insights into how global trade vulnerabilities evolve over time. Investigating cascading effects among interconnected commodities and industries could also significantly improve our understanding of disruption propagation within complex economic systems. Lastly, developing more sophisticated methods to aggregate individual node-level results into a compre-

hensive and meaningful metric of overall network vulnerability remains an essential area for further research.

Ultimately, this research provides meaningful insights and methodological contributions toward understanding trade vulnerabilities through the lens of graph neural networks. By highlighting critical methodological considerations and areas for improvement, this work lays essential groundwork for more accurate vulnerability assessments. Furthermore, it establishes a robust foundation for future efforts aimed at enhancing the resilience and robustness of global trade networks.

Appendix A

UN Comtrade Data

A.1 HS2012 Commodity Nomenclature

id	text
01	Live animals
02	Meat and edible meat offal
03	Fish and crustaceans, molluscs and other aquatic invertebrates
04	Dairy produce; birds' eggs; natural honey; edible products of animal origin, not elsewhere specified or included
05	Products of animal origin, not elsewhere specified or included
06	Live trees and other plants; bulbs, roots and the like; cut flowers and ornamental foliage
07	Edible vegetables and certain roots and tubers
08	Edible fruit and nuts; peel of citrus fruit or melons
09	Coffee, tea, mat and spices
10	Cereals
11	Products of the milling industry; malt; starches; inulin; wheat gluten
12	Oil seeds and oleaginous fruits; miscellaneous grains, seeds and fruit; industrial or medicinal plants; straw and fodder
13	Lac; gums, resins and other vegetable saps and extracts
14	Vegetable plaiting materials; vegetable products not elsewhere specified or included
15	Animal or vegetable fats and oils and their cleavage products; prepared edible fats; animal or vegetable waxes
16	Preparations of meat, of fish or of crustaceans, molluscs or other aquatic invertebrates
17	Sugars and sugar confectionery
18	Cocoa and cocoa preparations
19	Preparations of cereals, flour, starch or milk; pastrycooks' products
20	Preparations of vegetables, fruit, nuts or other parts of plants
21	Miscellaneous edible preparations
22	Beverages, spirits and vinegar
23	Residues and waste from the food industries; prepared animal fodder
24	Tobacco and manufactured tobacco substitutes
25	Salt; sulphur; earths and stone; plastering materials, lime and cement
26	Ores, slag and ash
27	Mineral fuels, mineral oils and products of their distillation; bituminous substances;

- mineral waxes
- 28 Inorganic chemicals; organic or inorganic compounds of precious metals, of rare-earth metals, of radioactive elements or of isotopes
- 29 Organic chemicals
- 30 Pharmaceutical products
- 31 Fertilisers
- 32 Tanning or dyeing extracts; tannins and their derivatives; dyes, pigments and other colouring matter; paints and varnishes; putty and other mastics; inks
- 33 Essential oils and resinoids; perfumery, cosmetic or toilet preparations
- 34 Soap, organic surface-active agents, washing preparations, lubricating preparations, artificial waxes, prepared waxes, polishing or scouring preparations, candles and similar articles, modelling pastes, "dental waxes" and dental preparations with a basis of plaster
- 35 Albuminoidal substances; modified starches; glues; enzymes
- 36 Explosives; pyrotechnic products; matches; pyrophoric alloys; certain combustible preparations
- 37 Photographic or cinematographic goods
- 38 Miscellaneous chemical products
- 39 Plastics and articles thereof
- 40 Rubber and articles thereof
- 41 Raw hides and skins (other than fur skins) and leather
- 42 Articles of leather; saddlery and harness; travel goods, handbags and similar containers; articles of animal gut (other than silk-worm gut)
- 43 Fur skins and artificial fur; manufactures thereof
- 44 Wood and articles of wood; wood charcoal
- 45 Cork and articles of cork
- 46 Manufactures of straw, of esparto or of other plaiting materials; basketware and wickerwork
- 47 Pulp of wood or of other fibrous cellulosic material; recovered (waste and scrap) paper or paperboard
- 48 Paper and paperboard; articles of paper pulp, of paper or of paperboard
- 49 Printed books, newspapers, pictures and other products of the printing industry; manuscripts, typescripts and plans
- 50 Silk
- 51 Wool, fine or coarse animal hair; horsehair yarn and woven fabric
- 52 Cotton
- 53 Other vegetable textile fibres; paper yarn and woven fabrics of paper yarn
- 54 Man-made filaments; strip and the like of man-made textile materials
- 55 Man-made staple fibres
- 56 Wadding, felt and nonwovens; special yarns; twine, cordage, ropes and cables and articles thereof
- 57 Carpets and other textile floor coverings
- 58 Special woven fabrics; tufted textile fabrics; lace; tapestries; trimmings; embroidery
- 59 Impregnated, coated, covered or laminated textile fabrics; textile articles of a kind suitable for industrial use
- 60 Knitted or crocheted fabrics
- 61 Articles of apparel and clothing accessories, knitted or crocheted
- 62 Articles of apparel and clothing accessories, not knitted or crocheted
- 63 Other made up textile articles; sets; worn clothing and worn textile articles; rags
- 64 Footwear, gaiters and the like; parts of such articles
- 65 Headgear and parts thereof
- 66 Umbrellas, sun umbrellas, walking-sticks, seat-sticks, whips, riding-crops and parts thereof
- 67 Prepared feathers and down and articles made of feathers or of down; artificial flowers; articles of human hair
- 68 Articles of stone, plaster, cement, asbestos, mica or similar materials

- 69 Ceramic products
- 70 Glass and glassware
- 71 Natural or cultured pearls, precious or semi-precious stones, precious metals, metals clad with precious metal, and articles thereof; imitation jewellery; coin
- 72 Iron and steel
- 73 Articles of iron or steel
- 74 Copper and articles thereof
- 75 Nickel and articles thereof
- 76 Aluminium and articles thereof
- 77 (Reserved for possible future use in the Harmonized System)
- 78 Lead and articles thereof
- 79 Zinc and articles thereof
- 80 Tin and articles thereof
- 81 Other base metals; cermets; articles thereof
- 82 Tools, implements, cutlery, spoons and forks, of base metal; parts thereof of base metal
- 83 Miscellaneous articles of base metal
- 84 Nuclear reactors, boilers, machinery and mechanical appliances; parts thereof
- 85 Electrical machinery and equipment and parts thereof; sound recorders and reproducers, television image and sound recorders and reproducers, and parts and accessories of such articles
- 86 Railway or tramway locomotives, rolling-stock and parts thereof; railway or tramway track fixtures and fittings and parts thereof; mechanical (including electro-mechanical) traffic signalling equipment of all kinds
- 87 Vehicles other than railway or tramway rolling-stock, and parts and accessories thereof
- 88 Aircraft, spacecraft, and parts thereof
- 89 Ships, boats and floating structures
- 90 Optical, photographic, cinematographic, measuring, checking, precision, medical or surgical instruments and apparatus; parts and accessories thereof
- 91 Clocks and watches and parts thereof
- 92 Musical instruments; parts and accessories of such articles
- 93 Arms and ammunition; parts and accessories thereof
- 94 Furniture; bedding, mattresses, mattress supports, cushions and similar stuffed furnishings; lamps and lighting fittings, not elsewhere specified or included; illuminated signs, illuminated name-plates and the like; prefabricated buildings
- 95 Toys, games and sports requisites; parts and accessories thereof
- 96 Miscellaneous manufactured articles
- 97 Works of art, collectors' pieces and antiques
- 99 Commodities not specified according to kind

A.2 Data tables: Reporters

The UN Comtrade provides a list of the countries that report data to the UN. The full table contains 255 rows and 6 columns (some omitted here, since they contain redundant information). The table provides a way to map the `reporterCode` to countries ISO2 (`reporterCodeIsoAlpha2`) and ISO3 (`reporterCodeIsoAlpha3`) codes. The full list can be seen in Table A.2.

reporterCode	reporterDesc	reporterCodeIsoAlpha2	reporterCodeIsoAlpha3
4	Afghanistan	AF	AFG
8	Albania	AL	ALB
12	Algeria	DZ	DZA
20	Andorra	AD	AND
24	Angola	AO	AGO
660	Anguilla	AI	AIA
28	Antigua and Barbuda	AG	ATG
886	Arab Rep. of Yemen (...1990)	YE	YEM
32	Argentina	AR	ARG
51	Armenia	AM	ARM
533	Aruba	AW	ABW
975	ASEAN	R4	R4
36	Australia	AU	AUS
40	Austria	AT	AUT
31	Azerbaijan	AZ	AZE
44	Bahamas	BS	BHS
48	Bahrain	BH	BHR
50	Bangladesh	BD	BGD
52	Barbados	BB	BRB
112	Belarus	BY	BLR
56	Belgium	BE	BEL
58	Belgium-Luxembourg (...1998)	BE	BEL
84	Belize	BZ	BLZ
204	Benin	BJ	BEN
60	Bermuda	BM	BMU
64	Bhutan	BT	BTN
68	Bolivia (Plurinational State of)	BO	BOL
535	Bonaire	BQ	BES
70	Bosnia Herzegovina	BA	BIH
72	Botswana	BW	BWA
92	Br. Virgin Isds	VG	VGB
76	Brazil	BR	BRA
96	Brunei Darussalam	BN	BRN
100	Bulgaria	BG	BGR
854	Burkina Faso	BF	BFA
108	Burundi	BI	BDI
132	Cabo Verde	CV	CPV
116	Cambodia	KH	KHM
120	Cameroon	CM	CMR
124	Canada	CA	CAN
136	Cayman Isds	KY	CYM
140	Central African Rep.	CF	CAF
148	Chad	TD	TCD
152	Chile	CL	CHL
156	China	CN	CHN
344	China, Hong Kong SAR	HK	HKG
446	China, Macao SAR	MO	MAC
170	Colombia	CO	COL
174	Comoros	KM	COM
178	Congo	CG	COG
184	Cook Isds	CK	COK
188	Costa Rica	CR	CRI
384	Cote d'Ivoire	CI	CIV
191	Croatia	HR	HRV
192	Cuba	CU	CUB

Continued on next page

reporterCode	reporterDesc	reporterCodeIsoAlpha2	reporterCodeIsoAlpha3
531	Cura ao	CW	CUW
196	Cyprus	CY	CYP
203	Czechia	CZ	CZE
200	Czechoslovakia (...1992)	CS	CSK
408	Dem. People's Rep. of Korea	KP	PRK
278	Dem. Rep. of Germany (...1990)	DD	DDR
180	Dem. Rep. of the Congo	CD	COD
866	Dem. Rep. of Vietnam (...1974)	VD	VDR
720	Dem. Yemen (...1990)	YD	YMD
208	Denmark	DK	DNK
262	Djibouti	DJ	DJI
212	Dominica	DM	DMA
214	Dominican Rep.	DO	DOM
588	East and West Pakistan (...1971)	PK	PAK
218	Ecuador	EC	ECU
818	Egypt	EG	EGY
222	El Salvador	SV	SLV
226	Equatorial Guinea	GQ	GNQ
232	Eritrea	ER	ERI
233	Estonia	EE	EST
748	Eswatini	SZ	SWZ
231	Ethiopia	ET	ETH
230	Ethiopia (...1992)	ET	ETH
97	European Union	EU	EUR
234	Faeroe Isds	FO	FRO
280	Fed. Rep. of Germany (...1990)	DE	DEU
242	Fiji	FJ	FJI
246	Finland	FI	FIN
251	France	FR	FRA
254	French Guiana (Overseas France)	GF	GUF
258	French Polynesia	PF	PYF
583	FS Micronesia	FM	FSM
266	Gabon	GA	GAB
270	Gambia	GM	GMB
268	Georgia	GE	GEO
276	Germany	DE	DEU
288	Ghana	GH	GHA
292	Gibraltar	GI	GIB
300	Greece	GR	GRC
304	Greenland	GL	GRL
308	Grenada	GD	GRD
312	Guadeloupe (Overseas France)	GP	GLP
320	Guatemala	GT	GTM
324	Guinea	GN	GIN
624	Guinea-Bissau	GW	GNB
328	Guyana	GY	GUY
332	Haiti	HT	HTI
336	Holy See (Vatican City State)	VA	VAT
340	Honduras	HN	HND
348	Hungary	HU	HUN
352	Iceland	IS	ISL
699	India	IN	IND
356	India (...1974)	IN	IND
360	Indonesia	ID	IDN
364	Iran	IR	IRN

