BACHELOR'S DEGREE THESIS

Degree in Statistics

Title: The Impact of ICT on International Trade: A Gravity Model Analysis

Author: Songrong Cao

Advisor: Raul Ramos Lobo

Department: Department of Econometry, Statistics and Applied Economics

Academic year: June 2024

UNIVERSITAT DE BARCELONA

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Abstract

This thesis investigates the impact of Information and Communications Technology (ICT) on bilateral trade flows using data from 188 countries between 2010 and 2020, sourced from CEPII and World Development Indicators (WDI). Employing a fixed effect PPML gravity model for panel data, the study examines the influence of ICT indicators, such as mobile cellular subscriptions, internet servers, broadband connections, and internet usage, on trade flows. The analysis categorizes countries into four income groups based on World Bank classifications to explore the varying effects across different economic contexts.

The findings indicate that traditional gravity model variables (GDP, distance, contiguity, and common language) are consistently significant predictors of trade flows and align well with established hypotheses. In contrast, the impact of ICT variables on trade flows presents mixed results. While internet usage shows a positive and significant effect in upper middle-income countries, most ICT indicators do not exhibit significant effects across the different income groups. Additionally, the interaction model reveals a weak but positive significant effect on trade flows when both trade partners invest in ICT simultaneously. However, the individual effects of ICT remain largely insignificant or negative. This highlights the need for further research to better understand the complex interplay between technological and economic factors in international trade.

Key word

ICT, international trade, gravity model, fixed effect, panel data, income groups, Poisson Pseudo Maximum Likelihood.

AMS CLASSIFICATION

The main classification codes for this thesis, according to the American Mathematical Society (AMS), based on the statistical methods used, are 62J05 (Linear Regression), 62J12 (Generalized Linear Models), and 91B60 (Trade Models).

Resumen

Esta tesis investiga el impacto de las Tecnologías de la Información y la Comunicación (TIC) en los flujos comerciales bilaterales utilizando datos de 188 países entre 2010 y 2020, obtenidos de CEPII y de los Indicadores de Desarrollo Mundial (WDI). Empleando un modelo de gravedad PPML de efectos fijos para datos de panel, el estudio examina la influencia de indicadores de TIC, como suscripciones de telefonía móvil, servidores de internet, conexiones de banda ancha y uso de internet, en los flujos comerciales. El análisis categoriza a los países en cuatro grupos de ingresos según las clasificaciones del Banco Mundial para explorar los efectos variables en diferentes contextos económicos.

Los hallazgos indican que las variables tradicionales del modelo de gravedad (PIB, distancia, contigüidad y lengua común) son consistentemente predictores significativos de los flujos comerciales y se alinean bien con las hipótesis establecidas. En contraste, el impacto de las variables TIC en los flujos comerciales presenta resultados mixtos. Mientras que el uso de internet muestra un efecto positivo y significativo en los países de ingresos medios-altos, la mayoría de los indicadores de TIC no presentan efectos significativos en los diferentes grupos de ingresos. Además, el modelo de interacción revela un efecto positivo pero débil en los flujos comerciales cuando ambos socios comerciales invierten simultáneamente en TIC. Sin embargo, los efectos individuales de las TIC siguen siendo en gran parte insignificantes o negativos. Esto resalta la necesidad de una mayor investigación para comprender mejor la compleja interacción entre los factores tecnológicos y económicos en el comercio internacional.

Palabras clave

TIC, comercio internacional, modelo gravitacional, efectos fijos, datos de panel, grupos de ingresos, Poisson Pseudo Máxima Verosimilitud.

Clasificación AMS

El código principal para este trabajo de fin de grado, de acuerdo con la Sociedad Americana de Matemáticas (AMS), debido a los métodos estadísticos que se han usado, es 62J05 (Regresión Lineal), 62J12 (Modelos Lineales Generalizados) y 91B60 (Modelos de Comercio).

Table of Contents

1 Introduction

In an era marked by rapid technological advancements, Information and Communications Technology (ICT) has emerged as a pivotal factor influencing various sectors, including international trade. The proliferation of ICT tools, such as mobile cellular networks, internet servers, broadband connections, and internet usage, has revolutionized how countries engage in trade, offering new opportunities and efficiencies. Despite the significant potential of ICT to enhance trade flows, there remains a need for empirical studies that explore this relationship across different income groups and over an extended period.

This thesis aims to analyze the impact of ICT on bilateral trade flows using data from 188 countries, sourced from the Centre d'Études Prospectives et d'Informations Internationales (CEPII), over the period from 2010 to 2020. ICT indicators, including mobile cellular subscriptions, internet servers, broadband connections, and internet usage, are obtained from the World Development Indicators (WDI). Employing a fixed effect gravity model for panel data, this study investigates whether these ICT variables are significantly related to trade flows.

To provide a nuanced analysis, this study categorizes countries into four income groups based on the World Bank's classification. These groups are defined as low-income, lower-middle-income, upper-middle-income, and high-income countries. This classification allows us to examine how the effects of ICT on trade performance vary across different economic contexts.

The central research questions guiding this thesis are: To what extent does the usage of different means of ICT contribute to trade performance? And what are the effects of ICT and other gravity variables on different income groups?

The thesis begins with an introduction outlining the research objectives and significance. It then reviews the theoretical foundations of the gravity model and the role of ICT in international trade. The methodology section describes the data sources, variables, and fixed effects model used for analysis. An exploratory data analysis follows, examining variable relationships and distributions. The results of the fixed effects models are presented, including analyses for different income groups and robustness check. The thesis concludes with a summary of findings, implications, limitations, and future research directions, followed by references and appendices for supplementary information.

2 Literature review

This section will provide an overview of the existing research on the relationship between Information and Communication Technology (ICT) and international trade, with a particular focus on the gravity model of trade. We will explore how various ICT components, such as the internet, mobile technologies, broadband connectivity, and secure internet servers, impact trade performance. Additionally, we will examine how these effects differ across countries with different income levels. This review will draw from a range of studies to provide a comprehensive understanding of the current state of knowledge in this field.

2.1 The Gravity Model of Trade

The gravity model, rooted in Newtonian physics, has become a fundamental tool in the empirical analysis of international trade flows. Tinbergen (1962) [1]and Pöyhönen (1963) first applied the model to trade, proposing that the volume of trade between two countries is directly proportional to their economic sizes (typically measured by GDP) and inversely proportional to the distance between them. Over the years, the gravity model has been extended to include various additional factors such as common language, colonial ties, and trade agreements, enhancing its explanatory power (Anderson and van Wincoop, 2003).

Anderson (1979) provided the first solid theoretical foundation for the gravity model by incorporating it into a general equilibrium framework. Subsequent studies by Helpman and Krugman (1985) and Bergstrand (1985) further refined the model, incorporating elements of monopolistic competition and economies of scale. These theoretical advancements have made the gravity model a standard tool for analyzing trade flows and evaluating the impact of trade policies and agreements.

2.2 ICT and International Trade

The advent of ICT has significantly transformed international trade by reducing transaction costs, enhancing communication, and facilitating the exchange of information. Freund and Weinhold (2002) found that internet adoption positively impacts export growth, particularly in developing countries, by providing access to a

[1] (Leibenstein and Tinbergen, 1966)

broader market and improving information flow. Similarly, Clarke and Wallsten (2006) highlighted that internet penetration boosts trade in services by enabling firms to offer digital products and services globally.

Empirical studies have shown that mobile phone penetration is another crucial factor influencing trade. Aker and Mbiti (2010) reviewed the impact of mobile phones on African economies, concluding that mobile phone access reduces information asymmetries, lowers transaction costs, and increases market efficiency. This, in turn, facilitates trade by connecting buyers and sellers more effectively.

Broadband connectivity, representing high-speed internet access, is also a significant driver of trade. A study by Czernich et al. (2011) demonstrated that broadband adoption leads to increased productivity and economic growth, which indirectly boosts trade. Furthermore, secure internet servers are critical for ensuring the safety and reliability of digital transactions, thus fostering international trade. A study by Zhongwei Xing and Julien Grollier (2018) demonstrated that ICT and e-commerce significantly enhance bilateral trade flows. Increasing fixed and mobile phone subscriptions, high-speed broadband, and secure internet services were found to boost trade. Greater internet adoption and increased usage of B2B and B2C platforms also stimulate trade by improving market access and efficiency, enhancing communication, and reducing transaction costs.

2.3 ICT and Trade Performance Across Income Groups

The impact of ICT on trade varies significantly across countries with different income levels. High-income countries typically have more advanced ICT infrastructure, which enhances their trade performance. A study by Freund and Weinhold (2004) indicated that high-income countries benefit more from internet penetration due to their established digital markets and better ICT capabilities.

2.4 Empirical Evidence on ICT and Trade

Numerous empirical studies have examined the relationship between ICT and international trade using various methodologies, including the gravity model. Freund and Weinhold (2002) used a gravity model to analyze the impact of internet penetration on bilateral trade flows, finding a positive and significant effect. Similarly, Choi (2010) employed a gravity model to study the impact of broadband adoption on trade, concluding that countries with higher broadband penetration experience greater trade flows.

2.5 The Role of Traditional Factors

Traditional factors, such as geographical distance, common language, and historical ties, can influence trade flows. Disdier and Head (2008) conducted a meta-analysis of gravity model estimates and found that distance remains a significant barrier to trade, although its impact has diminished with advancements in ICT. Similarly, common language and colonial ties have been shown to facilitate trade by reducing communication barriers and fostering trust (Melitz, 2008).

3 Background knowledge

This section provides a comprehensive overview of the fundamental concepts of ICT development and the gravity model, highlighting their significance in the context of international trade and economic activities.

3.1 ICT development

In the context of ICT (Information and Communication Technology) development, four critical components play a significant role in enhancing digital infrastructure and supporting economic activities, including international trade.

Internet servers are fundamental for hosting websites, online services, and applications, which facilitate the exchange of trade-related information and transactions. The presence of a robust network of internet servers within a country indicates advanced digital infrastructure, crucial for managing logistics, communication, and operational efficiency in businesses.

Mobile cellular subscriptions refer to the number of mobile phone connections. These are essential for enabling widespread communication and providing access to information on the go. Mobile cellular networks support the coordination of trade activities, ensuring timely delivery of goods and services, and maintaining continuous contact between trade partners.

Broadband subscriptions measure the extent of high-speed internet access in a country. Broadband connectivity is vital for various digital trade activities, such as e-commerce, online banking, and virtual meetings. Reliable and fast internet access through broadband supports businesses in operating efficiently within the digital marketplace.

Internet users denote the number of people who have access to and use the internet. A higher number of internet users reflects greater digital literacy and the widespread adoption of internet services. This is an important indicator of a country’s readiness to engage in digital trade, as it signifies the population's ability to effectively use online platforms and tools for business and communication.

These components collectively highlight the importance of ICT development in facilitating and enhancing international trade by providing the necessary digital infrastructure and connectivity.

3.2 Gravity model

The gravity model has become a cornerstone in the empirical analysis of international trade patterns. Initially lacking robust theoretical foundations, the model gained credibility with Anderson's (1979) work, which provided the first theoretical underpinnings. Since then, extensive literature has built upon these foundations, emphasizing the importance of proper specification to obtain unbiased results. The model has evolved to include a variety of covariates that capture natural, historical, cultural, and policy-related factors affecting bilateral trade.

The gravity model is widely used to analyze and understand the determinants of bilateral trade. Its applications extend to evaluating trade policies and assessing the effects of economic integration. The model's robustness and ability to incorporate various trade-influencing factors make it a valuable tool for policymakers and researchers.

3.2.1 Core Equation

The gravity model is conceptually derived from Newton's law of gravitation. In its simplest form, it posits that the trade flow (X_{ij}) between two countries is proportional to the product of their economic sizes (usually measured by GDP) and inversely proportional to the distance between them. The basic form of the gravity equation is:

$X_{ij}=AY_i \cdot Y_j / D_{ij}$ (1)

Where:

- X_{ij} is the trade flow from country i to country j.
- Y_i and Y_j are the GDPs of countries i and j, respectively.
- D_{ij} is the distance between the two countries.
- A is a constant.

To make the gravity model more applicable for econometric analysis, it is often transformed into a log-linear form. This transformation simplifies the multiplicative relationships into additive ones, facilitating the use of linear regression techniques. The log-linear form of the gravity model is expressed as:

$\log(X_{ij})=\log(A)+\beta_1\log(Y_i)+\beta_2\log(Y_j)-\beta_3\log(D_{ij})+\epsilon_{ij}$ (2)

This form allows for the inclusion of additional variables that capture other factors influencing trade, such as common language, colonial ties, and trade agreements.

3.2.2 OLS model

OLS model is the most basic, the intuitive gravity model takes the following log-linearized form. Is is written as:

$\log(Trade_{ij})=\beta_0+\beta_1\log(GDP_i)+\beta_2\log(GDP_j)+\beta 3\log(D_{ij})+\epsilon_{ij}$ (3)

where the objective is to estimate the unknow parameter β. As the name suggests, OLS minimizes the sum of squared errors ϵij. Under certain assumptions as to the error term , OLS gives parameter estimates that are not only intuitively appealing but have useful statistical properties that enable us to conduct hypothesis tests and draw inferences. Basic econometric theory lays down three necessary and sufficient conditions:

1. The errors must have mean zero and be uncorrelated with each of the explanatory variables (the orthogonality assumption).
2. The errors must be independently drawn from a normal distribution with a given (fixed) variance (the homoskedasticity assumption).
3. None of the explanatory variables is a linear combination of other explanatory variables (the full rank assumption).

If all three properties hold, then OLS estimates are consistent, unbiased, and efficient within the class of linear models.[2]

3.2.3 PPML model

$Trade_{ij}=\exp(\beta_0+\beta_1\log(GDP_i)+\beta_2\log(GDP_j)+\beta 3\log(D_{ij}))+\epsilon_{ij}$ (4)

The Poisson Pseudo Maximum Likelihood (PPML) gravity model is a robust method used to estimate trade flows while addressing specific limitations of the traditional Ordinary Least Squares (OLS) gravity model. The PPML approach is particularly effective in dealing with heteroskedasticity and the presence of zero trade flows, which are common issues in trade data.

[2] https://www.unescap.org/sites/default/d8files/5%20-%203.%20Estimating%20the%20Gravity%20Model_0.pdf

3.2.4 Fixed effect Model

The gravity model of trade is an empirical framework used to predict bilateral trade flows between two countries based on their economic sizes and the distance between them. The fixed effects estimation approach is a robust method within this framework, accounting for unobserved heterogeneity by including exporter and importer fixed effects. These fixed effects are dummy variables representing each exporter and importer, capturing all constant characteristics specific to each country. The fixed effects gravity model can be represented as follows:

$Trade_{ij}=C+\alpha_i+\gamma_j+\beta X_{ij}+\epsilon_{ij}$ (5)

Here, C is the regression constant, theoretically equal to world GDP. The terms α_i and γ_j are the exporter and importer fixed effects, respectively, which control for country-specific factors. X_{ij} represents the bilateral variables (e.g., distance, common language), and ϵ_{ij} is the error term. This approach ensures consistent and unbiased estimates by isolating the impact of bilateral trade determinants from country-specific influences.[3]

[3] Correia, S., Guimarães, P., Zylkin, T., 2020. ppmlhdfe: Fast Poisson Estimation with High-Dimensional Fixed Effects. The Stata Journal 20, 95–115. https://doi.org/10.1177/1536867X20909691

4 Hypothesis

In defining our gravity model to analyze international trade flows, we include various independent variables that are expected to influence trade between countries. The hypotheses for each variable and their expected effects on trade flows are outlined below.

Variable	Hypothesis	Expected Effect	Definition
Distance_{ij}	Greater distance between countries reduces trade flows.	Negative. Increased distance raises transportation costs and logistical complexities, thus diminishing trade.	Population-weighted distance between most populated cities (arithmetic mean)
Contigious_{ij}	Countries that share a common border will trade more with each other.	Positive. Shared borders reduce transportation costs and foster stronger economic ties.	Dummy equal to 1 if countries are contiguous
Language_{ij}	Countries sharing a common language will have higher trade flows.	Positive. A common language reduces communication barriers and transaction costs.	1 if countries share common official or primary language
Cellphone	Higher mobile phone penetration in a country enhances its trade capabilities.	Positive. Mobile technology improves communication and access to market information.	Mobile cellular subscriptions (per 100 people)
Broadband	Greater broadband connectivity boosts trade flows.	Positive. Broadband infrastructure facilitates efficient business operations and e-commerce.	Fixed broadband subscriptions (per 100 people)
InternetSecurity	A higher number of secure internet servers in a country increases its trade volume.	Positive. Secure internet infrastructure ensures safe and reliable digital transactions.	Secure Internet servers (per 1 million people)
InternetUser	Higher internet usage among the population promotes trade.	Positive. Greater internet adoption enhances market access and consumer reach.	Individuals using the Internet (% of population)
GDP	Larger economies have greater trade flows.	Positive. Higher GDP reflects greater production capacity and market size.	GDP (current US$)

5 Methodology

In this section, we will introduce the typology of the data we will be using, which is panel data. We will provide an overview of the databases used and their sources, followed by a discussion of the limitations and theoretical aspects of the traditional gravity model. Finally, we will outline the modeling methods, fixed effect PPML gravity models, that will be employed in this study to analyze the impact of ICT on global trade flows.

5.1 Panel data

Panel data, also known as longitudinal data, is a type of data that combines cross-sectional and time-series data. It follows multiple subjects (such as individuals, firms, or countries) over several time periods. This structure allows for the analysis of dynamic changes over time within the same subjects.

In mathematical terms, panel data can be represented as y_{ijt}, where:

- y is the dependent variable (e.g., trade flow between two countries).
- i denotes the cross-sectional unit (e.g., country of origin).
- j denotes another cross-sectional unit (e.g., country of destination).
- t denotes the time.

There are 2 types of panel data, balanced panel data and unbalanced panel data.

A **balanced panel** is a dataset in which *each* panel member (i.e., person) is observed *every* year. Consequently, if a balanced panel contains N panel members and T periods, the number of observations (n) in the dataset is necessarily $n = N \cdot T$.

An **unbalanced panel** is a dataset in which *at least one* panel member is not observed every period. Therefore, if an unbalanced panel contains N panel members and T periods, then the following strict inequality holds for the number of observations (n) in the dataset: $n < N \cdot T$.[4]

In this research we will use balanced panel data combined with gravity model to analyze the impact of ICT and e-commerce on the global trade flow.

[4] Panel data, 2024. . Wikipedia. https://en.wikipedia.org/wiki/Panel_data

5.2 Data source

The data source mainly comes from 2 databases, The CEPII Gravity Database and WDI. The CEPII Gravity database is a comprehensive resource for researchers and practitioners estimating gravity equations. Each observation in the database represents a combination of an exporter country, an importing country, and a year ("origin-destination-year"). It includes trade flows and a range of geographic, cultural, trade facilitation, and macroeconomic variables. The data spans from 1948 to 2021 and covers 252 countries, some of which only exist for shorter periods. The term "country" may include non-independent territories and historical territorial configurations.

The dataset is dynamic, reflecting changes in country boundaries over time, and is "squared," meaning every pair of countries appears each year, with missing values for non-existent countries. Gravity is the primary dataset, with countries identified using a unique variable, country_id, which combines the ISO3 alphabetic code with a number to account for territorial changes.

In the dataset, we extracted data from 2010 to 2020 to avoid the distortions caused by the COVID-19 pandemic, which began in 2020 and would likely affect the accuracy of the model. We selected specific variables from the CEPII Gravity database, which include:

- year: The year from 1990 to 2021.
- country_id_o: The country ID of the origin country (unilateral).
- country_id_d: The country ID of the destination country (unilateral).
- Tradeflow_imf_o : Trade flows as reported by the origin, in 1000 current USD. Source: IMF, bilateral.
- Distanceij: The population-weighted distance between the most populated cities (arithmetic mean) (bilateral).
- Contiguousij: A dummy variable equal to 1 if the countries are contiguous (bilateral).
- Languageij: A dummy variable equal to 1 if the countries share a common official or primary language (bilateral).
- Gdp_o: GDP (current thousands US$) of the origin country (unilateral)
- Gdp_d: GDP (current thousands US$) of the destination country (unilateral)

World Development Indicators (WDI) is the primary World Bank collection of development indicators, compiled from officially recognized international sources. It presents the most current and accurate global development data available, and includes national, regional and global estimates. [5]

In this paper, we extract data from the World Development Indicators (WDI) database for the period 2010 to 2020. We focus on indicators related to ICT development, including:

- Mobile cellular subscriptions (per 100 people): The number of mobile cellular subscriptions per 100 people, providing a measure of mobile phone penetration.
- Fixed broadband subscriptions (per 100 people): The number of fixed broadband subscriptions per 100 people, indicating the level of internet connectivity.
- Secure Internet servers (per 1 million people): The number of secure internet servers per 1 million people, reflecting the security infrastructure of the internet in each country.
- Individuals using the Internet (% of population): The percentage of the population that uses the internet, indicating the level of internet adoption and usage.

These indicators are essential for understanding the development and impact of ICT infrastructure across different countries during the selected period.

To analyze the data with different income group, we extract data from The World Bank Group[6], which categorizes world economies into four income groups based on their Gross National Income (GNI) per capita. The data is collected in the year 2020, and the income thresholds for each category are as follows:

- **Low income:** GNI per capita ≤ $1,045
- **Lower middle income:** GNI per capita between $1,046 and $4,095
- **Upper middle income:** GNI per capita between $4,096 and $12,695
- **High income:** GNI per capita > $12,695

[5] https://databank.worldbank.org/home.aspx

[6] World Bank Group country classifications by income level for FY24 (July 1, 2023- June 30, 2024) [WWW Document], n.d. . World Bank Blogs. URL https://blogs.worldbank.org/en/opendata/new-world-bank-group-country-classifications-income-level-fy24 (accessed 6.27.24).

5.3 Addressing Multicollinearity

Given that ICT variables are highly correlated with each other, multicollinearity can become a significant issue in our analysis. To mitigate this problem, we include each ICT variable in a separate model. This approach allows us to isolate the effect of each ICT variable on trade flows while reducing the multicollinearity that could otherwise distort our results. By doing so, we ensure that our estimates are more reliable and interpretable.

5.4 Analyzing Heterogeneity Across Income Groups

To explore the heterogeneity across different income groups, the countries are categorized into four income groups: Low Income (L), Lower Middle Income (LM), Upper Middle Income (UM), and High Income (H). Separate fixed effects models are estimated for each income group to examine the differences in how various factors, including ICT variables, influence trade flows within each group. This approach allows us to capture the specific characteristics and dynamics of trade within different income categories.

5.5 Model selection

To fit the data accurately, it is essential to choose the correct type of gravity model. Two common types of gravity models are Ordinary Least Squares (OLS) and Poisson Pseudo-Maximum Likelihood (PPML).

OLS is the most traditional method that predicts values using the ordinary least squares method. However, it has two significant shortcomings. First, OLS cannot handle zero trade flows because it involves taking the logarithm of the trade flow, which is undefined for zero values. Second, OLS assumes homoscedasticity, meaning it assumes that the variance of the error terms is constant across observations. This assumption is often violated in trade data, leading to inefficient and biased estimates.

To overcome these limitations, the PPML estimator is employed. The PPML model can handle zero trade flows naturally without needing to log-transform the dependent variable, making it suitable for datasets with many zero trade observations. Furthermore, the PPML estimator is robust to heteroscedasticity, providing more reliable and consistent estimates when the variance of the trade flows is not constant.

In addition to choosing PPML over OLS, the fixed effects model is utilized to control unobserved heterogeneity. By including fixed effects for both the exporting and importing countries as well as for time, the model accounts for country-specific and time-specific factors that could influence trade flows. This helps to control omitted variable bias, ensuring that the estimated effects of the explanatory variables are not confounded by unobserved factors.

The fixed effects PPML gravity model was chosen for this analysis due to its ability to handle zero trade flows, its robustness to heteroscedasticity, and its capability to control unobserved heterogeneity. This combination makes it a powerful and reliable tool for analyzing trade flows and the impact of various factors, including ICT variables, on trade.

5.6 Fixed effect PPML Gravity Model

As mentioned above, this thesis will use a fixed effect PPML gravity model to analyze the data. The form of the fixed effect PPML gravity model is as follows:

$E(Y_{ij,t}|X_{ij,t})=\exp(\beta_0+\beta_1 X_{1,ij,t}+\beta_2 X_{2,ij,t}+...+\beta_k X_{k,ij,t}+ \gamma_i + \delta_j +\lambda_t)$ (6)

Where:

- $Y_{ij.t}$ is the trade flow from country i to country j.
- $X_{k,ij,t}$ are the explanatory variables (e.g., GDP, distance, common language, etc.).
- β_k are the coefficients to be estimated.
- γ_i, δ_j, and λ_t are the fixed effects for origin country, destination country, and year, respectively.

5.6.1 Coefficient Interpretation

- for continue variables: a 1% change in X leads to a β /100 units change in trade flow

- for dummy variables: If the variable is present (i.e., it takes the value 1), it leads to a change of exp(β)-1 unit in trade flow compares to that the variable is not present.

5.7 Statistical Software

For the empirical analysis in this thesis, we used RStudio, an ideal software for econometric modeling. The models are estimated using the R programming language, with various packages such as fixest, dplyr, and ggplot2 employed for data manipulation, model estimation, and visualization. Specifically, the feglm function from the fixest package was utilized for fixed effect Poisson Pseudo-Maximum Likelihood (PPML) estimation.