Continued on next page

reporterCode	reporterDesc	reporterCodeIsoAlpha2	reporterCodeIsoAlpha3
368	Iraq	IQ	IRQ
372	Ireland	IE	IRL
376	Israel	IL	ISR
380	Italy	IT	ITA
388	Jamaica	JM	JAM
392	Japan	JP	JPN
400	Jordan	JO	JOR
398	Kazakhstan	KZ	KAZ
404	Kenya	KE	KEN
296	Kiribati	KI	KIR
414	Kuwait	KW	KWT
417	Kyrgyzstan	KG	KGZ
418	Lao People's Dem. Rep.	LA	LAO
428	Latvia	LV	LVA
422	Lebanon	LB	LBN
426	Lesotho	LS	LSO
430	Liberia	LR	LBR
434	Libya	LY	LYB
440	Lithuania	LT	LTU
442	Luxembourg	LU	LUX
450	Madagascar	MG	MDG
454	Malawi	MW	MWI
458	Malaysia	MY	MYS
462	Maldives	MV	MDV
466	Mali	ML	MLI
470	Malta	MT	MLT
584	Marshall Isds	MH	MHL
474	Martinique (Overseas France)	MQ	MTQ
478	Mauritania	MR	MRT
480	Mauritius	MU	MUS
175	Mayotte (Overseas France)	YT	MYT
484	Mexico	MX	MEX
496	Mongolia	MN	MNG
499	Montenegro	ME	MNE
500	Montserrat	MS	MSR
504	Morocco	MA	MAR
508	Mozambique	MZ	MOZ
104	Myanmar	MM	MMR
580	N. Mariana Isds	MP	MNP
516	Namibia	NA	NAM
520	Nauru	NR	NRU
524	Nepal	NP	NPL
528	Netherlands	NL	NLD
530	Netherlands Antilles (...2010)	AN	ANT
532	Netherlands Antilles and Aruba (...1985)	AN	ANT
540	New Caledonia	NC	NCL
554	New Zealand	NZ	NZL
558	Nicaragua	NI	NIC
562	Niger	NE	NER
566	Nigeria	NG	NGA
570	Niue	NU	NIU
807	North Macedonia	MK	MKD
579	Norway	NO	NOR
512	Oman	OM	OMN
490	Other Asia, nes	-	S19

Continued on next page

reporterCode	reporterDesc	reporterCodeIsoAlpha2	reporterCodeIsoAlpha3
582	Pacific Isds (...1991)	PC	PCI
586	Pakistan	PK	PAK
585	Palau	PW	PLW
591	Panama	PA	PAN
590	Panama, excl.Canal Zone (...1977)	PA	PAN
592	Panama-Canal-Zone (...1977)	PZ	PCZ
598	Papua New Guinea	PG	PNG
600	Paraguay	PY	PRY
459	Peninsula Malaysia (...1963)	_PM	-
604	Peru	PE	PER
608	Philippines	PH	PHL
616	Poland	PL	POL
620	Portugal	PT	PRT
634	Qatar	QA	QAT
410	Rep. of Korea	KR	KOR
498	Rep. of Moldova	MD	MDA
868	Rep. of Vietnam (...1974)	VN	VNM
638	R union (Overseas France)	RE	REU
717	Rhodesia Nyas (...1964)	_RN	-
642	Romania	RO	ROU
643	Russian Federation	RU	RUS
646	Rwanda	RW	RWA
461	Sabah (...1963)	_SH	-
652	Saint Barth lemy	BL	BLM
654	Saint Helena	SH	SHN
659	Saint Kitts and Nevis	KN	KNA
658	Saint Kitts, Nevis and Anguilla (...1980)	KN	KNA
662	Saint Lucia	LC	LCA
534	Saint Maarten	SX	SXM
666	Saint Pierre and Miquelon	PM	SPM
670	Saint Vincent and the Grenadines	VC	VCT
882	Samoa	WS	WSM
674	San Marino	SM	SMR
678	Sao Tome and Principe	ST	STP
682	Saudi Arabia	SA	SAU
686	Senegal	SN	SEN
688	Serbia	RS	SRB
891	Serbia and Montenegro (...2005)	CS	SCG
690	Seychelles	SC	SYC
694	Sierra Leone	SL	SLE
702	Singapore	SG	SGP
703	Slovakia	SK	SVK
705	Slovenia	SI	SVN
90	Solomon Isds	SB	SLB
706	Somalia	SO	SOM
710	South Africa	ZA	ZAF
728	South Sudan	SS	SSD
711	Southern African Customs Union (...1999)	-	ZA1
724	Spain	ES	ESP
144	Sri Lanka	LK	LKA
275	State of Palestine	PS	PSE
729	Sudan	SD	SDN

Continued on next page

reporterCode	reporterDesc	reporterCodeIsoAlpha2	reporterCodeIsoAlpha3
736	Sudan (...2011)	SD	SDN
740	Suriname	SR	SUR
752	Sweden	SE	SWE
757	Switzerland	CH	CHE
760	Syria	SY	SYR
762	Tajikistan	TJ	TJK
835	Tanganyika (...1964)	_TK	-
764	Thailand	TH	THA
626	Timor-Leste	TL	TLS
768	Togo	TG	TGO
772	Tokelau	TK	TKL
776	Tonga	TO	TON
780	Trinidad and Tobago	TT	TTO
788	Tunisia	TN	TUN
792	Turkiye	TR	TUR
795	Turkmenistan	TM	TKM
796	Turks and Caicos Isds	TC	TCA
798	Tuvalu	TV	TUV
800	Uganda	UG	UGA
804	Ukraine	UA	UKR
784	United Arab Emirates	AE	ARE
826	United Kingdom	GB	GBR
834	United Rep. of Tanzania	TZ	TZA
858	Uruguay	UY	URY
850	US Virgin Isds (...1980)	VI	VIR
842	USA	US	USA
841	USA and Puerto Rico (...1980)	US	USA
810	USSR (...1990)	SU	SUN
860	Uzbekistan	UZ	UZB
548	Vanuatu	VU	VUT
862	Venezuela	VE	VEN
704	Viet Nam	VN	VNM
876	Wallis and Futuna Isds	WF	WLF
887	Yemen	YE	YEM
890	Yugoslavia (...1991)	YU	YUG
894	Zambia	ZM	ZMB
836	Zanzibar and Pemba Isd (...1964)	_ZP	-
716	Zimbabwe	ZW	ZWE

A.3 UN Comtrade API

The UN Comtrade website (<https://comtradeplus.un.org/>) provides free access to detailed global trade data, continuously collected and released to the public. The data can be accessed via any of the following methods:

1. The website user interface: with a limit of 5,000 records;
2. An API, in three forms:
 - Free API: does not require authorization, but provides very limited access to the data.
 - Premium Individual API (the one used in this work): requires authorization token. It is free, but requires registration via email. Provides full access to the data, although it caps the data to 100,000 rows.
 - Premium Institutional Pro API: requires a paid subscription. Provides full access to the data, and allows for bulk downloads programmatically.
3. Bulk downloads: Only accessible through the paid subscription, allows to download all the data in bulks.

A.4 UN Comtrade Tariff Line Raw Data

UN Comtrade offers two primary types of datasets: Trade Data and Tariff Line Data. The latter is particularly detailed, offering granularity closer to the original data reported by countries¹. Additionally, the platform provides a Bilateral Data Comparison tool, facilitating straightforward comparisons of trade flows between two specified countries.

Table A.3 shows a representative extract from a tariff line dataset. The example shows transaction data for a given period (e.g., 2016), between a reporter country (e.g., Algeria) and a partner country (e.g., Italy), specifically focusing on a particular commodity class (e.g., '03'). Given that multiple rows of data may be reported throughout a single year, aggregation is typically required to obtain annual summaries. In

¹See <https://comtradeapi.un.org/files/v1/app/wiki/MethodologyGuidelineforComtradePlus.pdf>² for detailed methodology.

total, the full extracted bilateral dataset utilized in this research comprises 7,666,554 rows and 13 columns.

Furthermore, Table A.4 summarizes the complete set of attributes available through UN Comtrade datasets, which vary depending on subscription type and data extraction methods. A checkmark (✓) indicates an attribute's presence in the dataset, while an X denotes its absence.

A.4.1 Tariff Line Raw Data

A.4. UN Comtrade Tariff Line Raw Data

103

type	freq	Code	period	reporter	reporter	reporter	flow	partner	partner	partner	classification	cnd	cif	fob	primary
Code	Code	Code	Code	ISO	Desc	Code	Code	ISO	Desc	ISO	Code	Code	value	value	Value
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03061100	0	3208.74	3208.74	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03061100	0	2368.81	2368.81	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03061100	0	4272.64	4272.64	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03028900	0	346.99	346.99	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03028100	0	893.97	893.97	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03028100	0	2411.28	2411.28	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03021900	0	382.13	382.13	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03062100	0	5805.40	5805.40	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03062100	0	7639.78	7639.78	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03062100	0	10843.81	10843.81	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03062100	0	14444.33	14444.33	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03062100	0	10655.48	10655.48	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03062100	0	5981.98	5981.98	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03062100	0	24222.88	24222.88	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03062100	0	5079.00	5079.00	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03062100	0	1455.84	1455.84	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03062200	0	120.93	120.93	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03062790	0	1601.31	1601.31	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03075100	0	17.11	17.11	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03076000	0	16649.12	16649.12	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03076000	0	5516.32	5516.32	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03076000	0	6071.47	6071.47	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03076000	0	10550.74	10550.74	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03076000	0	19117.66	19117.66	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03076000	0	2577.49	2577.49	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03076000	0	5869.65	5869.65	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03076000	0	10318.23	10318.23	
C	A	2016	12	Algeria	DZA	X	380	Italy	ITA	H4	03076000	0	21601.86	21601.86	

Table A.3: Some columns have been ignored to fit the table in page.