6 Preprocessing

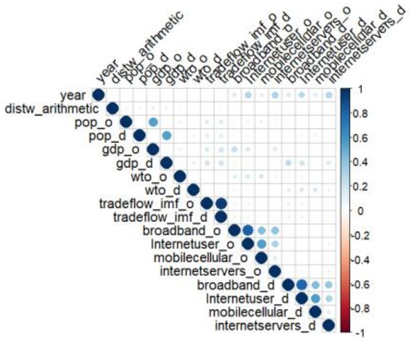
This section delves into international trade patterns using the gravity model framework. We begin by cleaning the data and creating a dataset with the necessary variables. Subsequently, we divide our data into four income groups to facilitate the econometric model analysis. Finally, we visualize the data to gain initial insights.

First, we imported data from three different sources: CEPII, the World Bank (WB), and the World Development Indicators (WDI). After extracting the relevant data, we created a consolidated database, eliminated NA strings, and factorized all dummy variables. Countries were classified into four income groups based on their income levels: Low Income (L), Lower Middle Income (LM), Upper Middle Income (UM), and High Income (H). This classification enabled a detailed examination of how trade flows and other variables differ across income groups. The classification was achieved by merging the main dataset with a country classification dataset and creating a new variable for income categories.

6.1 *Summary of dataset*

The dataset comprises 208,376 observations with 23 variables, capturing various economic and ICT-related metrics over multiple years for different countries. Spanning from 2010 to 2020, this dataset provides annual data for a wide range of countries, with 188 unique country identifiers. The primary focus includes trade flow indicators and ICT variables, offering insights into economic activities and technological adoption.

6.2 *Correlation Analysis*

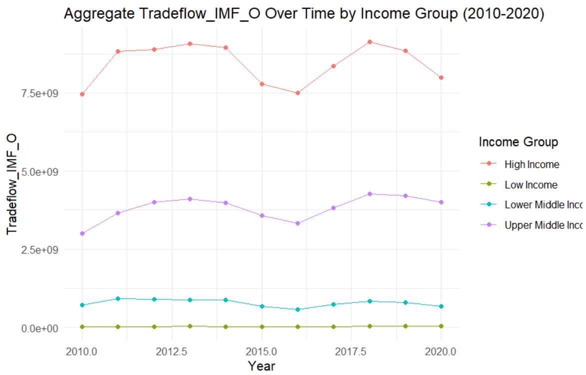


Var1	Var2	Correlatio
gdp_o	pop_o	0.5429
gdp_d	pop_d	0.5452
tradeflow_imf_d	tradeflow_imf_o	0.9477
Internetuser_o	broadband_o	0.8112
mobilecellular_o	Internetuser_o	0.5511
Internetuser_d	broadband_d	0.7997
mobilecellular_d	Internetuser_d	0.555

Figure 1. correlation graph Table 1. correlation table

A correlation matrix was calculated and visualized to understand the relationships between key numerical variables. This analysis helps identify potential multicollinearity issues and the strength of relationships between variables such as GDP, population, trade flow, and ICT metrics. Notable correlations include a strong positive relationship between populations and GDP, as well as between various ICT variables like Internet users and broadband subscriptions.

6.3 Trade Flow Distribution



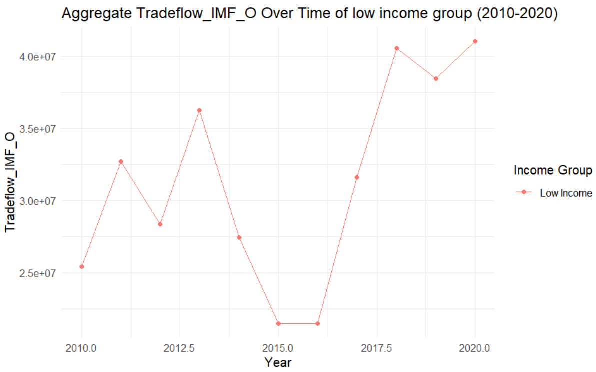


Figure 2. Trade flow across income groups Figure 3. Trade flow of low-income countries

The first plot illustrates the aggregate trade flow (Tradeflow_IMF_O) over time for different income groups from 2010 to 2020. High-income countries show the highest trade flow values, consistently above 7.5e+09, with noticeable fluctuations. The trade flow peaked around 2014-2015 and 2018-2019, followed by slight declines, particularly in 2015-2016, attributed to the global trade slowdown due to lower demand and increased protectionism, as detailed in the World Bank's research[7]. Upper middle-income countries have a stable trade flow, around 3.5e+09, with minor fluctuations, indicating a more consistent but lower level of trade activity compared to high-income countries. Lower middle-income countries show a stable and much lower trade flow, around 1.0e+09, reflecting limited global trade engagement. Low-income countries

[7] Global Trade: Slowdown, Factors, and Policies No. 12, February 2018 Global Knowledge & Research Hub in Malaysia Dorina Georgieva, Norman V. Loayza, and Fabian Mendez-Ramos https://documents1.worldbank.org/curated/en/698861520277541852/pdf/Global-trade-slowdown-factors-and-policies.pdf

exhibit minimal trade flows, often close to zero. However, when included in the main graph, the trade values of different income groups can distort the scale. Therefore, we have created a separate graph specifically for low-income countries.

The second plot focuses on the trade flow for low-income countries alone, revealing more detail on the variability within this group. The trade flow shows a fluctuating trend with peaks and troughs, particularly notable around 2015 and 2017. The sharp decline in 2015 aligns with the global trade slowdown, influenced by lower commodity prices, reduced demand from China, and increased policy uncertainty, as discussed in the World Bank's brief. Despite some recovery post-2016, the overall trade flow remains significantly lower compared to higher-income groups, indicating ongoing challenges and volatility in trade activities for low-income countries. This variability underscores the vulnerability of these countries to global economic changes and the importance of policy interventions to enhance trade resilience.

3. Distribution of ICT Variables by Income Group

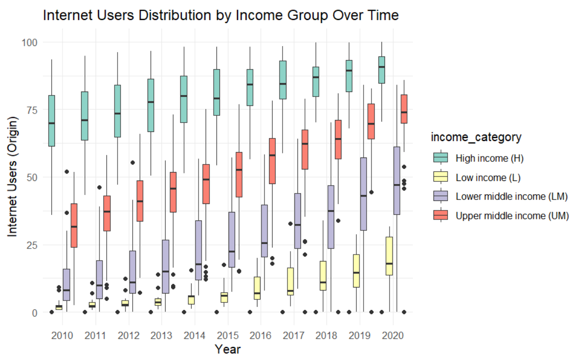


Figure 4. box plot of internet users

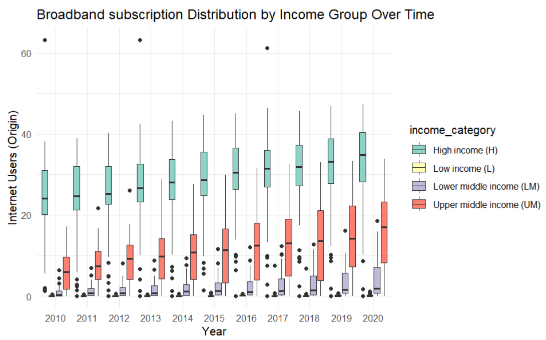


Figure 5. box plot of broadband subscription

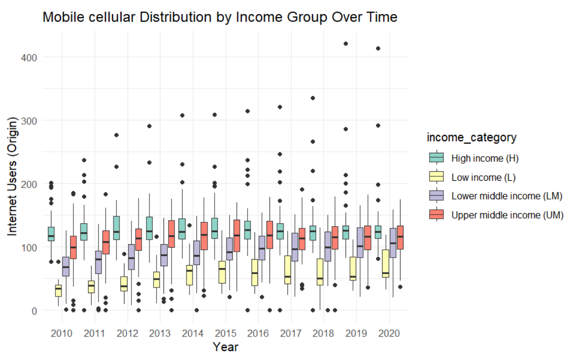


Figure 6. boxplot of mobile cellular

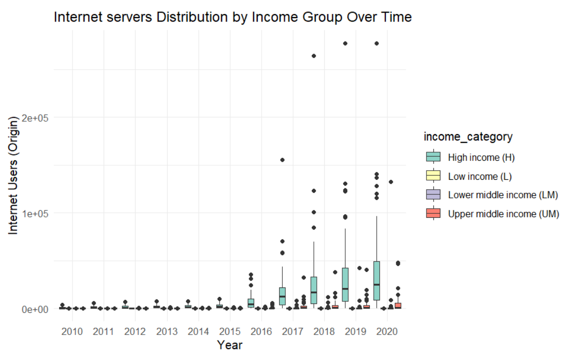


Figure 7. boxplot of internet server

The analysis of various ICT variables, including Internet users, broadband subscriptions, mobile cellular usage, and internet servers, across different income groups over time reveals clear patterns and disparities. The box plots indicate that higher income countries exhibit better ICT development, as evidenced by consistently higher values in all variables. Notably, in higher income countries, some outliers exist, especially in mobile cellular usage and internet server variables, suggesting exceptional cases of high adoption. Across all income groups, the variability in the number of Internet users is relatively similar, with a gradual increase from 2010 to 2020.

For broadband subscriptions, upper middle-income countries show greater variability, and low-income countries exhibit very low subscription rates. This suggests that broadband may not be a significant factor in econometric models for low-income countries due to minimal adoption. In the mobile cellular variable, outliers are present in each income group, but high-income countries show less variability, indicating more uniform adoption rates.

Lastly, in the internet server variable, only high and upper middle-income countries show substantial values, with low and lower middle-income countries having negligible counts. The number of internet servers has markedly increased since 2015 in high-income countries, highlighting a significant development phase. The low presence of internet servers in lower income groups indicates that this variable might not be significant in econometric models focusing on these countries due to insufficient data.

Overall, the data highlights the digital divide, with higher income countries maintaining superior and more consistent ICT infrastructure, while lower income countries lag with significant variability and lower adoption rates.

7 Results

This section presents the results of the econometric model. First, we apply a gravity model to all countries, and then we separately apply the gravity model to each income group. Finally, we conduct a robustness check for the gravity model of all countries to ensure the accuracy of our results.

7.1 Estimates Across the Complete Sample of Countries

This subsection provides an interpretation of the coefficients from the PPML gravity model for all the countries, highlighting the impact of various factors on trade flows. The analysis includes traditional gravity model variables (GDP, distance, contiguity, common language) as well as ICT-related variables (internet servers, mobile cellular subscriptions, broadband subscriptions, internet users).

Interpretation of Coefficients

We will interpret the content presented in Table 2, which shows the results of the PPML fixed effect model for all countries. The interpretation method, as outlined in section 5.6.1, explains that for continuous variables, a 1% change in X leads to a β/100 unit change in trade flow. For dummy variables, if the variable is present (i.e., it takes the value 1), it leads to a change of exp (β)−1 units in trade flow compared to when the variable is not present.

- **GDP of Origin (gdp_o) and GDP of Destination (gdp_d):**
  gdp_o ranges from 3.95E−11 to 4.72E-11
  gdp_d ranges from 3.94E-11 to 4.03E-11
  Both coefficients are positive and highly significant, suggesting that an increase in GDP of either the exporting or importing country is associated with higher trade flows.
  gdp_o: A coefficient of 4.38E−11means that a 1% increase in the GDP of the origin country results in a 4.38E−13 unit increase in trade flow.
  gdp_d: A coefficient of 3.98E−11 means that a 1% increase in the GDP of the destination country results in a 3.98E−13 unit increase in trade flow
- **Distance (distw_arith)**: The negative and highly significant coefficient suggests that greater distance between countries is associated with lower trade flows. This

aligns with the intuition that transportation costs and other frictions increase with distance. A coefficient of -1.57E-4 means that a 1% increase in distance results in a -1.57E-6 unit decrease in trade flow.

- **Contiguity (contig)**: The positive and highly significant coefficient indicates that countries sharing a common border trade significantly more with each other. A coefficient of 1.04 means that countries sharing a border have trade flows that are exp (1.04)−1≈1.83 units higher than those that do not share a border.
- **Common Language (comlang_eth)**: The positive and significant coefficient suggests that countries sharing a common language have higher trade flows, facilitating communication and reducing transaction costs. A coefficient of 0.22 means that countries sharing a common language have trade flows that are exp (0.22)−1≈0.25 unit higher than those that do not share a common language.
- **Internet servers** (internetservers_o, internetservers_d) : coefficients for internet servers are not statistically significant, suggesting that the number of internet servers in either the origin or destination country does not have a significant impact on trade flows.
- **Other ICT variables including** Broadband Subscriptions (broadband_o, broadband_d), Mobile Cellular Subscriptions (mobilecellular_o, mobilecellular_d) and Internet Users (internetusers_o, internetusers_d) are all not statistically significant.

The base model and models including ICT variables indicate that traditional gravity model variables (GDP, distance, contiguity, common language) are significant predictors of trade flows. However, the ICT variables included in these models do not show significant effects, suggesting that their influence on trade flows might be more complex or require different specifications or data to capture their impact.

Model Fit Statistics

• **Log-Likelihood**: Similar across models, indicating comparable fit.

• **Adjusted Pseudo R^2**: Very high (approximately 0.908), indicating the models explain a large portion of the variability in trade flows.

• **BIC**: Consistent across models. (A lower BIC value indicates a better fitting model)

• **Squared Correlation**: High values (approximately 0.858), suggesting a strong correlation between observed and predicted trade flows.

	Base			Internet Servers			Mobile Cellular			Broadband			Internet Users		
	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value
gdp_o	4.38E-11	7.41E-12	3.61e-09***	4.34E-11	8.23E-12	1.32e-07***	3.95E-11	7.97E-12	7.03e-07***	4.72E-11	1.44E-11	1.08e-03**	4.59E-11	8.65E-12	1.11e-07***
gdp_d	3.98E-11	5.50E-12	4.90e-13***	3.97E-11	5.93E-12	2.06e-11***	3.95E-11	6.14E-12	1.33e-10***	4.03E-11	6.37E-12	2.64e-10***	3.94E-11	5.39E-12	2.54e-13***
distw_arithmetic	-1.57E-04	1.71E-05	<2.2e-16***	-1.57E-04	1.72E-05	<2.2e-16***	-1.57E-04	1.71E-05	<2.2e-16***	-1.57E-04	1.71E-05	<2.2e-16***	-1.57E-04	1.71E-05	<2.2e-16***
contig	1.04	0.1	<2.2e-16***	1.04	0.1	<2.2e-16***	1.04	0.1	<2.2e-16***	1.04	0.1	<2.2e-16***	1.04	0.1	<2.2e-16***
comlang_ethno	0.22	0.09	1.57e-02*	0.22	0.09	1.57e-02*	0.22	0.09	1.57e-02*	0.22	0.09	1.57e-02*	0.22	0.09	1.57e-02*
internetservers_o				3.32E-07	5.89E-07	5.73E-01									
internetservers_d				8.41E-08	2.91E-07	7.73E-01									
mobilecellular_o							1.33E-03	1.30E-03	3.07E-01						
mobilecellular_d							-6.34E-04	5.13E-04	2.16E-01						
broadband_o										-2.70E-03	6.72E-03	6.88E-01			
broadband_d										-6.52E-04	2.49E-03	7.94E-01			
internetusers_o													-2.71E-03	2.85E-03	3.42E-01
internetusers_d													9.39E-04	9.51E-04	3.24E-01
Log-Likelihood	-4.25E+10			-4.25E+10			-4.24E+10			-4.25E+10			-4.24E+10		
Adj. Pseudo R2	0.908243			0.908249			0.908282			0.908249			0.908274		
BIC	8.49E+10			8.49E+10			8.49E+10			8.49E+10			8.49E+10		
Squared Cor.	0.858812			0.858801			0.859167			0.858956			0.858539		
Significance codes: * 0.001, ** 0.01, * 0.05, . 0.1**															

Table 2. result for all the countries

7.2 Estimates for Country Groups by Income Level

7.2.1 Results for low-income countries

The PPML gravity model results for lower-income countries (shows in table 3) reveal some challenges with the model's robustness and explanatory power, as evidenced by the insignificance of most variables. Traditional gravity variables such as contiguity consistently show significant positive impacts on trade flows, aligning with theoretical expectations. However, GDP variables and distance fail to reach statistical significance in most cases, indicating potential issues with model specification or data quality. Among the ICT variables, only origin internet servers and destination internet users show significant effects, suggesting a limited but notable role of ICT in trade. The adjusted pseudo-R-squared values (~0.64) suggest the model explains a considerable portion of the variance in trade flows, but the lack of significance across many variables highlights the need for improved data quality and possibly alternative model specifications. This underscores the methodological challenge, particularly in lower-income groups, where data scarcity and quality issues may hinder robust econometric analysis.

7.2.2 Results for lower middle-income countries

The table presents the results of lower middle-income countries (show in table 4). The significant variables in this model are the GDP of the origin country (gdp_o), distance (distw_arithmetic), and common language (comlang_ethno). The GDP of the origin country positively influences trade flows, while distance negatively impacts them, aligning with traditional gravity model hypotheses. However, the contiguity variable is not significant, contrary to expectations. Among the ICT variables, only internet servers in the destination country (internetservers_d) show marginal significance ($p = 0.086$), with a negative coefficient that contradicts our hypothesis that ICT development positively impacts trade flows. However, this variable is generally insignificant across most specifications.

This model partially supports the hypothesis that traditional gravity variables (GDP, distance, common language) significantly affect trade flows. However, ICT variables do not show a strong influence, indicating limited impact on trade flows within this income group. The highly adjusted pseudo-R-squared values (~0.798) and consistent BIC values suggest that the model fits well overall.

The lack of significance for most ICT variables might be due to data scarcity or the relatively lower levels of ICT infrastructure development in lower middle-income countries, limiting their impact on trade flows.

7.2.3 Results for upper middle-income countries

The PPML gravity model for upper middle-income countries (shows in table 5) shows a strong overall fit, with high adjusted pseudo R-squared values (~0.929) and consistent BIC values. This suggests that the model explains the variation in trade flows well for this income group. Significant variables include GDP of both the origin (gdp_o) and destination (gdp_d) countries, distance (distw_arithmetic), contiguity (contig), common language (comlang_ethno), and certain ICT variables.
Among the ICT variables, internet servers at the destination (internetservers_d) are significant with an estimate of -1.45E-06 and a p-value of 2.29e-03, although the coefficient is negative. This significant p-value suggests that the presence of internet servers at the destination plays a statistically significant role in trade flows, requiring further investigation to understand the complex dynamics at play. More notably, internet users at the origin (internetusers_o) with an estimate of 1.26E-02 and a p-value of 3.99e-03, and internet users at the destination (internetusers_d) with an estimate of 5.35E-03 and a p-value of 2.56e-03, are both positive and significant. This aligns with the hypothesis that ICT variables positively affect trade, indicating that higher internet usage at both the origin and destination enhances communication, connectivity, and overall trade flows.

7.2.4 Results for high-income countries

The PPML gravity model for high-income countries (shows in table 6) demonstrates several significant results, supporting the hypothesis regarding traditional variables' influence on trade flows. The GDP of both the origin (gdp_o) and destination (gdp_d) are highly significant across all models, indicating a strong positive relationship with trade flows. This aligns with traditional gravity model expectations that higher economic output enhances trade.

Distance (distw_arithmetic) shows a significant negative relationship, which is consistent with the hypothesis that greater distances reduce trade flows. Contiguity (contig) is also highly significant and positive, suggesting that neighboring countries trade more intensively, again supporting the traditional gravity model predictions.

Among the ICT variables, none show a significant positive effect on trade flows. This lack of significance could be due to the already high levels of ICT infrastructure in high-income countries, where additional improvements may not substantially impact trade flows.

The high values of adjusted pseudo-R-squared (approximately 0.921) and squared correlation indicate that the model fits the data well. The significance of traditional variables like GDP and distance, combined with the insignificance of ICT variables, suggests that while economic size and geographic proximity are crucial determinants of trade in high-income countries, the role of ICT infrastructure may be less pronounced due to saturation effects.

This model's results underline the need for further research to explore why ICT variables are not as impactful in high-income countries, potentially focusing on different aspects of digital trade facilitation or more nuanced ICT indicators.

	Base			Internet Servers			Mobile Cellular			Broadband			Internet Users		
	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value
gdp_o	1.76E-08	1.25E-08	1.57E-01	2.17E-08	1.08E-08	4.52e-02*	1.83E-08	1.24E-08	1.38E-01	1.99E-08	1.17E-08	8.92e-02.	1.88E-08	1.23E-08	1.27E-01
gdp_d	-1.36E-12	7.62E-11	9.86E-01	-3.47E-12	6.51E-11	9.57E-01	-1.53E-11	7.53E-11	8.39E-01	-3.77E-11	8.37E-11	6.53E-01	-8.49E-12	7.74E-11	9.13E-01
distw_arithmetic	-6.77E-05	8.23E-05	4.11E-01	-6.52E-05	8.10E-05	4.21E-01	-6.72E-05	8.20E-05	4.13E-01	-6.76E-05	8.22E-05	4.11E-01	-6.73E-05	8.20E-05	4.12E-01
contig	1.53	0.39	9.05e-05***	1.54	0.38	5.86e-05***	1.54	0.39	8.28e-05***	1.52	0.39	1.03e-04***	1.53	0.39	8.72e-05***
comlang_ethno	0.19	0.35	5.93E-01	0.19	0.35	5.86E-01	0.19	0.35	5.88E-01	0.19	0.35	5.89E-01	0.19	0.35	5.90E-01
internetservers_o				1.38E-02	4.41E-03	1.73e-03**									
internetservers_d				-5.88E-06	2.99E-06	4.98e-02*									
mobilecellular_o							2.19E-03	4.15E-03	5.97E-01						
mobilecellular_d							6.98E-03	2.97E-03	1.86e-02*						
broadband_o										7.04E-01	1.80E-01	9.54e-05***			
broadband_d										2.17E-02	1.66E-02	1.91E-01			
internetusers_o													-3.23E-03	1.34E-02	8.09E-01
internetusers_d													8.13E-03	3.97E-03	4.08e-02*
Log-Likelihood	-3.46E+08			-3.43E+08			-3.45E+08			-3.44E+08			-3.46E+08		
Adj. Pseudo R2	0.640429			0.643941			0.641542			0.642591			0.64072		
BIC	6.92E+08			6.86E+08			6.90E+08			6.88E+08			6.92E+08		
Squared Cor.	0.403897			0.417837			0.411825			0.407949			0.406708		
Significance codes: *** 0.001, ** 0.01, * 0.05, . 0.1															

Table 3. result for low income countries

	Base			Internet Servers			Mobile Cellular			Broadband			Internet Users		
	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value
gdp_o	4.47E-10	1.47E-10	2.46e-03**	4.46E-10	1.50E-10	2.96e-03**	4.43E-10	1.42E-10	1.78e-03**	4.49E-10	1.78E-10	1.178e-02*	3.48E-10	1.52E-10	2.25e-02*
gdp_d	1.66E-11	1.53E-11	2.79E-01	1.40E-11	1.36E-11	3.05E-01	1.12E-11	1.87E-11	5.50E-01	2.03E-11	2.21E-11	3.59E-01	1.56E-11	1.56E-11	3.15E-01
distw_arithmetic	-1.57E-04	3.10E-05	4.44e-07***	-1.57E-04	3.10E-05	4.33e-07***	-1.57E-04	3.10E-05	4.48e-07***	-1.57E-04	3.10E-05	4.35e-07***	-1.57E-04	3.10E-05	4.41e-07***
contig	0.54	0.43	2.14E-01	0.54	0.43	2.13E-01	0.54	0.43	2.14E-01	0.54	0.43	2.12E-01	0.54	0.43	2.13E-01
comlang_ethno	0.67	0.16	3.85e-05***	0.67	0.16	3.86e-05***	0.67	0.16	3.88e-05***	0.67	0.16	3.83e-05***	0.67	0.16	3.81e-05***
internetservers_o				-5.51E-06	1.49E-05	7.11E-01									
internetservers_d				7.28E-07	7.16E-07	3.09E-01									
mobilecellular_o							2.27E-03	3.57E-03	5.25E-01						
mobilecellular_d							2.21E-03	1.43E-03	1.23E-01						
broadband_o										-5.41E-04	2.81E-02	9.85E-01			
broadband_d										-3.42E-03	8.97E-03	7.03E-01			
internetusers_o													-7.34E-03	4.28E-03	8.62e-02.
internetusers_d													4.71E-04	2.30E-03	8.38E-01
Log-Likelihood	-5.01E+09			-5.01E+09			-5.00E+09			-5.01E+09			-5.00E+09		
Adj. Pseudo R2	0.797589			0.797625			0.797816			0.797605			0.797813		
BIC	1.00E+10			1.00E+10			1.00E+10			1.00E+10			1.00E+10		
Squared Cor.	0.481466			0.482662			0.480466			0.482134			0.481893		
Significance codes: *** 0.001, ** 0.01, * 0.05, . 0.1															