dataItem	description	[1]	[2]	[3]	[4]
datasetCode	Combination of several keys to identify specific datasetCode	✓	✓	✓	✓
typeCode	Product type: Goods or Services	✓	✓	✓	✓
freqCode	The time interval at which observations occur	✓	✓	✓	✓
refPeriodId	The period of time to which the measured observation is intended to refer	✓	✓	✓	✓
refYear	Year of observation	✓	✓	✓	✓
refMonth	Month of observation. For annual, it would be set to 52	✓	✓	✓	✓
period	Combination of year and month (for monthly), year for (annual)	✓	✓	✓	✓
reporterCode	The country or geographic area to which the measured statistical phenomenon relates	✓	✓	✓	✓
reporterISO	ISO 3 code of reporter	✓	✓	✓	✓
reporterDesc	Description of reporter	✓	✓	✓	✓
flowCategory	Simplified trade flow: exports or imports	✓	✓	✓	✓
flowCode	Trade flow or sub-flow (exports, re-exports, imports, re-imports, etc.)	✓	✓	✓	✓
flowDesc	Description of trade flows	✓	✓	✓	✓
partnerCode	The primary partner country or geographic area for the respective trade flow	✓	✓	✓	✓
partnerISO	ISO 3 code of 1st partner	✓	✓	✓	✓
partnerDesc	Description of 1st partner	✓	✓	✓	✓
partner2Code	A secondary partner country or geographic area for the respective trade flow	✓	✓	✓	✓
partner2ISO	ISO 3 code of 2nd partner	✓	✓	✓	✓
partner2Desc	Description of 2nd partner	✓	✓	✓	✓
classificationSearchCode	Additional product classification to assist data search	✓	✓	✓	✓
classificationCode	Indicates the product classification used and which version (HS, SITC)	✓	✓	✓	✓
isOriginalClassification	Flag to indicate whether the classification is reported by country or not	✓	✓	✓	✓
cmdCode	Product code in conjunction with classification code	✓	✓	✓	✓
cndDesc	Description of commodity / service category	✓	✓	✓	✓
aggrLevel	Hierarchical level of commodity / service category	✓	✓	✓	✓
isLeaf	Identification whether a product code has the most basic level (i.e., sub-heading for HS)	✓	✓	✓	✓
customsCode	Customs or statistical procedure	✓	✓	✓	✓
customsDesc	Description of customs procedure	✓	✓	✓	✓
mosCode	The mode of supply on delivery of services (trade in services only)	✓	✓	✓	✓
mosDesc	Description of mode of supply (trade in services only)	✓	✓	✓	✓
motCode	The mode of transport used when goods enter or leave the economic territory of a country	✓	✓	✓	✓
motDesc	Description of mode of transport	✓	✓	✓	✓
qtyUnitCode	Unit of primary quantity	✓	✓	✓	✓
qtyUnitAbbr	Abbreviation of primary quantity unit	✓	✓	✓	✓
qty	Value of primary quantity	✓	✓	✓	✓
isQtyEstimated	Flag whether primary quantity is estimated or not	✓	✓	✓	✓

altQtyUnitCode	Unit of secondary quantity
altQtyUnitAbbr	Abbreviation of secondary quantity unit
altQty	Value of secondary quantity
isAltQtyEstimated	Flag whether secondary quantity is estimated or not
netWgt	Net weight
isNetWgtEstimated	Flag whether net weight is estimated or not
grossWgt	Gross weight
isGrossWgtEstimated	Flag whether gross weight is estimated or not
cifValue	Trade values in CIF
fobValue	Trade values in FOB
primaryValue	Primary trade values (taken from CIF or FOB values)
legacyEstimationFlag	Legacy quantity estimation flag
isReported	Flag to indicate whether a record is reported by country
isAggregate	Flag to indicate whether a record is aggregated by UNSD

Table A.4: Variables present in UN Comtrade datasets. NOTES: [1] isInDataAPIServices; [2] isInDataFinal; [3] isInBulkFileServices; [4] isInBulkFileFinal.

A.5 Countries with no bilateral data

ID	Name	ISO Code	Reason
10	Antarctica	ATA	[1]
16	American Samoa	ASM	[1]
74	Bouvet Island	BVT	[1]
86	British Indian Ocean Territory	IOT	[1]
92	British Virgin Islands	VGB	[2]
148	Chad	TCD	[2]
158	Taiwan	TWN	[1]
162	Christmas Island	CXR	[1]
166	Cocos (Keeling) Islands	CCK	[1]
175	Mayotte	MYT	[2]
184	Cook Islands	COK	[2]
226	Equatorial Guinea	GNQ	[2]
232	Eritrea	ERI	[2]
234	Faroe Islands	FRO	[2]
238	Islas Malvinas	FLK	[1]
239	South Georgia and South Sandwich Islds.	SGS	[1]
260	French Southern and Antarctic Lands	ATF	[1]
292	Gibraltar	GIB	[2]
316	Guam	GUM	[1]
332	Haiti	HTI	[2]
334	Heard and McDonald Islands	HMD	[1]
336	Vatican City	VAT	[2]
408	North Korea	PRK	[2]
520	Nauru	NRU	[2]
534	St-Martin / St Maarten	SXM	[2]
535	Bonaire	BES	[2]
548	Vanuatu	VUT	[2]
570	Niue	NIU	[2]
574	Norfolk Island	NFK	[1]
580	Northern Mariana Islands	MNP	[2]
581	US Minor Outlying Islands	UMI	[1]
584	Marshall Islands	MHL	[2]
612	Pitcairn	PCN	[1]
652	Saint Barthelemy	BLM	[2]
654	Saint Helena, Ascension and Tristan da Cunha	SHN	[2]
660	Anguilla	AIA	[2]
666	Saint Pierre and Miquelon	SPM	[2]
674	San Marino	SMR	[2]
706	Somalia	SOM	[2]
728	South Sudan	SSD	[2]
732	Western Sahara	ESH	[1]
760	Syria	SYR	[2]
772	Tokelau	TKL	[2]
795	Turkmenistan	TKM	[2]
798	Tuvalu	TUV	[2]
876	Wallis and Futuna	WLF	[2]
999	Undeclared	ANS	[1]

Table A.5: Countries with missing bilateral. Reasons: [1] = Country not present in UN Comtrade dataset; [2] = No reported data.

Appendix B

Atlas of Economic Complexity dataset

Harvard provides free access to the data used in the Atlas of Economic Complexity through its Growth Lab¹. The datasets used in this work include:

- Country Trade by Partner and Product ², for years 2012 to 2023. Contains harmonized trade data for country-partner-product-year combinations. Appendix B.1 shows an extract of this table. The full table contains 59,709,816 rows and 9 columns.
- Location Country³. Provides a list of the countries present in the dataset, along with ISO3 codes and other attributes. Appendix B.2 shows the full table, containing 252 rows and 4 columns.
- Product HS12³. Provides a table with the products contained in the data, with a mapping from Product ID to HS Nomenclature. Appendix B.3 shows an extract of this table. The full table contains 6,540 rows and 6 columns.

¹Link: <https://atlas.hks.harvard.edu/data-downloads> (Last accessed: 12/05/2025).

²The Growth Lab at Harvard University, 2025, "International Trade Data (HS12)", <https://doi.org/10.7910/DVN/YAVJDF>, Harvard Dataverse

³The Growth Lab at Harvard University, 2025, "Classifications Data", <https://doi.org/10.7910/DVN/3BAL10>, Harvard Dataverse

B.1 Data tables: Trade Data

country_id	partner_country_id	year	product_id	export_value	import_value	coi	eci	pci
...								
12	380	2012	669	0	102,705	-1.076	-1.233	-1.567
12	380	2012	671	9555	0	-1.076	-1.233	-1.963
12	380	2012	668	0	61,733	-1.076	-1.233	-1.854
12	380	2012	667	282,065	0	-1.076	-1.233	-1.514
12	380	2012	672	179,216	0	-1.076	-1.233	-1.781
...								

Table B.1: Extract of the Country Trade by Partner and Product table from [43]. `country_id` and `partner_country_id` can be matched to `country_id` in Table B.2, while `product_id` can be matched to `product_id` in Table B.3. Trade discrepancies are already smoothed.

B.2 Data tables: Countries

country_id	name_short_en	iso3_code	legacy_location_id
533	Aruba	ABW	0
4	Afghanistan	AFG	1
24	Angola	AGO	2
660	Anguilla	AIA	3
8	Albania	ALB	4
20	Andorra	AND	5
530	Netherlands Antilles	ANT	6
784	United Arab Emirates	ARE	7
32	Argentina	ARG	8
51	Armenia	ARM	9
16	American Samoa	ASM	10
10	Antarctica	ATA	11
260	French Southern and Antarctic Lands	ATF	12
28	Antigua and Barbuda	ATG	13
36	Australia	AUS	14
40	Austria	AUT	15
31	Azerbaijan	AZE	16
108	Burundi	BDI	17
56	Belgium	BEL	18
204	Benin	BEN	19
535	Bonaire	BES	20
854	Burkina Faso	BFA	21
50	Bangladesh	BGD	22
100	Bulgaria	BGR	23
48	Bahrain	BHR	24
44	The Bahamas	BHS	25
70	Bosnia and Herzegovina	BIH	26
652	Saint Barthélemy	BLM	27

Continued on next page

country_id	name_short_en	iso3_code	legacy_location_id
112	Belarus	BLR	28
84	Belize	BLZ	29
60	Bermuda	BMU	30
68	Bolivia	BOL	31
76	Brazil	BRA	32
52	Barbados	BRB	33
96	Brunei	BRN	34
64	Bhutan	BTN	35
74	Bouvet Island	BVT	36
72	Botswana	BWA	37
140	Central African Republic	CAF	38
124	Canada	CAN	39
166	Cocos (Keeling) Islands	CCK	40
756	Switzerland	CHE	41
152	Chile	CHL	42
156	China	CHN	43
384	Côte d'Ivoire	CIV	44
120	Cameroon	CMR	45
180	Democratic Republic of the Congo	COD	46
178	Republic of the Congo	COG	47
184	Cook Islands	COK	48
170	Colombia	COL	49
174	Comoros	COM	50
132	Cabo Verde	CPV	51
188	Costa Rica	CRI	52
200	Czechoslovakia	CSK	53
192	Cuba	CUB	54
531	Curaçao	CUW	55
162	Christmas Island	CXR	56
136	Cayman Islands	CYM	57
196	Cyprus	CYP	58
203	Czechia	CZE	59
278	East Germany	DDR	60
276	Germany	DEU	61
262	Djibouti	DJI	62
212	Dominica	DMA	63
208	Denmark	DNK	64
214	Dominican Republic	DOM	65
12	Algeria	DZA	66
218	Ecuador	ECU	67
818	Egypt	EGY	68
232	Eritrea	ERI	69
732	Western Sahara	ESH	70
724	Spain	ESP	71
233	Estonia	EST	72
231	Ethiopia	ETH	73
246	Finland	FIN	74
242	Fiji	FJI	75
238	Falkland Islands	FLK	76
250	France	FRA	77
234	Faroe Islands	FRO	78
583	Micronesia	FSM	79
266	Gabon	GAB	80
826	United Kingdom	GBR	81
268	Georgia	GEO	82
288	Ghana	GHA	83

Continued on next page

country_id	name_short_en	iso3_code	legacy_location_id
292	Gibraltar	GIB	84
324	Guinea	GIN	85
312	Guadeloupe	GLP	86
270	The Gambia	GMB	87
624	Guinea-Bissau	GNB	88
226	Equatorial Guinea	GNQ	89
300	Greece	GRC	90
308	Grenada	GRD	91
304	Greenland	GRL	92
320	Guatemala	GTM	93
254	French Guiana	GUF	94
316	Guam	GUM	95
328	Guyana	GUY	96
344	Hong Kong	HKG	97
334	Heard and McDonald Islands	HMD	98
340	Honduras	HND	99
191	Croatia	HRV	100
332	Haiti	HTI	101
348	Hungary	HUN	102
360	Indonesia	IDN	103
356	India	IND	104
86	British Indian Ocean Territory	IOT	105
372	Ireland	IRL	106
364	Iran	IRN	107
368	Iraq	IRQ	108
352	Iceland	ISL	109
376	Israel	ISR	110
380	Italy	ITA	111
388	Jamaica	JAM	112
400	Jordan	JOR	113
392	Japan	JPN	114
398	Kazakhstan	KAZ	115
404	Kenya	KEN	116
417	Kyrgyzstan	KGZ	117
116	Cambodia	KHM	118
296	Kiribati	KIR	119
659	Saint Kitts and Nevis	KNA	120
410	South Korea	KOR	121
414	Kuwait	KWT	122
418	Laos	LAO	123
422	Lebanon	LBN	124
430	Liberia	LBL	125
434	Libya	LYB	126
662	Saint Lucia	LCA	127
144	Sri Lanka	LKA	128
426	Lesotho	LSO	129
440	Lithuania	LTU	130
442	Luxembourg	LUX	131
428	Latvia	LVA	132
446	Macao	MAC	133
504	Morocco	MAR	134
498	Moldova	MDA	135
450	Madagascar	MDG	136
462	Maldives	MDV	137
484	Mexico	MEX	138
584	Marshall Islands	MHL	139

Continued on next page

country_id	name_short_en	iso3_code	legacy_location_id
807	North Macedonia	MKD	140
466	Mali	MLI	141
470	Malta	MLT	142
104	Myanmar	MMR	143
499	Montenegro	MNE	144
496	Mongolia	MNG	145
580	Northern Mariana Islands	MNP	146
508	Mozambique	MOZ	147
478	Mauritania	MRT	148
500	Montserrat	MSR	149
474	Martinique	MTQ	150
480	Mauritius	MUS	151
454	Malawi	MWI	152
458	Malaysia	MYS	153
175	Mayotte	MYT	154
516	Namibia	NAM	155
540	New Caledonia	NCL	156
562	Niger	NER	157
574	Norfolk Island	NFK	158
566	Nigeria	NGA	159
558	Nicaragua	NIC	160
570	Niue	NIU	161
528	Netherlands	NLD	162
578	Norway	NOR	163
524	Nepal	NPL	164
520	Nauru	NRU	165
554	New Zealand	NZL	166
512	Oman	OMN	167
586	Pakistan	PAK	168
591	Panama	PAN	169
582	Pacific Islands	PCI	170
612	Pitcairn	PCN	171
592	Panama Canal Zone	PCZ	172
604	Peru	PER	173
608	Philippines	PHL	174
585	Palau	PLW	175
598	Papua New Guinea	PNG	176
616	Poland	POL	177
408	North Korea	PRK	178
620	Portugal	PRT	179
600	Paraguay	PRY	180
275	Palestine	PSE	181
258	French Polynesia	PYF	182
634	Qatar	QAT	183
638	Réunion	REU	184
642	Romania	ROU	185
643	Russia	RUS	186
646	Rwanda	RWA	187
682	Saudi Arabia	SAU	188
891	Serbia and Montenegro	SCG	189
729	Sudan	SDN	190
686	Senegal	SEN	191
702	Singapore	SGP	192
239	South Georgia and South Sandwich Islds.	SGS	193
654	Saint Helena, Ascension and Tristan da Cunha	SHN	194
90	Solomon Islands	SLB	195

Continued on next page

country_id	name_short_en	iso3_code	legacy_location_id
694	Sierra Leone	SLE	196
222	El Salvador	SLV	197
674	San Marino	SMR	198
706	Somalia	SOM	199
666	Saint Pierre and Miquelon	SPM	200
688	Serbia	SRB	201
728	South Sudan	SSD	202
678	Sao Tome and Principe	STP	203
810	Soviet Union	SUN	204
740	Suriname	SUR	205
703	Slovakia	SVK	206
705	Slovenia	SVN	207
752	Sweden	SWE	208
748	Eswatini	SWZ	209
534	St-Martin / St Maarten	SXM	210
690	Seychelles	SYC	211
760	Syria	SYR	212
796	Turks and Caicos Islands	TCA	213
148	Chad	TCD	214
768	Togo	TGO	215
764	Thailand	THA	216
762	Tajikistan	TJK	217
772	Tokelau	TKL	218
795	Turkmenistan	TKM	219
626	Timor-Leste	TLS	220
776	Tonga	TON	221
780	Trinidad and Tobago	TTO	222
788	Tunisia	TUN	223
792	Turkiye	TUR	224
798	Tuvalu	TUV	225
834	Tanzania	TZA	226
800	Uganda	UGA	227
804	Ukraine	UKR	228
581	US Minor Outlying Islands	UMI	229
858	Uruguay	URY	230
840	United States of America	USA	231
860	Uzbekistan	UZB	232
336	Vatican City	VAT	233
670	Saint Vincent and the Grenadines	VCT	234
868	North Vietnam	VDR	235
862	Venezuela	VEN	236
92	British Virgin Islands	VGB	237
850	US Virgin Islands	VIR	238
704	Vietnam	VNM	239
548	Vanuatu	VUT	240
876	Wallis and Futuna	WLF	241
882	Samoa	WSM	242
887	Yemen	YEM	243
720	Democratic Republic of Yemen	YMD	244
890	Yugoslavia	YUG	245
710	South Africa	ZAF	246
894	Zambia	ZMB	247
716	Zimbabwe	ZWE	248
158	Taiwan	TWN	249
914	Services Partners	USP	250
999	Undeclared	ANS	250

B.3 Data tables: Products

product_id	code	name_short_en	product_level	...
1	1	Textiles	1	
2	2	Agriculture	1	
3	3	Stone	1	..
4	4	Minerals	1	
5	5	Metals	1	
		...		

Table B.3: List of products, from [43]. The column `code` matches the code from the HS nomenclature. Some columns have been ignored to fit the table.

B.4 Smoothing discrepancies:

Authors in [43] smooth discrepancies in trading flows by using mirrored trade flows to fill gaps while accounting for FOB-CIF variations. To standardize these values, they apply a discount rate of 8%, based on prior research. Let $x_{c,c',p}$ represent the exports of country c to country c' in product p and $m_{c',c,p}$ the imports of country c' from country c . Under this notation the corrected flows are defined as:

$$\hat{x}_{c,c',p} = \max \left\{ x_{c,c',p}, \frac{m_{c',c,p}}{1.08} \right\}$$

$$\hat{m}_{c,c',p} = \max \left\{ x_{c',c,p}, \frac{m_{c,c',p}}{1.08} \right\}$$

For more details, visit <https://observatory-economic-complexity.github.io/oec-documentation/data-processing.html#filling-gaps-in-our-data>⁴

B.5 Economic Complexity Index (ECI)

A detailed definition of the Economic Complexity Index (ECI) calculation can be found in [36]. There, the authors state the following: Following the product-matrix from 4.2.4, with 1 if country c produces product p , and otherwise, the authors in [36] define diversity and ubiquity simply by summing over the rows or columns of that matrix. Formally defined as:

⁴Last accessed:12/05/2025.

$$\text{Diversity} = k_{c,0} = \sum_p X_{c,p} \quad (\text{B.1})$$

$$\text{Ubiquity} = k_{p,0} = \sum_c X_{c,p} \quad (\text{B.2})$$

In order to correct the information that diversity and ubiquity carry, and thus generate a more accurate measure of the number of capabilities of a country, they calculate the average ubiquity of the products that it exports, the average diversity of the countries that make those products and so forth. This can be expressed by the recursion:

$$k_{c,N} = \frac{1}{k_{c,0}} \sum_p M_{c,p} \cdot k_{p,N-1} \quad (\text{B.3})$$

$$k_{p,N} = \frac{1}{k_{p,0}} \sum_c M_{c,p} \cdot k_{c,N-1} \quad (\text{B.4})$$

Inserting (B.4) into (B.3) we obtain

$$k_{c,N} = \frac{1}{k_{c,0}} \sum_p M_{c,p} \cdot \frac{1}{k_{p,0}} \sum_{c'} M_{c,p} \cdot k_{c,N-2} \quad (\text{B.5})$$

$$k_{c,N} = \sum_{c'} k_{c',N-2} \sum \frac{M_{cp} M_{c'p}}{k_{c,0} k_{p,0}} \quad (\text{B.6})$$

This can be rewritten as:

$$k_{c,N} = \sum_{c'} \tilde{M}_{cc'} k_{c',N-2} \quad (\text{B.7})$$

where

$$\tilde{M}_{cc'} = \sum_p \frac{M_{cp} M_{c'p}}{k_{c,0} k_{p,0}} \quad (\text{B.8})$$

We note (B.6) is satisfied when $k_{c,N} = k_{c,N-2} = 1$. This is the eigenvector of $M_{cc'}$ which is associated with the largest eigenvalue. Since this eigenvector is a vector of ones, it is not informative, and thus authors turn to the second largest eigenvalue —the eigenvector that captures the largest amount of variance in the system, and the measure of economic

complexity. Formally, they define the Economic Complexity Index (ECI) as:

$$ECI = \frac{\vec{K} - \langle \vec{K} \rangle}{\text{stdev}(\vec{K})} \quad (\text{B.9})$$

where $\langle \cdot \rangle$ represents an average, stdev stands for the standard deviation and \vec{K} = Eigenvector of $M_{cc'}$ associated with the second largest eigenvalue.

B.6 Product Complexity Index (PCI)

Following the previous section (B.5), PCI is defined as:

$$PCI = \frac{\vec{Q} - \langle \vec{Q} \rangle}{\text{stdev}(\vec{Q})} \quad (\text{B.10})$$

where \vec{Q} = Eigenvector of $M_{pp'}$ associated with the second largest eigenvalue.

B.7 Complexity Outlook Index

A detailed definition of the Economic Complexity Outlook Index (COI) can be found in <https://atlas.hks.harvard.edu/glossary>⁵. To calculate COI we first need to calculate distance of every product to existing production (from 0 to 1). We then sum the ‘closeness,’ i.e. 1 minus the distance to the products that the country is not currently making, weighted by the level of complexity of these products. Formally,

$$COI_c = \sum_p (1 - d_{cp})(1 - M_{cp})PCI_p \quad (\text{B.11})$$

where PCI is the Product Complexity Index of product p . The term $1 - M_{cp}$ ensures we only count the products that a country is not currently producing. The distance d relies on the calculation of proximity, based on the minimum conditional probability that a country that exports product p will also export product r . Since conditional probabilities are not symmetric, we take the minimum probability of product p , given product r , and vice versa.

⁵Last accessed: 12/05/2025.

Appendix C

BACI dataset

The BACI dataset contains one row per each year-exporter-importer-product combination, resulting in 130,258,877 rows x 6 columns. Exporter and importer IDs match those of the UN Comtrade dataset, and product IDs match the HS 6-digit nomenclature. Export values are mirrored (harmonized), and expressed in thousand USD, while quantities are expressed in metric tons. An extract of this data can be seen in Table C.1.

year	exporter	importer	product_id	export_value	qty
...					
2012	12	380	030219	82.96	20.54
2012	12	380	030229	1.73	0.11
2012	12	380	030285	31.43	1.99
2012	12	380	030289	193.36	12.51
2012	12	380	030621	29.37	0.47
2012	12	380	030622	1.24	0.03
2012	12	380	030629	155.75	3.30
2012	12	380	030751	0.93	0.24
2012	12	380	030760	461.74	114.36
...					

Table C.1: Extract of trade data from the BACI dataset. Contains one row per each year-exporter-importer-product combination. Exporter and importer IDs match those of the UN Comtrade dataset, and product IDs match the HS 6-digit nomenclature. Export values are mirrored, and expressed in thousand USD, and quantities in metric tons.

Appendix D

Additional Data Sources

D.1 Global Preferential Trade Agreements Database

A trade agreement describes any contractual arrangement between countries that regulates their trade relationships. These may be bilateral (between two countries) or multilateral (involving more than two countries), typically encompassing a broad range of provisions over trade conditions; such as stipulations on customs duties, fiscal and commercial regulations, transit arrangements for goods, customs valuation methods, administrative procedures, quotas, etc. These conditions can significantly influence trade dynamics between countries, potentially serving as an important source of information in network analyses.

The Global Preferential Trade Agreements Database from the World Bank contains 349 records and 7 columns¹. A representative extract of this dataset is provided in Table D.1.

Common Name	Membership	...	Type of Agreement	In-force Status
AEC	Algeria; Angola; ...; Zimbabwe	...	Regional/Plurilateral FTA	in force
Afghanistan - India	Afghanistan; India		Bilateral FTA	in force
Agadir	Egypt; Jordan; Morocco; Tunisia	...	Regional/Plurilateral FTA	unknown
Albania - Croatia	Albania; Croatia		Bilateral FTA	expired
FYROM	Albania; North Macedonia	...	Bilateral FTA	expired

Table D.1: Extract from the Global Preferential Trade Agreements Database. FTA stands for Free Trade Agreement. Some columns are ignored to fit the table.

¹Last accessed: 14-04-2025

D.2 Discrepancy Index

The Discrepancy Index data used in this work ranges from 2017 to 2022, and contains 129,443,445 rows and 9 columns. There is one row per country-country-product-year combination. Country codes are expressed in ISO3 form, while Product codes follow the HS nomenclature. An extract of this data can be seen in Table D.2.