Table 4. result for lower middle-income countries

	Base			Internet Servers			Mobile Cellular			Broadband			Internet Users		
	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value
gdp_o	5.02E-11	9.41E-12	9.28e-08***	5.34E-11	1.05E-11	3.66e-07***	4.45E-11	1.08E-11	3.51e-05***	3.81E-11	1.20E-11	1.53e-03**	5.81E-11	8.72E-12	2.60e-11***
gdp_d	5.43E-11	1.41E-11	1.11e-04***	6.93E-11	1.33E-11	1.84e-07***	5.39E-11	1.30E-11	3.24e-05***	5.54E-11	1.33E-11	3.15e-05***	5.32E-11	1.27E-11	2.57e-05***
distw_arithmetic	-1.10E-04	2.86E-05	1.14e-04***	-1.10E-04	2.86E-05	1.14e-04***	-1.10E-04	2.85E-05	1.13e-04***	-1.10E-04	2.86E-05	1.14e-04***	-1.10E-04	2.85E-05	1.11e-04***
contig	1.01	0.14	1.29e-13***	1.01	0.14	1.30e-13***	1.01	0.14	1.28e-13***	1.01	0.14	1.32e-13***	1.01	0.14	1.12e-13***
comlang_ethno	0.66	0.24	5.73e-03**	0.66	0.24	5.71e-03**	0.66	0.24	5.71e-03**	0.66	0.24	5.67e-03**	0.66	0.24	5.67e-03**
internetservers_o				2.66E-06	4.94E-06	5.90E-01									
internetservers_d				-1.45E-06	4.76E-07	2.29e-03**									
mobilecellular_o							1.04E-03	1.49E-03	4.87E-01						
mobilecellular_d							2.57E-04	7.61E-04	7.35E-01						
broadband_o										6.42E-03	6.70E-03	3.38E-01			
broadband_d										-1.25E-03	4.68E-03	7.90E-01			
internetusers_o													1.26E-02	4.37E-03	3.99e-03**
internetusers_d													5.35E-03	1.77E-03	2.56e-03**
Log-Likelihood	-1.06E+10			-1.06E+10			-1.06E+10			-1.06E+10			-1.05E+10		
Adj. Pseudo R2	0.928692			0.928807			0.928707			0.928704			0.929038		
BIC	2.12E+10			2.11E+10			2.11E+10			2.12E+10			2.11E+10		
Squared Cor.	0.927496			0.927955			0.927133			0.927222			0.92759		
Significance codes: * 0.001, ** 0.01, * 0.05, . 0.1**															

Table 5. result for upper middle-income countries

	Base			Internet Servers			Mobile Cellular			Broadband			Internet Users		
	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value
gdp_o	2.70E-11	8.28E-12	1.13e-03**	1.73E-11	7.30E-12	1.78e-02*	2.59E-11	8.84E-12	3.43e-03**	2.36E-11	7.90E-12	2.78e-03**	3.35E-11	9.69E-12	5.46e-04***
gdp_d	3.42E-11	6.06E-12	1.70e-08***	3.35E-11	6.23E-12	7.21e-08***	3.42E-11	7.77E-12	1.06e-05***	3.26E-11	8.17E-12	6.69e-05***	3.53E-11	6.62E-12	9.49e-08***
distw_arithmetic	-1.80E-04	1.57E-05	<2.2e-16***	-1.80E-04	1.57E-05	<2.2e-16***	-1.80E-04	1.57E-05	<2.2e-16***	-1.80E-04	1.57E-05	<2.2e-16***	-1.80E-04	1.57E-05	<2.2e-16***
contig	1.28	0.11	<2.2e-16***	1.28	0.11	<2.2e-16***	1.28	0.11	<2.2e-16***	1.28	0.11	<2.2e-16***	1.28	0.11	<2.2e-16***
comlang_ethno	0.01	0.08	8.80E-01	0.01	0.08	8.80E-01	0.01	0.08	8.81E-01	0.01	0.08	8.81E-01	0.01	0.08	8.80E-01
internetservers_o				8.89E-07	5.80E-07	1.25E-01									
internetservers_d				4.59E-07	3.09E-07	1.38E-01									
mobilecellular_o							1.55E-03	1.91E-03	4.18E-01						
mobilecellular_d							-1.04E-03	6.94E-04	1.32E-01						
broadband_o										-1.25E-02	1.03E-02	2.24E-01			
broadband_d										3.69E-04	2.90E-03	8.99E-01			
internetusers_o													-6.15E-03	4.11E-03	1.34E-01
internetusers_d													-6.48E-04	8.68E-04	4.55E-01
Log-Likelihood	-2.04E+10			-2.04E+10			-2.04E+10			-2.04E+10			-2.04E+10		
Adj. Pseudo R2	0.921134			0.921191			0.921203			0.921242			0.921271		
BIC	4.08E+10			4.07E+10			4.07E+10			4.07E+10			4.07E+10		
Squared Cor.	0.922021			0.921435			0.922104			0.922222			0.921636		
Significance codes: *** 0.001, ** 0.01, * 0.05, . 0.1															

Table 6. result for high-income countries

7.3 Robustness check

In this section, we conduct a robustness check on the original model, which includes all countries, to ensure the reliability and stability of the results under different model specifications. This allows us to verify whether the observed relationships hold true when the model is modified. Specifically, we perform two robustness tests: the first one examines the interaction between ICT variables at the origin and destination to see if simultaneous investment in ICT by both countries influences trade flow. The second test adds quadratic terms for the ICT variables to detect any non-linear effects.

7.3.1 Results of interaction

Based on the robustness checks conducted for the interaction between ICT variables at the origin and destination (shown in Table 7), we observe that all the individual ICT variables are either negative or insignificant, contrary to our hypothesis. In contrast, all traditional variables are significant and align with our hypothesis.

For the interaction term between the origin and destination countries' ICT variables, the results show that broadband and internet users have a positive and significant effect on trade flows when both trade partners invest in them.

These results consistently indicate that while individual ICT variables show no significant influence on trade flows, simultaneous investments in ICT by both trade partners have a positive effect.

7.3.2 Quadratic term

To investigate whether there is any undetected non-linear relationship between ICT (show in table 8) variables and trade flow, we added quadratic terms for the ICT variables to our models. The results show that none of the ICT variables, including their quadratic terms, are significant. Traditional variables such as GDP and distance continue to be significant across all models (MC, BB, IU, IS).

This robust analysis confirms that our results are consistent: ICT variables do not significantly affect trade flows, whereas traditional variables such as GDP and distance play a critical role.

Variable	Coefficient	Std. Error	z value	p-value	Significance	Coefficient	Std. Error	z value	p-value	Significance	Coefficient	Std. Error	z value	p-value	Significance	Coefficient	Std. Error	z value	p-value	Significance
	Model MC					Model BB					Model IU					Model IS				
gdp_o	4.13E-11	8.16E-12	5.054632	4.31E-07	***	4.73E-11	1.47E-11	3.22571	1.26E-03	**	4.81E-11	1.01E-11	4.74195	2.12E-06	***	4.39E-11	7.92E-12	5.536045	3.09E-08	***
gdp_d	4.13E-11	6.90E-12	5.981113	2.22E-09	***	4.04E-11	6.75E-12	5.98455	2.17E-09	***	4.01E-11	5.45E-12	7.36028	1.84E-13	***	4.02E-11	5.96E-12	6.744043	1.54E-11	***
distw_arithmetic	-1.57E-04	1.71E-05	-9.15094	< 2.2e-16	***	-1.56E-04	1.67E-05	-9.30453	< 2.2e-16	***	-1.57E-04	1.69E-05	-9.29755	< 2.2e-16	***	-1.57E-04	1.71E-05	-9.16051	< 2.2e-16	***
contig1	1.043034	0.101207	10.30598	< 2.2e-16	***	1.025227	0.101332	10.11756	< 2.2e-16	***	1.021117	0.100073	10.20374	< 2.2e-16	***	1.043415	0.099703	10.46522	< 2.2e-16	***
comlang_ethno1	0.231469	0.090241	2.565021	1.03E-02	*	0.219678	0.092283	2.38048	1.73E-02	*	0.232826	0.092267	2.5234	1.16E-02	*	0.222025	0.09115	2.435824	1.49E-02	*
mobilecellular_o	-0.00211	0.003465	-0.60926	5.42E-01																
mobilecellular_d	-0.0042	0.002585	-1.62384	1.04E-01																
mobilecellular_o:mobilecellular_o	2.71E-05	2.02E-05	1.341891	1.80E-01																
broadband_o						-0.01235	0.007538	-1.63881	1.01E-01											
broadband_d						-0.01095	0.003673	-2.98241	2.86E-03	**										
broadband_o:broadband_o						3.96E-04	1.29E-04	3.07827	2.08E-03	**										
Internetuser_o											-0.00825	0.004049	-2.03732	4.16E-02	*					
Internetuser_d											-0.00438	0.002952	-1.48409	1.38E-01						
Internetuser_o:Internetuser_o											9.48E-05	4.78E-05	1.98044	4.77E-02	*					
internetservers_o																2.37E-08	6.57E-07	0.03602	9.71E-01	
internetservers_d																-1.96E-07	3.52E-07	-0.55766	5.77E-01	
internetservers_o:internetservers_d																8.35E-12	9.65E-12	0.865228	3.87E-01	
Log-Likelihood	-4.24E+10					-4.22E+10					-4.23E+10					-4.25E+10				
Adj. Pseudo R2	0.908406					0.908745					0.908563					0.908263				
BIC	8.48E+10					8.45E+10					8.46E+10					8.49E+10				
Squared Cor.	0.858928					0.857797					0.854944					0.857974				

Table 7. result for all the countries with interactions

Variable	Coefficient	Std. Error	z value	p-value	Significance	Coefficient	Std. Error	z value	p-value	Significance	Coefficient	Std. Error	z value	p-value	Significance	Coefficient	Std. Error	z value	p-value	Significance
	Model MC					Model BB					Model IU					Model IS				
gdp_o	4.21E-11	1.02E-11	4.111347	3.93E-05	***	4.65E-11	1.35E-11	3.452714	5.55E-04	***	4.48E-11	6.83E-12	6.552475	5.66E-11	***	4.41E-11	9.17E-12	4.806221	1.54E-06	***
gdp_d	3.57E-11	5.91E-12	6.038617	1.55E-09	***	4.08E-11	7.02E-12	5.814522	6.08E-09	***	3.95E-11	5.39E-12	7.325409	2.38E-13	***	3.97E-11	5.93E-12	6.693541	2.18E-11	***
distw_arithmetic	-1.57E-04	1.72E-05	-9.15487	< 2.2e-16	***	-1.57E-04	1.71E-05	-9.16126	< 2.2e-16	***	-1.57E-04	1.71E-05	-9.15996	< 2.2e-16	***	-1.57E-04	1.72E-05	-9.15701	< 2.2e-16	***
contig1	1.044133	0.099701	10.47266	< 2.2e-16	***	1.044012	0.099753	10.466	< 2.2e-16	***	1.043835	0.099925	10.44615	< 2.2e-16	***	1.04402	0.099783	10.46289	< 2.2e-16	***
comlang_ethno1	0.220097	0.090996	2.418761	1.56E-02	*	0.220037	0.091053	2.416592	1.57E-02	*	0.220258	0.091095	2.417899	1.56E-02	*	0.219976	0.091074	2.415367	1.57E-02	*
mobilecellular_o	0.000362	0.002213	0.163458	8.70E-01																
mobilecellular_d	0.001515	0.001683	0.900476	3.68E-01																
I(mobilecellular_o^2)	3.21E-06	4.52E-06	0.708864	4.78E-01																
I(mobilecellular_d^2)	-6.82E-06	5.14E-06	-1.32598	1.85E-01																
broadband_o						-0.0157	0.011816	-1.32829	1.84E-01											
broadband_d						0.000213	0.005163	0.041255	9.67E-01											
I(broadband_o^2)						3.46E-04	2.04E-04	1.694331	9.02E-02	.										
I(broadband_d^2)						-2.02E-05	1.09E-04	-0.1858	8.53E-01											
Internetuser_o											0.001875	0.00409	0.458458	6.47E-01						
Internetuser_d											0.002476	0.002609	0.948824	3.43E-01						
I(Internetuser_o^2)											-4.33E-05	3.87E-05	-1.12071	2.62E-01						
I(Internetuser_d^2)											-1.54E-05	1.94E-05	-0.79316	4.28E-01						
internetservers_o																1.17E-06	1.20E-06	0.973354	3.30E-01	
internetservers_d																-1.91E-07	6.44E-07	-0.296	7.67E-01	
I(internetservers_o^2)																-5.45E-12	5.60E-12	-0.97226	3.31E-01	
I(internetservers_d^2)																1.82E-12	2.47E-12	0.735043	4.62E-01	
Log-Likelihood	-4.24E+10					-4.24E+10					-4.24E+10					-4.25E+10				
Adj. Pseudo R2	0.908302					0.908327					0.908304					0.90826				
BIC	8.49E+10					8.48E+10					8.49E+10					8.49E+10				
Squared Cor.	0.859689					0.858329					0.85856					0.858833				

Table 8. result for all the countries with quadratic terms

7.4 key findings

1. After comparing the results across different income groups and the model encompassing all countries, we observe that traditional gravity model variables (GDP, distance, contiguity, common language) are consistently significant and align well with our hypotheses.