CountryA	CountryB	ProductCode	...	Year	DI
...					
ARE	ZAF	01		2017	1.00
ARM	GEO	01		2017	0.06
BRA	ITA	01	..	2017	-0.93
CAN	GTM	01		2017	-1.00
CAN	MLI	01		2017	-1.00
...					

Table D.2: Extract from the Discrepancy Index data. CountryA and CountryB columns follow the ISO 3166-1 alpha-3 codes. ProductCode follows the HS nomenclature. Some columns are omitted to fit the table.

D.3 Corruption Perception Index

The Corruption Perception Index data used in this work ranges from 2012 to 2022, containing 181 rows and 14 columns. The full list can be seen in Table D.3.

Country	2022	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012
Afghanistan	24	16	19	16	16	15	15	11	12	8	8
Albania	36	35	36	35	36	38	39	36	33	31	33
Algeria	33	33	36	35	35	33	34	36	36	36	34
Angola	33	29	27	26	19	19	18	15	19	23	22
Argentina	38	38	42	45	40	39	36	32	34	34	35
Armenia	46	49	49	42	35	35	33	35	37	36	34
Australia	75	73	77	77	77	77	79	79	80	81	85
Austria	71	74	76	77	76	75	75	76	72	69	69
Azerbaijan	23	30	30	30	25	31	30	29	29	28	27
Bahamas	64	64	63	64	65	65	66	-	71	71	71
Bahrain	44	42	42	42	36	36	43	51	49	48	51
Bangladesh	25	26	26	26	26	28	26	25	25	27	26
Barbados	65	65	64	62	68	68	61	-	74	75	76
Belarus	39	41	47	45	44	44	40	32	31	29	31
Belgium	73	73	76	75	75	75	77	77	76	75	75
Benin	43	42	41	41	40	39	36	37	39	36	36
Bhutan	68	68	68	68	68	67	65	65	65	63	63

Continued on next page

Country	2022	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012
Bolivia	31	30	31	31	29	33	33	34	35	34	34
Bosnia and Herzegovina	34	35	35	36	38	38	39	38	39	42	42
Botswana	60	55	60	61	61	61	60	63	63	64	65
Brazil	38	38	38	35	35	37	40	38	43	42	43
Brunei Darussalam	-	-	60	60	63	62	58	-	-	60	55
Bulgaria	43	42	44	43	42	43	41	41	43	41	41
Burkina Faso	42	42	40	40	41	42	42	38	38	38	38
Burundi	17	19	19	19	17	22	20	21	20	21	19
Cabo Verde	60	58	58	58	57	55	59	55	57	58	60
Cambodia	24	23	21	20	20	21	21	21	21	20	22
Cameroon	26	27	25	25	25	25	26	27	27	25	26
Canada	74	74	77	77	81	82	82	83	81	81	84
Central African Republic	24	24	26	25	26	23	20	24	24	25	26
Chad	19	20	21	20	19	20	20	22	22	19	19
Chile	67	67	67	67	67	67	66	70	73	71	72
China	45	45	42	41	39	41	40	37	36	40	39
Colombia	39	39	39	37	36	37	37	37	37	36	36
Comoros	19	20	21	25	27	27	24	26	26	28	28
Congo	21	21	19	19	19	21	20	23	23	22	26
Costa Rica	54	58	57	56	56	59	58	55	54	53	54
Cote d'Ivoire	37	36	36	35	35	36	34	32	32	27	29
Croatia	50	47	47	47	48	49	49	51	48	48	46
Cuba	45	46	47	48	47	47	47	47	46	46	48
Cyprus	52	53	57	58	59	57	55	61	63	63	66
Czechia	56	54	54	56	59	57	55	56	51	48	49
Dem. Rep. of the Congo	20	19	18	18	20	21	21	22	22	22	21
Denmark	90	88	88	87	88	88	90	91	92	91	90
Djibouti	30	30	27	30	31	31	30	34	34	36	36
Dominica	55	55	55	55	57	57	59	-	58	58	58
Dominican Republic	32	30	28	28	30	29	31	33	32	29	32
Ecuador	36	36	39	38	34	32	31	32	33	35	32
Egypt	30	33	33	35	35	32	34	36	37	32	32
El Salvador	33	34	36	34	35	33	36	39	39	38	38
Equatorial Guinea	17	17	16	16	16	17	-	-	-	19	20
Eritrea	22	22	21	23	24	20	18	18	18	20	25
Estonia	74	74	75	74	73	71	70	70	69	68	64
Eswatini	30	32	33	34	38	39	-	-	43	39	37
Ethiopia	38	39	38	37	34	35	34	33	33	33	33
Fiji	53	55	-	-	-	-	-	-	-	-	-
Finland	87	88	85	86	85	85	89	90	89	89	90
France	72	71	69	69	72	70	69	70	69	71	71
Gabon	29	31	30	31	31	32	35	34	37	34	35
Gambia	34	37	37	37	37	30	26	28	29	28	34
Georgia	56	55	56	56	58	56	57	52	52	49	52
Germany	79	80	80	80	80	81	81	81	79	78	79
Ghana	43	43	43	41	41	40	43	47	48	46	45
Greece	52	49	50	48	45	48	44	46	43	40	36
Grenada	52	53	53	53	52	52	56	-	-	-	-
Guatemala	24	25	25	26	27	28	28	28	32	29	33
Guinea	25	25	28	29	28	27	27	25	25	24	24
Guinea Bissau	21	21	19	18	16	17	16	17	19	19	25
Guyana	40	39	41	40	37	38	34	29	30	27	28
Haiti	17	20	18	18	20	22	20	17	19	19	19
Honduras	23	23	24	26	29	29	30	31	29	26	28
Hong Kong	76	76	77	76	76	77	77	75	74	75	77
Hungary	42	43	44	44	46	45	48	51	54	54	55

Continued on next page

Country	2022	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012
Iceland	74	74	75	78	76	77	78	79	79	78	82
India	40	40	40	41	41	40	40	38	38	36	36
Indonesia	34	38	37	40	38	37	37	36	34	32	32
Iran	25	25	25	26	28	30	29	27	27	25	28
Iraq	23	23	21	20	18	18	17	16	16	16	18
Ireland	77	74	72	74	73	74	73	75	74	72	69
Israel	63	59	60	60	61	62	64	61	60	61	60
Italy	56	56	53	53	52	50	47	44	43	43	42
Jamaica	44	44	44	43	44	44	39	41	38	38	38
Japan	73	73	74	73	73	73	72	75	76	74	74
Jordan	47	49	49	48	49	48	48	53	49	45	48
Kazakhstan	36	37	38	34	31	31	29	28	29	26	28
Kenya	32	30	31	28	27	28	26	25	25	27	27
Korea, North	17	16	18	17	14	17	12	8	8	8	8
Korea, South	63	62	61	59	57	54	53	54	55	55	56
Kosovo	41	39	36	36	37	39	36	33	33	33	34
Kuwait	42	43	42	40	41	39	41	49	44	43	44
Kyrgyzstan	27	27	31	30	29	29	28	28	27	24	24
Laos	31	30	29	29	29	29	30	25	25	26	21
Latvia	59	59	57	56	58	58	57	56	55	53	49
Lebanon	24	24	25	28	28	28	28	28	27	28	30
Lesotho	37	38	41	40	41	42	39	44	49	49	45
Liberia	26	29	28	28	32	31	37	37	37	38	41
Libya	17	17	17	18	17	17	14	16	18	15	21
Lithuania	62	61	60	60	59	59	59	59	58	57	54
Luxembourg	77	81	80	80	81	82	81	85	82	80	80
Madagascar	26	26	25	24	25	24	26	28	28	28	32
Malawi	34	35	30	31	32	31	31	31	33	37	37
Malaysia	47	48	51	53	47	47	49	50	52	50	49
Maldives	40	40	43	29	31	33	36	-	-	-	-
Mali	28	29	30	29	32	31	32	35	32	28	34
Malta	51	54	53	54	54	56	55	60	55	56	57
Mauritania	30	28	29	28	27	28	27	31	30	30	31
Mauritius	50	54	53	52	51	50	54	53	54	52	57
Mexico	31	31	31	29	28	29	30	31	35	34	34
Moldova	39	36	34	32	33	31	30	33	35	35	36
Mongolia	33	35	35	35	37	36	38	39	39	38	36
Montenegro	45	46	45	45	45	46	45	44	42	44	41
Morocco	38	39	40	41	43	40	37	36	39	37	37
Mozambique	26	26	25	26	23	25	27	31	31	30	31
Myanmar	23	28	28	29	29	30	28	22	21	21	15
Namibia	49	49	51	52	53	51	52	53	49	48	48
Nepal	34	33	33	34	31	31	29	27	29	31	27
Netherlands	80	82	82	82	82	82	83	84	83	83	84
New Zealand	87	88	88	87	87	89	90	91	91	91	90
Nicaragua	19	20	22	22	25	26	26	27	28	28	29
Niger	32	31	32	32	34	33	35	34	35	34	33
Nigeria	24	24	25	26	27	27	28	26	27	25	27
North Macedonia	40	39	35	35	37	35	37	42	45	44	43
Norway	84	85	84	84	84	85	85	88	86	86	85
Oman	44	52	54	52	52	44	45	45	45	47	47
Pakistan	27	28	31	32	33	32	32	30	29	28	27
Panama	36	36	35	36	37	37	38	39	37	35	38
Papua New Guinea	30	31	27	28	28	29	28	25	25	25	25
Paraguay	28	30	28	28	29	29	30	27	24	24	25
Peru	36	36	38	36	35	37	35	36	38	38	38

Continued on next page

Country	2022	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012
Philippines	33	33	34	34	36	34	35	35	38	36	34
Poland	55	56	56	58	60	60	62	63	61	60	58
Portugal	62	62	61	62	64	63	62	64	63	62	63
Qatar	58	63	63	62	62	63	61	71	69	68	68
Romania	46	45	44	44	47	48	48	46	43	43	44
Russia	28	29	30	28	28	29	29	29	27	28	28
Rwanda	51	53	54	53	56	55	54	54	49	53	53
Saint Lucia	55	56	56	55	55	55	60	-	71	71	71
St. Vinc. and the Grenadines	60	59	59	59	58	58	60	-	62	62	62
Sao Tome and Principe	45	45	47	46	46	46	46	42	42	42	42
Saudi Arabia	51	53	53	53	49	49	46	52	49	46	44
Senegal	43	43	45	45	45	45	45	44	43	41	36
Serbia	36	38	38	39	39	41	42	40	41	42	39
Seychelles	70	70	66	66	66	60	-	55	55	54	52
Sierra Leone	34	34	33	33	30	30	30	29	31	30	31
Singapore	83	85	85	85	85	84	84	85	84	86	87
Slovakia	53	52	49	50	50	50	51	51	50	47	46
Slovenia	56	57	60	60	60	61	61	60	58	57	61
Solomon Islands	42	43	42	42	44	39	42	-	-	-	-
Somalia	12	13	12	9	10	9	10	8	8	8	8
South Africa	43	44	44	44	43	43	45	44	44	42	43
South Sudan	13	11	12	12	13	12	11	15	15	14	-
Spain	60	61	62	62	58	57	58	58	60	59	65
Sri Lanka	36	37	38	38	38	38	36	37	38	37	40
Sudan	22	20	16	16	16	16	14	12	11	11	13
Suriname	40	39	38	44	43	41	45	36	36	36	37
Sweden	83	85	85	85	85	84	88	89	87	89	88
Switzerland	82	84	85	85	85	85	86	86	86	85	86
Syria	13	13	14	13	13	14	13	18	20	17	26
Taiwan	68	68	65	65	63	63	61	62	61	61	61
Tajikistan	24	25	25	25	25	21	25	26	23	22	22
Tanzania	38	39	38	37	36	36	32	30	31	33	35
Thailand	36	35	36	36	36	37	35	38	38	35	37
Timor-Leste	42	41	40	38	35	38	35	28	28	30	33
Togo	30	30	29	29	30	32	32	32	29	29	30
Trinidad and Tobago	42	41	40	40	41	41	35	39	38	38	39
Tunisia	40	44	44	43	43	42	41	38	40	41	41
Turkey	36	38	40	39	41	40	41	42	45	50	49
Turkmenistan	19	19	19	19	20	19	22	18	17	17	17
Uganda	26	27	27	28	26	26	25	25	26	26	29
Ukraine	33	32	33	30	32	30	29	27	26	25	26
United Arab Emirates	67	69	71	71	70	71	66	70	70	69	68
United Kingdom	73	78	77	77	80	82	81	81	78	76	74
United States of America	69	67	67	69	71	75	74	76	74	73	73
Uruguay	74	73	71	71	70	70	71	74	73	73	72
Uzbekistan	31	28	26	25	23	22	21	19	18	17	17
Vanuatu	48	45	43	46	46	43	-	-	-	-	-
Venezuela	14	14	15	16	18	18	17	17	19	20	19
Vietnam	42	39	36	37	33	35	33	31	31	31	31
Yemen	16	16	15	15	14	16	14	18	19	18	23
Zambia	33	33	33	34	35	37	38	38	38	38	37
Zimbabwe	23	23	24	24	22	22	22	21	21	21	20

D.4 Country Borders

The Country Borders dataset used in this work contains 252 rows and 19 columns. Countries are expressed in their ISO2, ISO3, ISO numeric, FIPS country code, and English name. The column **neighbours**, containing the neighbouring countries of a given country, is expressed in ISO2 code. The full list of countries (with some omitted columns) can be seen in Table D.4.

#ISO	ISO3	Country	neighbours
AD	AND	Andorra	ES,FR
AE	ARE	United Arab Emirates	SA,OM
AF	AFG	Afghanistan	TM,CN,IR,TJ,PK,UZ
AG	ATG	Antigua and Barbuda	
AI	AIA	Anguilla	
AL	ALB	Albania	MK,GR,ME,RS,XK
AM	ARM	Armenia	GE,IR,AZ,TR
AO	AGO	Angola	CD,NA,ZM,CG
AQ	ATA	Antarctica	
AR	ARG	Argentina	CL,BO,UY,PY,BR
AS	ASM	American Samoa	
AT	AUT	Austria	CH,DE,HU,SK,CZ,IT,SI,LI
AU	AUS	Australia	
AW	ABW	Aruba	
AX	ALA	Aland Islands	
AZ	AZE	Azerbaijan	GE,IR,AM,TR,RU
BA	BIH	Bosnia and Herzegovina	HR,ME,RS
BB	BRB	Barbados	
BD	BGD	Bangladesh	MM,IN
BE	BEL	Belgium	DE,NL,LU,FR
BF	BFA	Burkina Faso	NE,BJ,GH,CI,TG,ML
BG	BGR	Bulgaria	MK,GR,RO,TR,RS
BH	BHR	Bahrain	
BI	BDI	Burundi	TZ,CD,RW
BJ	BEN	Benin	NE,TG,BF,NG
BL	BLM	Saint Barthelemy	
BM	BMU	Bermuda	
BN	BRN	Brunei	MY
BO	BOL	Bolivia	PE,CL,PY,BR,AR
BQ	BES	Bonaire, Saint Eustatius and Saba	
BR	BRA	Brazil	SR,PE,BO,UY,GY,PY,GF,VE,CO,AR
BS	BHS	Bahamas	
BT	BTN	Bhutan	CN,IN
BV	BVT	Bouvet Island	
BW	BWA	Botswana	ZW,ZA,NA
BY	BLR	Belarus	PL,LT,UA,RU,LV
BZ	BLZ	Belize	GT,MX
CA	CAN	Canada	US
CC	CCK	Cocos Islands	
CD	COD	Democratic Republic of the Congo	TZ,CF,SS,RW,ZM,BI,UG,CG,AO
CF	CAF	Central African Republic	TD,SD,CD,SS,CM,CG
CG	COG	Republic of the Congo	CF,GA,CD,CM,AO

Continued on next page

#ISO	ISO3	Country	neighbours
CH	CHE	Switzerland	DE,IT,LI,FR,AT
CI	CIV	Ivory Coast	LR,GH,GN,BF,ML
CK	COK	Cook Islands	
CL	CHL	Chile	PE,BO,AR
CM	CMR	Cameroon	TD,CF,GA,GQ,CG,NG
CN	CHN	China	LA,BT,TJ,KZ,MN,AF,NP,MM,KG,PK,KP,RU,VN,IN
CO	COL	Colombia	EC,PE,PA,BR,VE
CR	CRI	Costa Rica	PA,NI
CU	CUB	Cuba	US
CV	CPV	Cabo Verde	
CW	CUW	Curacao	
CX	CXR	Christmas Island	
CY	CYP	Cyprus	
CZ	CZE	Czechia	PL,DE,SK,AT
DE	DEU	Germany	CH,PL,NL,DK,BE,CZ,LU,FR,AT
DJ	DJI	Djibouti	ER,ET,SO
DK	DNK	Denmark	DE
DM	DMA	Dominica	
DO	DOM	Dominican Republic	HT
DZ	DZA	Algeria	NE,EH,LY,MR,TN,MA,ML
EC	ECU	Ecuador	PE,CO
EE	EST	Estonia	RU,LV
EG	EGY	Egypt	LY,SD,IL,PS
EH	ESH	Western Sahara	DZ,MR,MA
ER	ERI	Eritrea	ET,SD,DJ
ES	ESP	Spain	AD,PT,GI,FR,MA
ET	ETH	Ethiopia	ER,KE,SD,SS,SO,DJ
FI	FIN	Finland	NO,RU,SE
FJ	FJI	Fiji	
FK	FLK	Falkland Islands	
FM	FSM	Micronesia	
FO	FRO	Faroe Islands	
FR	FRA	France	CH,DE,BE,LU,IT,AD,MC,ES
GA	GAB	Gabon	CM,GQ,CG
GB	GBR	United Kingdom	IE
GD	GRD	Grenada	
GE	GEO	Georgia	AM,AZ,TR,RU
GF	GUF	French Guiana	SR,BR
GG	GGY	Guernsey	
GH	GHA	Ghana	CI,TG,BF
GI	GIB	Gibraltar	ES
GL	GRL	Greenland	
GM	GMB	Gambia	SN
GN	GIN	Guinea	LR,SN,SL,CI,GW,ML
GP	GLP	Guadeloupe	
GQ	GNQ	Equatorial Guinea	GA,CM
GR	GRC	Greece	AL,MK,TR,BG
GS	SGS	S. Georgia & the South Sandwich Islands	
GT	GTM	Guatemala	MX,HN,BZ,SV
GU	GUM	Guam	
GW	GNB	Guinea-Bissau	SN,GN
GY	GUY	Guyana	SR,BR,VE
HK	HKG	Hong Kong	
HM	HMD	Heard Island and McDonald Islands	
HN	HND	Honduras	GT,NL,SV

Continued on next page

#ISO	ISO3	Country	neighbours
HR	HRV	Croatia	HU,SI,BA,ME,RS
HT	HTI	Haiti	DO
HU	HUN	Hungary	SK,SI,RO,UA,HR,AT,RS
ID	IDN	Indonesia	PG,TL,MY
IE	IRL	Ireland	GB
IL	ISR	Israel	SY,JO,LB,EG,PS
IM	IMN	Isle of Man	
IN	IND	India	CN,NP,MM,BT,PK,BD
IO	IOT	British Indian Ocean Territory	
IQ	IRQ	Iraq	SY,SA,IR,JO,TR,KW
IR	IRN	Iran	TM,AF,IQ,AM,PK,AZ,TR
IS	ISL	Iceland	
IT	ITA	Italy	CH,VA,SI,SM,FR,AT
JE	JEY	Jersey	
JM	JAM	Jamaica	
JO	JOR	Jordan	SY,SA,IQ,IL,PS
JP	JPN	Japan	
KE	KEN	Kenya	ET,TZ,SS,SO,UG
KG	KGZ	Kyrgyzstan	CN,TJ,UZ,KZ
KH	KHM	Cambodia	LA,TH,VN
KI	KIR	Kiribati	
KM	COM	Comoros	
KN	KNA	Saint Kitts and Nevis	
KP	PRK	North Korea	CN,KR,RU
KR	KOR	South Korea	KP
XK	XKX	Kosovo	RS,AL,MK,ME
KW	KWT	Kuwait	SA,IQ
KY	CYM	Cayman Islands	
KZ	KAZ	Kazakhstan	TM,CN,KG,UZ,RU
LA	LAO	Laos	CN,MM,KH,TH,VN
LB	LBN	Lebanon	SY,IL
LC	LCA	Saint Lucia	
LI	LIE	Liechtenstein	CH,AT
LK	LKA	Sri Lanka	
LR	LBR	Liberia	SL,CI,GN
LS	LSO	Lesotho	ZA
LT	LTU	Lithuania	PL,BY,RU,LV
LU	LUX	Luxembourg	DE,BE,FR
LV	LVA	Latvia	LT,EE,BY,RU
LY	LBY	Libya	TD,NE,DZ,SD,TN,EG
MA	MAR	Morocco	DZ,EH,ES
MC	MCO	Monaco	FR
MD	MDA	Moldova	RO,UA
ME	MNE	Montenegro	AL,HR,BA,RS,XK
MF	MAF	Saint Martin	SX
MG	MDG	Madagascar	
MH	MHL	Marshall Islands	
MK	MKD	North Macedonia	AL,GR,BG,RS,XK
ML	MLI	Mali	SN,NE,DZ,CI,GN,MR,BF
MM	MMR	Myanmar	CN,LA,TH,BD,IN
MN	MNG	Mongolia	CN,RU
MO	MAC	Macao	
MP	MNP	Northern Mariana Islands	
MQ	MTQ	Martinique	
MR	MRT	Mauritania	SN,DZ,EH,ML
MS	MSR	Montserrat	

Continued on next page

#ISO	ISO3	Country	neighbours
MT	MLT	Malta	
MU	MUS	Mauritius	
MV	MDV	Maldives	
MW	MWI	Malawi	TZ,MZ,ZM
MX	MEX	Mexico	GT,US,BZ
MY	MYS	Malaysia	BN,TH,ID
MZ	MOZ	Mozambique	ZW,TZ,SZ,ZA,ZM,MW
NA	NAM	Namibia	ZA,BW,ZM,AO
NC	NCL	New Caledonia	
NE	NER	Niger	TD,BJ,DZ,LY,BF,NG,ML
NF	NFK	Norfolk Island	
NG	NGA	Nigeria	TD,NE,BJ,CM
NI	NIC	Nicaragua	CR,HN
NL	NLD	The Netherlands	DE,BE
NO	NOR	Norway	FI,RU,SE
NP	NPL	Nepal	CN,IN
NR	NRU	Nauru	
NU	NIU	Niue	
NZ	NZL	New Zealand	
OM	OMN	Oman	SA,YE,AE
PA	PAN	Panama	CR,CO
PE	PER	Peru	EC,CL,BO,BR,CO
PF	PYF	French Polynesia	
PG	PNG	Papua New Guinea	ID
PH	PHL	Philippines	
PK	PAK	Pakistan	CN,AF,IR,IN
PL	POL	Poland	DE,LT,SK,CZ,BY,UA,RU
PM	SPM	Saint Pierre and Miquelon	
PN	PCN	Pitcairn	
PR	PRI	Puerto Rico	
PS	PSE	Palestinian Territory	JO,IL,EG
PT	PRT	Portugal	ES
PW	PLW	Palau	
PY	PRY	Paraguay	BO,BR,AR
QA	QAT	Qatar	SA
RE	REU	Reunion	
RO	ROU	Romania	MD,HU,UA,BG,RS
RS	SRB	Serbia	AL,HU,MK,RO,HR,BA,BG,ME,XK
RU	RUS	Russia	GE,CN,BY,UA,KZ,LV,PL,EE,LT,FI,MN,NO,AZ,KP
RW	RWA	Rwanda	TZ,CD,BI,UG
SA	SAU	Saudi Arabia	QA,OM,IQ,YE,JO,AE,KW
SB	SLB	Solomon Islands	
SC	SYC	Seychelles	
SD	SDN	Sudan	SS,TD,EG,ET,ER,LY,CF
SS	SSD	South Sudan	CD,CF,ET,KE,SD,UG
SE	SWE	Sweden	NO,FI
SG	SGP	Singapore	
SH	SHN	Saint Helena	
SI	SVN	Slovenia	HU,IT,HR,AT
SJ	SJM	Svalbard and Jan Mayen	
SK	SVK	Slovakia	PL,HU,CZ,UA,AT
SL	SLE	Sierra Leone	LR,GN
SM	SMR	San Marino	IT
SN	SEN	Senegal	GN,MR,GW,GM,ML
SO	SOM	Somalia	ET,KE,DJ
SR	SUR	Suriname	GY,BR,GF