2. The impact of ICT variables on trade flows shows mixed results. While internet users in the upper middle-income group are positive and significant, most ICT indicators are not statistically significant across the models for each income group. We cannot conclusively prove the effect of ICT development on trade flows.

3. The analysis across different income groups reveals the inadequacy of the econometric model for the lower-income group due to data scarcity. This data limitation affects the robustness and reliability of the results for this group, highlighting a critical methodological challenge.

4. The high value of the adjusted pseudo-R-squared (approximately 0.921) and squared correlation indicates that the models fit the data well, particularly in the upper middle-income and high-income groups.

5. The robustness check shows that while both trade partners invest in ICT (broadband and internet servers), there may be a weak positive significant effect on trade flows. However, this cannot conclusively prove that ICT is the direct cause, as all the individual effects of ICT on trade flows are either insignificant or negative.

8 Conclusions

The purpose of this study was to analyze the impact of different means of ICT on trade performance and to examine the effects of ICT and other gravity variables across different income groups. The main findings indicate that traditional gravity model variables such as GDP, distance, contiguity, and common language are consistently significant and align well with our hypotheses. Specifically, higher GDP levels, shared borders, and common language positively impact trade flows, while greater distances reduce them.

The impact of ICT variables on trade flows shows mixed results. While internet users in the upper middle-income group have a positive and significant effect, most ICT indicators are not statistically significant across the models for each income group. The analysis reveals that individual ICT variables generally do not significantly affect trade flows. However, when both trade partners invest in ICT simultaneously, particularly in broadband and internet servers, there is a weak but positive significant effect on trade flows.

These results suggest that traditional economic and geographic factors have a more substantial impact on trade flows than ICT investments. The findings imply that while ICT has potential, its direct impact on trade flows is limited. Policymakers should consider focusing on coordinated ICT infrastructure investments to maximize trade benefits.

Considering these findings, future research should explore different aspects of digital trade facilitation and more nuanced ICT indicators to better capture their true influence on trade flows. Addressing data scarcity, especially in lower-income groups, is also crucial for improving the robustness and reliability of econometric models.

In conclusion, while ICT investments show some potential when made collaboratively between trade partners, traditional gravity model variables remain the key determinants of international trade flows.

9 References

- Aker, J.C., Mbiti, I.M., 2010. Mobile Phones and Economic Development in Africa. Journal of Economic Perspectives 24, 207–232. https://doi.org/10.1257/jep.24.3.207
- Anderson, J.E., Van Wincoop, E., 2003. Gravity with Gravitas: A Solution to the Border Puzzle. American Economic Review 93, 170–192. https://doi.org/10.1257/000282803321455214
- CEPII - Data [WWW Document], n.d. URL http://www.cepii.fr/cepii/en/bdd_modele/bdd_modele.asp (accessed 6.27.24).
- Choi, C., 2010. The effect of the Internet on service trade. Economics Letters 109, 102–104. https://doi.org/10.1016/j.econlet.2010.08.005
- Clarke, G.R.G., Wallsten, S.J., 2006. Has the Internet Increased Trade? Developed and Developing Country Evidence. Economic Inquiry 44, 465–484. https://doi.org/10.1093/ei/cbj026
- Correia, S., Guimarães, P., Zylkin, T., 2020. ppmlhdfe: Fast Poisson Estimation with High-Dimensional Fixed Effects. The Stata Journal 20, 95–115. https://doi.org/10.1177/1536867X20909691
- Czernich, N., Falck, O., Kretschmer, T., Woessmann, L., 2011. Broadband Infrastructure and Economic Growth. The Economic Journal 121, 505–532. https://doi.org/10.1111/j.1468-0297.2011.02420.x
- Disdier, A.-C., Head, K.C., 2004. The Puzzling Persistence of the Distance Effect on Bilateral Trade. SSRN Journal. https://doi.org/10.2139/ssrn.665083
- Egger, P., Larch, M., 2011. An assessment of the Europe agreements' effects on bilateral trade, GDP, and welfare. European Economic Review 55, 263–279. https://doi.org/10.1016/j.euroecorev.2010.05.002
- Freund, C., Weinhold, D., 2002. The Internet and International Trade in Services. American Economic Review 92, 236–240. https://doi.org/10.1257/000282802320189320
- Freund, C.L., Weinhold, D., 2004. The effect of the Internet on international trade. Journal of International Economics 62, 171–189. https://doi.org/10.1016/S0022-1996(03)00059-X
- Gravity model of trade, 2024. . Wikipedia.

- Helpman, E., Krugman, P.R., 2005. Market structure and foreign trade: increasing returns, imperfect competition, and the international economy, 9. print. ed. MIT Press, Cambridge, Mass.
- Kitenge, E., Lahiri, S., 2023. Estimating gravity coefficients with multiple layers of heterogeneity. Rev International Economics roie.12721. https://doi.org/10.1111/roie.12721
- Leibenstein, H., Tinbergen, J., 1966. Shaping the World Economy: Suggestions for an International Economic Policy. The Economic Journal 76, 92. https://doi.org/10.2307/2229041
- Melitz, J., 2008. Language and foreign trade. European Economic Review 52, 667–699. https://doi.org/10.1016/j.euroecorev.2007.05.002
- Panel data, 2024. . Wikipedia.
- Pöyhönen, P., 1963. A Tentative Model for the Volume of Trade between Countries. Weltwirtschaftliches Archiv 90, 93–100.
- Raihan, D.S., n.d. The OLS Estimation of a basic gravity model.
- Shepherd, B., n.d. The gravity model of international trade: a user guide (R version).
- Visser, R., 2019. The effect of the internet on the margins of trade. Information Economics and Policy 46, 41–54. https://doi.org/10.1016/j.infoecopol.2018.12.001
- World Bank Group country classifications by income level for FY24 (July 1, 2023- June 30, 2024) [WWW Document], n.d. . World Bank Blogs. URL https://blogs.worldbank.org/en/opendata/new-world-bank-group-country-classifications-income-level-fy24 (accessed 6.27.24).
- World Development Indicators | DataBank [WWW Document], n.d. URL https://databank.worldbank.org/source/world-development-indicators (accessed 6.27.24).
- Xing, Z., 2018. The impacts of Information and Communications Technology (ICT) and E-commerce on bilateral trade flows. Int Econ Econ Policy 15, 565–586. https://doi.org/10.1007/s10368-017-0375-5

10 Anexos

10.1 Country list

Abbreviation1	Country1	Abbreviation2	Country2	Abbreviation3	Country3	Abbreviation4	Country4
ABW	Aruba	DOM	Dominican Re	KWT	Kuwait	PYF	French Polyne
AGO	Angola	DZA	Algeria	LAO	Laos	QAT	Qatar
ALB	Albania	ECU	Ecuador	LBN	Lebanon	ROU	Romania
ARE	United Arab E	EGY	Egypt	LBR	Liberia	RUS	Russia
ARG	Argentina	ERI	Eritrea	LBY	Libya	RWA	Rwanda
ARM	Armenia	ESP	Spain	LCA	St. Lucia	SAU	Saudi Arabia
ATG	Antigua & Bar	EST	Estonia	LKA	Sri Lanka	SEN	Senegal
AUS	Australia	FIN	Finland	LSO	Lesotho	SGP	Singapore
AUT	Austria	FJI	Fiji	LTU	Lithuania	SLB	Solomon Islan
AZE	Azerbaijan	FRA	France	LUX	Luxembourg	SLE	Sierra Leone
BDI	Burundi	FRO	Faroe Islands	LVA	Latvia	SLV	El Salvador
BEL	Belgium	FSM	Micronesia (F	MAC	Macao SAR Cl	SMR	San Marino
BEN	Benin	GAB	Gabon	MAR	Morocco	SOM	Somalia
BFA	Burkina Faso	GBR	United Kingdo	MDA	Moldova	SRB	Serbia
BGD	Bangladesh	GEO	Georgia	MDG	Madagascar	SSD	South Sudan
BGR	Bulgaria	GHA	Ghana	MDV	Maldives	STP	São Tomé & P
BHR	Bahrain	GIN	Guinea	MEX	Mexico	SUR	Suriname
BHS	Bahamas	GMB	Gambia	MHL	Marshall Islan	SVK	Slovakia
BIH	Bosnia & Herz	GNB	Guinea-Bissau	MKD	North Maced	SVN	Slovenia
BLR	Belarus	GNQ	Equatorial Gu	MLI	Mali	SWE	Sweden
BLZ	Belize	GRC	Greece	MLT	Malta	SWZ	Eswatini
BMU	Bermuda	GRD	Grenada	MMR	Myanmar (Bu	SYC	Seychelles
BOL	Bolivia	GRL	Greenland	MNE	Montenegro	SYR	Syria
BRA	Brazil	GTM	Guatemala	MNG	Mongolia	TCD	Chad
BRB	Barbados	GUY	Guyana	MOZ	Mozambique	TGO	Togo
BRN	Brunei	HKG	Hong Kong SA	MRT	Mauritania	THA	Thailand
BTN	Bhutan	HND	Honduras	MUS	Mauritius	TJK	Tajikistan
BWA	Botswana	HRV	Croatia	MWI	Malawi	TKM	Turkmenistan
CAF	Central Africa	HTI	Haiti	NAM	Namibia	TLS	Timor-Leste
CAN	Canada	HUN	Hungary	NCL	New Caledoni	TON	Tonga
CHE	Switzerland	IND	India	NER	Niger	TTO	Trinidad & Tol
CHL	Chile	IRL	Ireland	NGA	Nigeria	TUN	Tunisia
CHN	China	IRN	Iran	NIC	Nicaragua	TUR	Turkey
CIV	Côte d’Ivoire	IRQ	Iraq	NLD	Netherlands	TUV	Tuvalu
CMR	Cameroon	ISL	Iceland	NOR	Norway	TZA	Tanzania
COD	Congo - Kinsh	ISR	Israel	NPL	Nepal	UGA	Uganda
COG	Congo - Brazz	ITA	Italy	NRU	Nauru	UKR	Ukraine
COL	Colombia	JAM	Jamaica	NZL	New Zealand	URY	Uruguay
COM	Comoros	JOR	Jordan	OMN	Oman	USA	United States
CPV	Cape Verde	JPN	Japan	PAN	Panama	UZB	Uzbekistan
CRI	Costa Rica	KAZ	Kazakhstan	PER	Peru	VCT	St. Vincent & (
CUB	Cuba	KEN	Kenya	PHL	Philippines	VEN	Venezuela
CYP	Cyprus	KGZ	Kyrgyzstan	PLW	Palau	VUT	Vanuatu
CZE	Czechia	KHM	Cambodia	PNG	Papua New G	WSM	Samoa
DJI	Djibouti	KIR	Kiribati	POL	Poland	ZAF	South Africa
DMA	Dominica	KNA	St. Kitts & Ne	PRT	Portugal	ZMB	Zambia
DNK	Denmark	KOR	South Korea	PRY	Paraguay	ZWE	Zimbabwe

10.2 Code and statistical results

R Notebook

Preprocessing

```
setwd("C:/Users/rongr/Desktop/TFG/final")
Gravity_V202211 <-
readRDS("C:/Users/rongr/Desktop/TFG/final/Gravity_V202211.rds")

library(sqldf)
## Loading required package: gsubfn
## Loading required package: proto
## Loading required package: RSQLite
library(readxl)
library(dplyr)
##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##     filter, lag
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
library(tidyr)
library(stats)
library(ggplot2)
## Warning: package 'ggplot2' was built under R version 4.3.3
library(corrplot)
## corrplot 0.92 loaded
library(sjPlot)
## Warning: package 'sjPlot' was built under R version 4.3.3
library(gravity)
## Warning: package 'gravity' was built under R version 4.3.3
## Gravity 1.1
## If this package is useful, please support it.
https://www.buymeacoffee.com/pacha.
##
## Attaching package: 'gravity'
## The following object is masked from 'package:stats':
##
##     nls
library(broom)
cepii <- sqldf('
  SELECT
    year,
    country_id_o,
    country_id_d,
    distw_arithmetic,
    contig,
    comlang_ethno,
    pop_o,
    pop_d,
    gdp_o,
    gdp_d,
    wto_o,
```

```
    wto_d,
    rta_coverage,
    CASE
      WHEN rta_coverage IS NULL THEN 0
      ELSE 1
    END AS rta,
    tradeflow_imf_o,
    tradeflow_imf_d
  FROM Gravity_V202211
  WHERE year < 2021 AND year > 2009
')
library(readxl)
wdi <- read_excel("wdigravity.xlsx")
class <- read_excel("groups.xlsx")
## New names:
## • `` -> `...4`
class <- class %>%
  mutate(across(where(is.character), as.factor))
data<-sqldf('
  SELECT
    c.*,
    w_o.broadband AS broadband_o,
    w_o.Internet AS Internetuser_o,
    w_o.mobilecellular AS mobilecellular_o,
    w_o.internetservers AS internetservers_o,
    w_d.broadband AS broadband_d,
    w_d.Internet AS Internetuser_d,
    w_d.mobilecellular AS mobilecellular_d,
    w_d.internetservers AS internetservers_d
  FROM cepii c
  LEFT JOIN wdi w_o
  ON c.year = w_o.years AND c.country_id_o = w_o.country
  LEFT JOIN wdi w_d
  ON c.year = w_d.years AND c.country_id_d = w_d.country
')

#data structure
data<-data[,-(13)]
data$contig <- as.factor(data$contig)
data$comlang_ethno <- as.factor(data$comlang_ethno)
data$rta<-as.factor(data$rta)
data<-na.omit(data)
```

Group classification

```
library(dplyr)
merged_data <- data %>%
  left_join(class, by = c("country_id_o" = "country_code"))
merged_data <- merged_data %>%
  mutate(income_category = case_when(
    income_level == "L"  ~ "Low income (L)",
    income_level == "LM" ~ "Lower middle income (LM)",
    income_level == "UM" ~ "Upper middle income (UM)",
    income_level == "H"  ~ "High income (H)",
  ))
merged_data <- merged_data %>%
```

```r
 filter(!is.na(income_category))
merged_data <- merged_data %>%
  select(-FY22, -`...4`)
```

```r
head(merged_data)
##   year country_id_o country_id_d distw_arithmetic contig
comlang_ethno   pop_o
## 1 2010          ABW          ARE            12772      0
0 101.597
## 2 2011          ABW          ARE            12772      0
0 101.932
## 3 2017          ABW          ARE            12772      0
0 105.361
## 4 2018          ABW          ARE            12772      0
0 105.846
## 5 2019          ABW          ARE            12772      0
0 106.310
## 6 2020          ABW          ARE            12772      0
0 106.766
##      pop_d   gdp_o     gdp_d wto_o wto_d rta tradeflow_imf_o
tradeflow_imf_d
## 1 8441.537 2467704 286049305     0     1   0           0.030
168.407
## 2 8925.096 2584464 347454046     0     1   0         165.430
2.839
## 3 9487.206 3092179 385605507     0     1   0          19.333
538.983
## 4 9630.966 3202235 422215044     0     1   0           0.115
1538.489
## 5 9770.526 3310056 417215560     0     1   0          16.493
889.712
## 6 9890.400 2496648 358868765     0     1   0          30.793
430.352
##   broadband_o Internetuser_o mobilecellular_o internetservers_o
broadband_d
## 1    19.15150          62.00         131.3508          89.69414
9.276577
## 2     0.00000          69.00           0.0000         128.34689
10.110172
## 3     0.00000          97.17           0.0000         976.86814
30.455887
## 4     0.00000           0.00           0.0000        1151.35615
33.090905
## 5    17.85010           0.00         132.4665        1399.82338
33.066331
## 6    17.82615           0.00         132.2888        1585.58897
34.941553
##   Internetuser_d mobilecellular_d internetservers_d income_level
## 1       68.00000         128.8177          92.55143            H
## 2       78.00000         136.7594         131.07558            H
## 3       94.81992         218.6323        1344.46427            H
## 4       98.45000         219.7007        1563.64724            H
## 5       99.15000         212.8044        2028.08246            H
## 6      100.00000         197.8439        1496.77694            H
```

```
##   income_category
## 1 High income (H)
## 2 High income (H)
## 3 High income (H)
## 4 High income (H)
## 5 High income (H)
## 6 High income (H)
dataL <- merged_data %>% filter(income_category == "Low income (L)")
dataLM <- merged_data %>% filter(income_category == "Lower middle
income (LM)")
dataUM <- merged_data %>% filter(income_category == "Upper middle
income (UM)")
dataH <- merged_data %>% filter(income_category == "High income (H)")
head(dataL)
##   year country_id_o country_id_d distw_arithmetic contig
comlang_ethno
## 1 2015          BDI          AGO             1882      0
0
## 2 2018          BDI          AGO             1882      0
0
## 3 2020          BDI          AGO             1879      0
0
## 4 2010          BDI          ARE             4203      0
0
## 5 2011          BDI          ARE             4203      0
0
## 6 2012          BDI          ARE             4203      0
0
##       pop_o     pop_d   gdp_o     gdp_d wto_o wto_d rta
tradeflow_imf_o
## 1 10199.270 27859.304 3097325 102962242     1     1   0
10.008
## 2 11175.379 30809.788 2660124  77792940     1     1   0
24.273
## 3 11890.781 32866.268 2780511  53619071     1     1   0
18.324
## 4  9232.753  8441.537 2026864 286049305     1     1   0
107.333
## 5  9540.362  8925.096 2355652 347454046     1     1   0
54.160
## 6  9849.569  9205.651 2472385 372313981     1     1   0
48.251
##   tradeflow_imf_d broadband_o Internetuser_o mobilecellular_o
internetservers_o
## 1           1.836 0.025831656       2.000000         46.58923
0.7457714
## 2          29.141 0.034236826       2.700000         54.97003
5.8293960
## 3          23.288 0.034614744       3.787954         54.26373
8.4286487
## 4       13709.713 0.003856855       1.000000         18.38612
0.4382791
## 5       19654.275 0.005256071       1.110000         20.24789
0.4230238
```

```
## 6      152599.328 0.007166571       1.220000         22.94044
0.4083516
##   broadband_d Internetuser_d mobilecellular_d internetservers_d
income_level
## 1   0.5459774       22.00000         49.36245           8.07033
L
## 2   0.3503314       29.00000         42.49095          11.03169
L
## 3   0.3635103       32.55015         43.81009          19.74364
L
## 4   9.2765768       68.00000        128.81766          92.55143
L
## 5  10.1101723       78.00000        136.75942         131.07558
L
## 6  11.0212499       84.99999        158.97634         250.08745
L
##   income_category
## 1  Low income (L)
## 2  Low income (L)
## 3  Low income (L)
## 4  Low income (L)
## 5  Low income (L)
## 6  Low income (L)
head(dataLM)
##   year country_id_o country_id_d distw_arithmetic contig
comlang_ethno    pop_o
## 1 2010          AGO          ARE             5921      0
0 19549.12
## 2 2011          AGO          ARE             5921      0
0 20180.49
## 3 2012          AGO          ARE             5921      0
0 20820.52
## 4 2013          AGO          ARE             5921      0
0 21471.62
## 5 2014          AGO          ARE             5921      0
0 26920.47
## 6 2015          AGO          ARE             5917      0
0 27859.30
##      pop_d     gdp_o     gdp_d wto_o wto_d rta tradeflow_imf_o
tradeflow_imf_d
## 1 8441.537  82470896 286049305     1     1   0        366245.2
189747.9
## 2 8925.096 104115872 347454046     1     1   0        382914.8
301582.1
## 3 9205.651 115341558 372313981     1     1   0        656894.0
1234968.9
## 4 9346.129 124178244 402340119     1     1   0        790460.6
757063.8
## 5 9070.867 126776877 403197690     1     1   0        867779.3
812542.2
## 6 9154.302 102962242 357949211     1     1   0        665253.6
1056613.9
##   broadband_o Internetuser_o mobilecellular_o internetservers_o
broadband_d
```

```
## 1  0.06420083            2.8          40.24692           1.284017
9.276577
## 2  0.06520025            4.7          49.76777           1.854973
10.110172
## 3  0.08143466            7.7          50.75814           3.176079
11.021250
## 4  0.08521818           13.0          50.80964           4.551191
11.905303
## 5  0.32346251           21.4          51.80029           5.897892
12.355918
## 6  0.54597740           22.0          49.36245           8.070330
13.843905
##   Internetuser_d mobilecellular_d internetservers_d income_level
## 1       68.00000         128.8177          92.55143           LM
## 2       78.00000         136.7594         131.07558           LM
## 3       84.99999         158.9763         250.08745           LM
## 4       88.00000         183.5446         289.42462           LM
## 5       90.40000         190.3476         382.75450           LM
## 6       90.50000         201.2197         609.40468           LM
##              income_category
## 1 Lower middle income (LM)
## 2 Lower middle income (LM)
## 3 Lower middle income (LM)
## 4 Lower middle income (LM)
## 5 Lower middle income (LM)
## 6 Lower middle income (LM)
head(dataUM)
##   year country_id_o country_id_d distw_arithmetic contig
comlang_ethno    pop_o
## 1 2010          ALB          ARE             3724      0
0 2856.673
## 2 2011          ALB          ARE             3724      0
0 2829.337
## 3 2012          ALB          ARE             3724      0
0 2801.681
## 4 2013          ALB          ARE             3724      0
0 2773.620
## 5 2014          ALB          ARE             3724      0
0 2889.104
## 6 2015          ALB          ARE             3724      0
0 2880.703
##      pop_d    gdp_o     gdp_d wto_o wto_d rta tradeflow_imf_o
tradeflow_imf_d
## 1 8441.537 11926957 286049305     1     1   0       11702.782
826.576
## 2 8925.096 12890867 347454046     1     1   0         844.501
2141.489
## 3 9205.651 12344530 372313981     1     1   0        2780.774
4197.677
## 4 9346.129 12923240 402340119     1     1   0        1613.173
2120.302
## 5 9070.867 13219857 403197690     1     1   0        1543.602
4450.583
## 6 9154.302 11390366 357949211     1     1   0        4592.151
```

```
4163.680
##   broadband_o Internetuser_o mobilecellular_o internetservers_o
broadband_d
## 1    3.622537           45.0          92.4134          4.119435
9.276577
## 2    4.420038           47.0         106.8725          8.261063
10.110172
## 3    5.535177           49.4         121.0154         25.513713
11.021250
## 4    6.323350           51.8         127.6746         35.577453
11.905303
## 5    7.210702           54.3         116.4887         52.265339
12.355918
## 6    8.425731           56.9         117.9871         68.386085
13.843905
##   Internetuser_d mobilecellular_d internetservers_d income_level
## 1       68.00000         128.8177          92.55143           UM
## 2       78.00000         136.7594         131.07558           UM
## 3       84.99999         158.9763         250.08745           UM
## 4       88.00000         183.5446         289.42462           UM
## 5       90.40000         190.3476         382.75450           UM
## 6       90.50000         201.2197         609.40468           UM
##            income_category
## 1 Upper middle income (UM)
## 2 Upper middle income (UM)
## 3 Upper middle income (UM)
## 4 Upper middle income (UM)
## 5 Upper middle income (UM)
## 6 Upper middle income (UM)
head(dataH)
##   year country_id_o country_id_d distw_arithmetic contig
comlang_ethno   pop_o
## 1 2010          ABW          ARE            12772      0
0 101.597
## 2 2011          ABW          ARE            12772      0
0 101.932
## 3 2017          ABW          ARE            12772      0
0 105.361
## 4 2018          ABW          ARE            12772      0
0 105.846
## 5 2019          ABW          ARE            12772      0
0 106.310
## 6 2020          ABW          ARE            12772      0
0 106.766
##      pop_d   gdp_o     gdp_d wto_o wto_d rta tradeflow_imf_o
tradeflow_imf_d
## 1 8441.537 2467704 286049305     0     1   0           0.030
168.407
## 2 8925.096 2584464 347454046     0     1   0         165.430
2.839
## 3 9487.206 3092179 385605507     0     1   0          19.333
538.983
## 4 9630.966 3202235 422215044     0     1   0           0.115
1538.489
```

```
## 5 9770.526 3310056 417215560      0      1    0           16.493
889.712
## 6 9890.400 2496648 358868765      0      1    0           30.793
430.352
##   broadband_o Internetuser_o mobilecellular_o internetservers_o
broadband_d
## 1    19.15150          62.00         131.3508          89.69414
9.276577
## 2     0.00000          69.00           0.0000         128.34689
10.110172
## 3     0.00000          97.17           0.0000         976.86814
30.455887
## 4     0.00000           0.00           0.0000        1151.35615
33.090905
## 5    17.85010           0.00         132.4665        1399.82338
33.066331
## 6    17.82615           0.00         132.2888        1585.58897
34.941553
##   Internetuser_d mobilecellular_d internetservers_d income_level
## 1       68.00000         128.8177          92.55143            H
## 2       78.00000         136.7594         131.07558            H
## 3       94.81992         218.6323        1344.46427            H
## 4       98.45000         219.7007        1563.64724            H
## 5       99.15000         212.8044        2028.08246            H
## 6      100.00000         197.8439        1496.77694            H
##   income_category
## 1 High income (H)
## 2 High income (H)
## 3 High income (H)
## 4 High income (H)
## 5 High income (H)
## 6 High income (H)
```