Continued on next page

#ISO	ISO3	Country	neighbours
ST	STP	Sao Tome and Principe	
SV	SLV	El Salvador	GT,HN
SX	SXM	Sint Maarten	MF
SY	SYR	Syria	IQ,JO,IL,TR,LB
SZ	SWZ	Eswatini	ZA,MZ
TC	TCA	Turks and Caicos Islands	
TD	TCD	Chad	NE,LY,CF,SD,CM,NG
TF	ATF	French Southern Territories	
TG	TGO	Togo	BJ,GH,BF
TH	THA	Thailand	LA,MM,KH,MY
TJ	TJK	Tajikistan	CN,AF,KG,UZ
TK	TKL	Tokelau	
TL	TLS	Timor Leste	ID
TM	TKM	Turkmenistan	AF,IR,UZ,KZ
TN	TUN	Tunisia	DZ,LY
TO	TON	Tonga	
TR	TUR	Turkey	SY,GE,IQ,IR,GR,AM,AZ,BG
TT	TT0	Trinidad and Tobago	
TV	TUV	Tuvalu	
TW	TWN	Taiwan	
TZ	TZA	Tanzania	MZ,KE,CD,RW,ZM,BI,UG,MW
UA	UKR	Ukraine	PL,MD,HU,SK,BY,RO,RU
UG	UGA	Uganda	TZ,KE,SS,CD,RW
UM	UMI	United States Minor Outlying Islands	
US	USA	United States	CA,MX,CU
UY	URY	Uruguay	BR,AR
UZ	UZB	Uzbekistan	TM,AF,KG,TJ,KZ
VA	VAT	Vatican	IT
VC	VCT	Saint Vincent and the Grenadines	
VE	VEN	Venezuela	GY,BR,CO
VG	VGB	British Virgin Islands	
VI	VIR	U.S. Virgin Islands	
VN	VNM	Vietnam	CN,LA,KH
VU	VUT	Vanuatu	
WF	WLF	Wallis and Futuna	
WS	WSM	Samoa	
YE	YEM	Yemen	SA,OM
YT	MYT	Mayotte	
ZA	ZAF	South Africa	ZW,SZ,MZ,BW,NA,LS
ZM	ZMB	Zambia	ZW,TZ,MZ,CD,NA,MW,AO
ZW	ZWE	Zimbabwe	ZA,MZ,BW,ZM
CS	SCG	Serbia and Montenegro	AL,HU,MK,RO,HR,BA,BG
AN	ANT	Netherlands Antilles	GP

D.5 CERDI Maritime Distance Database

The CERDI Maritime Distance Database contains a row for each country to each other country connected by a relevant seaport, totalling 51,302 rows and 5 columns. Countries are expressed in their ISO3 code, while distances in kilometres. An extract of this data can be seen in Table D.5.

iso1	iso2	seadistance	capitalport1	capitalport2	roaddistance	short
...						
ABW	AFG	16498.80		1471.30		0
ABW	AGO	9437.06		1184.17		0
ABW	AIA	956.85				0
ABW	ALB	8790.06		247.95		0
ABW	AND	7685.42		453.33		0
...						

Table D.5: Extract from the CERDI Maritime Distance Database. There is a row for each country to each other country in the dataset. Distances are expressed in kilometres.

Appendix E

Data Source Integration Diagram

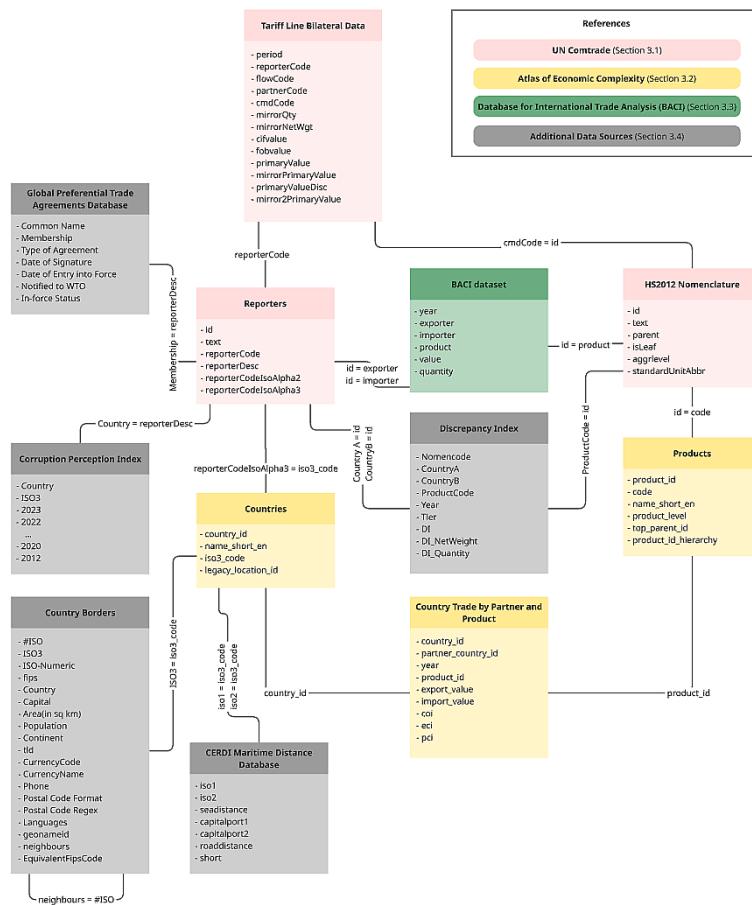


Figure E.1: Diagram of data sources schema and joins.

Appendix F

Optuna

Optuna [56] is an open source hyperparameter optimization framework to automate hyperparameter search. Its key features include:

- **Eager search spaces:** Automated search for optimal hyperparameters using Python conditionals, loops, and syntax.
- **State-of-the-art algorithms:** Efficiently search large spaces and prune unpromising trails for faster results.
- **Easy parallelization:** Parallelize hyperparameter searches over multiple threads or processes without modifying code.

In addition to this, Optuna provides a friendly user interface, where each study can be seen in a dashboard (Figure F.1), helping to keep track and monitor all hyperparameter experiments.



Figure F.1: Snippet of one of Optuna’s dashboards used in this work. The dashboard provides information on the different trials (i.e., different combinations of hyperparameters), the hyperparameters’ importance, the best trial so far, and trials timeline.

F.1 Hyperparameter tuning results

F.1.1 Hyperparameter tuning results

135

Graph	Model	Layers	Channels	Heads	Learn. Rate	Weight Decay	Dropout	Optimizer	Aggregator	Normalize	Project	Bias	Res. Connect.
E	MLP	1	64	-	0.0024	-	0.1817	AdamW	-	-	-	-	-
	GCN	2	64	-	0.0023	0.0014	0.1158	AdamW	-	-	-	-	-
	SAGE	2	128	7	0.0016	0.0058	0.1192	AdamW	Max	FALSE	FALSE	TRUE	TRUE
	GATv2	2	128	4	0.0010	0.0074	0.4000	AdamW	-	-	-	-	-
	MLP	1	64	-	0.0037	-	0.3337	AdamW	-	-	-	-	-
	GCN	2	64	-	0.0027	0.0029	0.1778	AdamW	-	-	-	-	-
T	SAGE	3	128	-	0.0003	0.0083	0.1005	AdamW	Max	FALSE	TRUE	TRUE	-
	GATv2	2	128	8	0.0005	0.0089	0.6000	AdamW	-	-	FALSE	TRUE	TRUE
	MLP	1	64	-	0.0033	-	0.2283	AdamW	-	-	-	-	-
	GCN	2	32	-	0.0063	0.0003	0.1108	AdamW	-	-	-	-	-
	SAGE	2	128	-	0.0095	0.0001	0.1372	AdaGrad	Max	FALSE	TRUE	FALSE	-
	GATv2	3	128	8	0.0100	3E-05	0.4000	AdaGrad	-	-	FALSE	TRUE	TRUE
E + L	MLP	1	128	-	0.0079	-	0.3845	AdamW	-	-	-	-	-
	GCN	2	128	-	0.0090	0.0006	0.1826	AdamW	-	-	-	-	-
	SAGE	2	8	-	0.0097	0.0093	0.1196	AdamW	Max	FALSE	FALSE	TRUE	-
	GATv2	3	32	8	0.0010	0.0065	0.4000	AdamW	-	-	FALSE	TRUE	TRUE
	MLP	1	128	-	0.0069	-	0.2920	AdamW	-	-	-	-	-
	GCN	2	64	-	0.0087	0.0023	0.2789	AdamW	-	-	-	-	-
MG T	SAGE	1	64	-	0.0096	0.0024	0.4650	AdamW	Max	FALSE	FALSE	TRUE	-
	GATv2	2	32	4	0.0050	0.0033	0.8000	AdamW	-	-	TRUE	TRUE	TRUE

Appendix G

Computational Resources

All experiments were conducted using the ITU HPC Cluster (<http://hpc.itu.dk> [internal network only]). Most runs included the following setup:

- CPU: AMD EPYC 7443
- RAM: 16 GB
- GPU: NVIDIA RTX 6000 (48 GB)
- Frameworks: PyTorch Geometric 2.7.0, Optuna 4.2.1

CO₂ Emissions from Experiments

A total of approximately 771.3 hours of computation was performed on the above-described hardware. The experiments were conducted on private infrastructure (ITU HPC) with an estimated carbon intensity of 0.173 kgCO₂eq/kWh. Using the Machine Learning Impact calculator from [72], total emissions are estimated at approximately **40.03 kgCO₂eq**, with no direct carbon offsetting applied.

Appendix H

Use of AI Assistance

Generative AI systems, specifically ChatGPT (GPT-4o), were used during the preparation of this thesis to support language refinement. Their application was strictly limited to paraphrasing and improving the clarity of the author's original text. No novel ideas or substantive content suggestions were provided by these tools. This use is comparable to conventional language assistance resources such as Grammarly, spell checkers¹, dictionaries, and thesauri. As a non-native English speaker, I employed these tools exclusively for linguistic enhancement, ensuring that the meaning and intent of the content remained entirely my own. This usage is compliant with ITU's **Generative AI Guidelines for Students** (v3, March 2025).

¹Also used, as they are integrated into Overleaf.

Bibliography

- [1] Yuval Noah Harari. *21 Lessons for the 21st Century: 'Truly mind-expanding... Ultra-topical'* *Guardian*. Random House, 2018.
- [2] Shijie Sun. Research on the application of network analysis methods in macroeconomics. *Science, Technology and Social Development Proceedings Series*, 1:10–70088, 2024.
- [3] Julian Maluck and Reik V Donner. A network of networks perspective on global trade. *PloS one*, 10(7):e0133310, 2015.
- [4] World Trade Organization - Statistics: Trade Evolution. https://www.wto.org/english/res_e/statis_e/trade_evolution_e/trade_wto_e.htm.
- [5] Nassim Nicholas Taleb. *Antifragile: Things that gain from disorder*, volume 3. Random House Trade Paperbacks, 2014.
- [6] Ricardo David. On the principles of political economy and taxation, 1971.
- [7] Adam Smith. *The wealth of nations [1776]*, volume 11937. na, 1937.
- [8] Daniel M Bernhofen and John C Brown. A direct test of the theory of comparative advantage: the case of japan. *Journal of Political Economy*, 112(1):48–67, 2004.
- [9] Jacob Viner. Cost curves and supply curves. In *The Foundations of Price Theory Vol 4*, pages 181–218. Routledge, 2024.
- [10] Bertil Gotthard Ohlin. Interregional and international trade. (*No Title*), 1967.

- [11] Stephen S Golub and Chang-Tai Hsieh. Classical ricardian theory of comparative advantage revisited. *Review of international economics*, 8(2):221–234, 2000.
- [12] GV Vijayasri. The importance of international trade in the world. *International Journal of Marketing, Financial Services & Management Research*, 2(9):111–119, 2013.
- [13] Christopher Bren d'Amour, Leonie Wenz, Matthias Kalkuhl, Jan Christoph Steckel, and Felix Creutzig. Teleconnected food supply shocks. *Environmental Research Letters*, 11(3):035007, 2016.
- [14] Philippe Marchand, Joel A Carr, Jampel Dell'Angelo, Marianela Fader, Jessica A Gephart, Matti Kummu, Nicholas R Magliocca, Miina Porkka, Michael J Puma, Zak Ratajczak, et al. Reserves and trade jointly determine exposure to food supply shocks. *Environmental Research Letters*, 11(9):095009, 2016.
- [15] Saleh Shahriar, Lu Qian, Sokvibol Kea, and Nazir Muhammad Abdullahi. The gravity model of trade: A theoretical perspective. *Review of Innovation and Competitiveness: A Journal of Economic and Social Research*, 5(1):21–42, 2019.
- [16] Alan V Deardorff. *Determinants of bilateral trade: does gravity work in a neoclassical world?*, volume 5377. National Bureau of Economic Research Cambridge, MA, USA, 1995.
- [17] Frank Schweitzer, Giorgio Fagiolo, Didier Sornette, Fernando Vega-Redondo, Alessandro Vespignani, and Douglas R White. Economic networks: The new challenges. *science*, 325(5939):422–425, 2009.
- [18] Raja Kali and Javier Reyes. Financial contagion on the international trade network. *Economic Inquiry*, 48(4):1072–1101, 2010.
- [19] Daron Acemoglu, Ufuk Akcigit, and William Kerr. Networks and the macroeconomy: An empirical exploration. *Nber macroeconomics annual*, 30(1):273–335, 2016.
- [20] Ester Gutiérrez-Moya, Belarmino Adenso-Díaz, and Sebastián Lozano. Analysis and vulnerability of the international wheat trade network. *Food security*, 13:113–128, 2021.

- [21] Diego Garlaschelli and Maria I Loffredo. Structure and evolution of the world trade network. *Physica A: Statistical Mechanics and its Applications*, 355(1):138–144, 2005.
- [22] Ma Angeles Serrano and Marián Boguná. Topology of the world trade web. *Physical Review E*, 68(1):015101, 2003.
- [23] Giorgio Fagiolo, Javier Reyes, and Stefano Schiavo. On the topological properties of the world trade web: A weighted network analysis. *Physica A: Statistical Mechanics and its Applications*, 387(15):3868–3873, 2008.
- [24] Marco Grassia, Giuseppe Mangioni, Stefano Schiavo, and Silvio Traverso. Insights into countries' exposure and vulnerability to food trade shocks from network-based simulations. *Scientific Reports*, 12(1):4644, 2022.
- [25] Zoltán Lakner, Erzsébet Szabó, Viktória Szűcs, and András Székács. Network and vulnerability analysis of international spice trade. *Food Control*, 83:141–146, 2018.
- [26] Gaojian Ji, Honglin Zhong, Harold L Feukam Nzudie, Peng Wang, and Peipei Tian. The structure, dynamics, and vulnerability of the global food trade network. *Journal of Cleaner Production*, 434:140439, 2024.
- [27] Jessica A Gephart, Elena Rovenskaya, Ulf Dieckmann, Michael L Pace, and Åke Brännström. Vulnerability to shocks in the global seafood trade network. *Environmental Research Letters*, 11(3):035008, 2016.
- [28] Bharti Khemani, Shruti Patil, Ketan Kotecha, and Sudeep Tanwar. A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions. *Journal of Big Data*, 11(1):18, 2024.
- [29] Naoto Minakawa, Kiyoshi Izumi, and Hiroki Sakaji. Bilateral trade flow prediction by gravity-informed graph auto-encoder. In 2022 *IEEE International Conference on Big Data (Big Data)*, pages 2327–2332. IEEE, 2022.

- [30] Kobby Panford-Quainoo, Avishek Joey Bose, and Michaël Defferrard. Bilateral trade modelling with graph neural networks. In *ICLR workshop on practical ML for developing countries*, 2020.
- [31] Anderson Monken, Flora Haberkorn, Munisamy Gopinath, Laura Freeman, and Feras A Batarseh. Graph neural networks for modeling causality in international trade. In *The international FLAIRS conference proceedings*, volume 34, 2021.
- [32] Bassem Sellami, Chahinez Ounoughi, Tarmo Kalvet, Marek Tiits, and Diego Rincon-Yanez. Harnessing graph neural networks to predict international trade flows. *Big Data and Cognitive Computing*, 8(6):65, 2024.
- [33] Debasish Jana, Sven Malama, Sriram Narasimhan, and Ertugrul Taciroglu. Edge-based graph neural network for ranking critical road segments in a network. *Plos one*, 18(12):e0296045, 2023.
- [34] Vicente Balmaseda, María Coronado, and Gonzalo de Cadenas-Santiago. Predicting systemic risk in financial systems using deep graph learning. *Intelligent Systems with Applications*, 19:200240, 2023.
- [35] Siddhesh Vishwanath Kaushik and Sonja Mitikj. Bridging the gap in trade reporting: Insights from the discrepancy index, 2024.
- [36] Hausmann R., Hidalgo C., Bustos S., and Coscia M. and Chung S. and Jimenez J. and Simoes A. and Yildirim M. *The Atlas of Economic Complexity*. Cambridge, MA: MIT Press., 2013.
- [37] Chuke Chen, Zhihan Jiang, Nan Li, Heming Wang, Peng Wang, Zhihe Zhang, Chao Zhang, Fengmei Ma, Yuanyi Huang, Xiaohui Lu, et al. Advancing un comtrade for physical trade flow analysis: review of data quality issues and solutions. *Resources, Conservation and Recycling*, 186:106526, 2022.
- [38] Célestin Coquidé, Leonardo Ermann, José Lages, and Dima L Shepelyansky. Influence of petroleum and gas trade on eu economies from the reduced google matrix analysis of un comtrade data. *The European Physical Journal B*, 92:1–14, 2019.
- [39] Guillaume Gaulier and Soledad Zignago. Baci: international trade database at the product-level (the 1994-2007 version). 2010.

- [40] Thomas Cantens. Mirror analysis: Customs risk analysis and fraud detection. *Global Trade and Customs Journal*, 10(6), 2015.
- [41] Cyril Chalendard, Ana M Fernandes, Gael Raballand, and Bob Rijkers. Technology (ab) use and corruption in customs. *nd* <https://www.freit.org/ETOS/papers/fernandes.pdf>, 2020.
- [42] Banu Demir and Beata Javorcik. Trade policy changes, tax evasion and benford's law. *Journal of Development Economics*, 144:102456, 2020.
- [43] The Growth Lab at Harvard University. Growth projections and complexity rankings, v2 [data set], 2019.
- [44] Simone Bertoli, Michaël Goujon, and Olivier Santoni. The cerdi-seadistance database, 2016.
- [45] M Ángeles Serrano, Marián Boguñá, and Alessandro Vespignani. Patterns of dominant flows in the world trade web. *Journal of Economic Interaction and Coordination*, 2(2):111–124, 2007.
- [46] Zsolt Tibor Kosztýán, Dénes Kiss, and Beáta Fehérvölgyi. Trade network dynamics in a globalized environment and on the edge of crises. *Journal of Cleaner Production*, 465:142699, 2024.
- [47] Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 855–864, 2016.
- [48] Albert O Hirschman. *National power and the structure of foreign trade*, volume 105. Univ of California Press, 1980.
- [49] Matti Kummu, Pekka Kinnunen, Elina Lehikoinen, Miina Porkka, Cibele Queiroz, Elin Röös, Max Troell, and Charlotte Weil. Interplay of trade and food system resilience: Gains on supply diversity over time at the cost of trade independency. *Global Food Security*, 24:100360, 2020.
- [50] Mauro Barone and Michele Coscia. Birds of a feather scam together: trustworthiness homophily in a business network. *Social Networks*, 54:228–237, 2018.

- [51] Ville Satopaa, Jeannie Albrecht, David Irwin, and Barath Raghavan. Finding a "kneedle" in a haystack: Detecting knee points in system behavior. In *2011 31st international conference on distributed computing systems workshops*, pages 166–171. IEEE, 2011.
- [52] Jean-Noël Barrot and Julien Sauvagnat. Input specificity and the propagation of idiosyncratic shocks in production networks. *The Quarterly Journal of Economics*, 131(3):1543–1592, 2016.
- [53] Will Hamilton, Zhitao Ying, and Jure Leskovec. Inductive representation learning on large graphs. *Advances in neural information processing systems*, 30, 2017.
- [54] Thomas Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *ArXiv*, abs/1609.02907, 2016.
- [55] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.
- [56] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019.
- [57] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [58] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.
- [59] Ilya Sutskever, James Martens, George Dahl, and Geoffrey Hinton. On the importance of initialization and momentum in deep learning. In *International conference on machine learning*, pages 1139–1147. PMLR, 2013.
- [60] Tijmen Tieleman. Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. *COURSERA: Neural networks for machine learning*, 4(2):26, 2012.
- [61] John Duchi, Elad Hazan, and Yoram Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of machine learning research*, 12(7), 2011.

- [62] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- [63] Shaked Brody, Uri Alon, and Eran Yahav. How attentive are graph attention networks? *arXiv preprint arXiv:2105.14491*, 2021.
- [64] Vinod Nair and Geoffrey E Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pages 807–814, 2010.
- [65] Gabriel Bénédict, Vincent Koops, Daan Odijk, and Maarten de Rijke. Sigmoidf1: A smooth f1 score surrogate loss for multilabel classification. *arXiv preprint arXiv:2108.10566*, 2021.
- [66] Wayne H Wagner and Sheila C Lau. The effect of diversification on risk. *Financial Analysts Journal*, 27(6):48–53, 1971.
- [67] Corrado Gini. Concentration and dependency ratios. *Rivista di politica economica*, 87:769–792, 1997.
- [68] Derek Kellenberg and Arik Levinson. Misreporting trade: Tariff evasion, corruption, and auditing standards. *Review of International Economics*, 27(1):106–129, 2019.
- [69] Eyke Hüllermeier and Willem Waegeman. Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods. *Machine learning*, 110(3):457–506, 2021.
- [70] Chirag Agarwal, Owen Queen, Himabindu Lakkaraju, and Marinka Zitnik. Evaluating explainability for graph neural networks. *Scientific Data*, 10(1):144, 2023.
- [71] Zhitao Ying, Dylan Bourgeois, Jiaxuan You, Marinka Zitnik, and Jure Leskovec. Gnnexplainer: Generating explanations for graph neural networks. *Advances in neural information processing systems*, 32, 2019.
- [72] Alexandre Lacoste, Alexandra Luccioni, Victor Schmidt, and Thomas Dandres. Quantifying the carbon emissions of machine learning. *arXiv preprint arXiv:1910.09700*, 2019.