Descriptive analysis

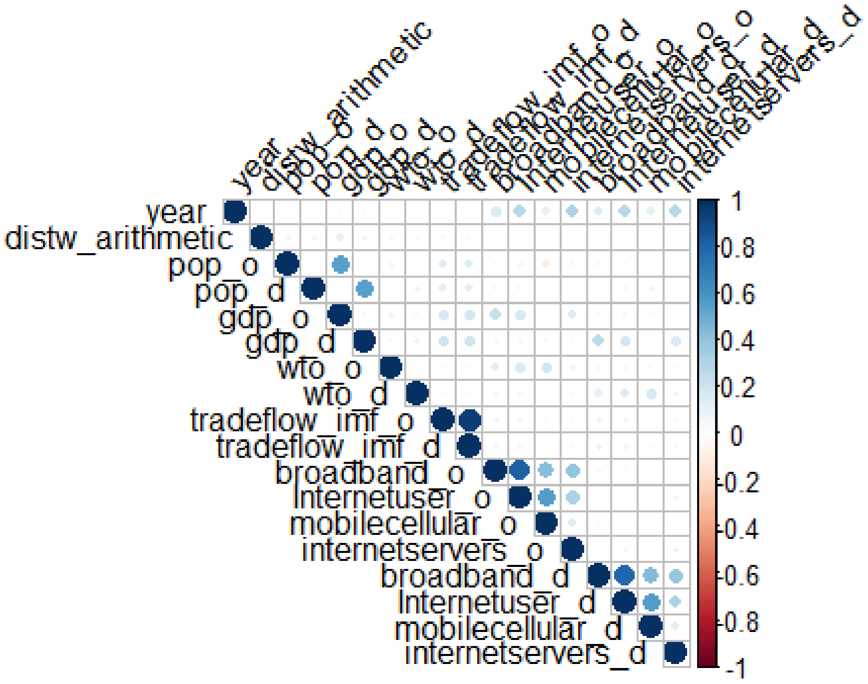
summary

```
summary(data)
##       year      country_id_o       country_id_d
distw_arithmetic
##  Min.   :2010   Length:208376      Length:208376      Min.   :   55
##  1st Qu.:2012   Class :character   Class :character   1st Qu.: 3491
##  Median :2015   Mode  :character   Mode  :character   Median : 6600
##  Mean   :2015                                         Mean   : 7090
##  3rd Qu.:2018                                         3rd Qu.: 9957
##  Max.   :2020                                         Max.   :19680
##  contig     comlang_ethno     pop_o                pop_d
##  0:203106   0:175131      Min.   :      9.8   Min.   :      9.8
##  1:  5270   1: 33245      1st Qu.:   3516.8   1st Qu.:   2951.7
##                           Median :   9897.2   Median :   9762.3
##                           Mean   :  51496.7   Mean   :  49429.3
##                           3rd Qu.:  34005.3   3rd Qu.:  31444.3
##                           Max.   :1411100.0   Max.   :1411100.0
##      gdp_o              gdp_d                wto_o            wto_d
##  Min.   :3.182e+04   Min.   :3.182e+04   Min.   :0.0000   Min.
:0.0000
##  1st Qu.:1.516e+07   1st Qu.:1.275e+07   1st Qu.:1.0000   1st
```

```
Qu.:1.0000
##  Median :6.690e+07   Median :5.621e+07   Median :1.0000   Median
:1.0000
##  Mean   :6.329e+08   Mean   :6.055e+08   Mean   :0.8921   Mean
:0.8825
##  3rd Qu.:3.703e+08   3rd Qu.:3.453e+08   3rd Qu.:1.0000   3rd
Qu.:1.0000
##  Max.   :2.137e+10   Max.   :2.137e+10   Max.   :1.0000   Max.
:1.0000
##  rta        tradeflow_imf_o     tradeflow_imf_d      broadband_o
##  0:158467   Min.   :        0   Min.   :        0   Min.   : 0.000
##  1: 49909   1st Qu.:      176   1st Qu.:      241   1st Qu.: 1.621
##             Median :     4069   Median :     5191   Median :12.210
##             Mean   :   690920   Mean   :   685614   Mean   :15.469
##             3rd Qu.:    66477   3rd Qu.:    74349   3rd Qu.:27.743
##             Max.   :480688672   Max.   :539503424   Max.   :63.202
##  Internetuser_o   mobilecellular_o internetservers_o   broadband_d
##  Min.   :  0.00   Min.   :  0.00   Min.   :     0.0   Min.   :
0.000
##  1st Qu.: 29.00   1st Qu.: 90.86   1st Qu.:    18.3   1st Qu.:
1.159
##  Median : 59.20   Median :113.12   Median :   212.5   Median
:10.569
##  Mean   : 54.22   Mean   :110.45   Mean   :  6211.6   Mean
:14.500
##  3rd Qu.: 79.98   3rd Qu.:130.86   3rd Qu.:  2330.5   3rd
Qu.:26.554
##  Max.   :100.00   Max.   :420.85   Max.   :277330.6   Max.
:63.202
##  Internetuser_d   mobilecellular_d internetservers_d
##  Min.   :  0.00   Min.   :  0.00   Min.   :     0.00
##  1st Qu.: 24.50   1st Qu.: 87.47   1st Qu.:    14.36
##  Median : 55.90   Median :111.62   Median :   168.41
##  Mean   : 52.12   Mean   :109.11   Mean   :  5648.73
##  3rd Qu.: 78.70   3rd Qu.:130.20   3rd Qu.:  1811.95
##  Max.   :100.00   Max.   :420.85   Max.   :277330.58
```

Correlation Analysis

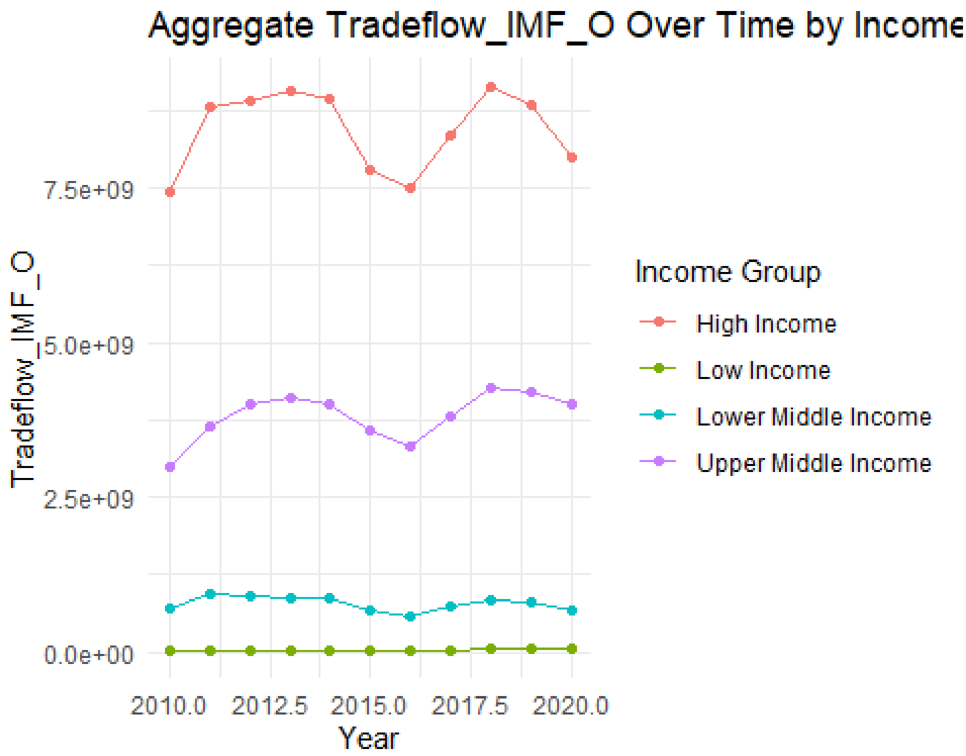
```
library(corrplot)

# Select only the numerical columns from the dataframe
numerical_vars <- data[, sapply(data, is.numeric)]

# Calculate the correlation matrix
correlation_matrix <- cor(numerical_vars, use = "complete.obs")

# Visualize the correlation matrix
corrplot(correlation_matrix, method = "circle", type = "upper", tl.col
= "black", tl.srt = 45)
```




```
# Set a threshold for high correlation
threshold <- 0.5

# Find highly correlated pairs
highly_correlated <- which(abs(correlation_matrix) > threshold,
arr.ind = TRUE)
highly_correlated <- highly_correlated[highly_correlated[, 1] !=
highly_correlated[, 2], ]

# Create a dataframe to list highly correlated variables
highly_correlated_vars <- data.frame(
  Var1 = rownames(correlation_matrix)[highly_correlated[, 1]],
  Var2 = colnames(correlation_matrix)[highly_correlated[, 2]],
  Correlation = correlation_matrix[highly_correlated]
)

# Remove duplicate pairs
highly_correlated_vars <-
highly_correlated_vars[!duplicated(t(apply(highly_correlated_vars, 1,
sort))), ]

# Print the dataframe of highly correlated variables
print(highly_correlated_vars)
##                  Var1            Var2 Correlation
## 1               gdp_o           pop_o   0.5429033
## 2               gdp_d           pop_d   0.5452043
## 5     tradeflow_imf_d tradeflow_imf_o   0.9476544
## 7      Internetuser_o     broadband_o   0.8111873
## 9    mobilecellular_o  Internetuser_o   0.5511160
## 11     Internetuser_d     broadband_d   0.7996659
## 13   mobilecellular_d  Internetuser_d   0.5549758
```

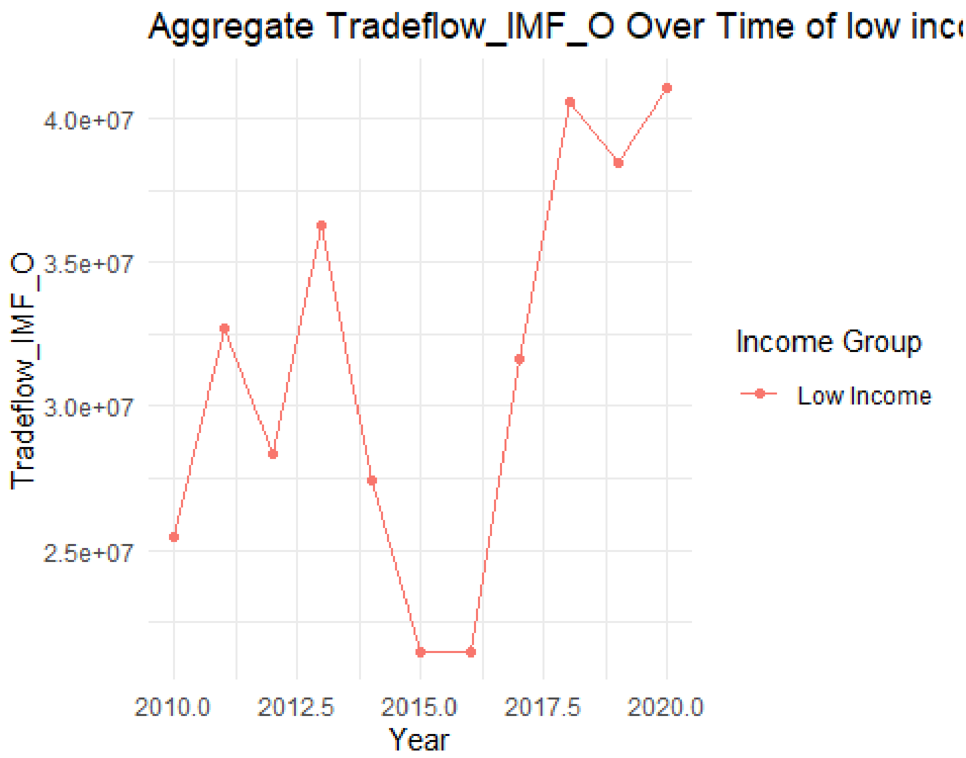
Trade Flow Distribution over time

```
# Visualizations
library(ggplot2)
dataL$income_category <- "Low Income"
dataLM$income_category <- "Lower Middle Income"
dataUM$income_category <- "Upper Middle Income"
dataH$income_category <- "High Income"

# Combine all datasets into one
combined_data <- bind_rows(dataL, dataLM, dataUM, dataH)

# Convert year to numeric if it's not already
combined_data$year <- as.numeric(as.character(combined_data$year))

# Plot the combined data
ggplot(combined_data, aes(x = year, y = tradeflow_imf_o,
                          color = income_category, group =
income_category)) +
  geom_line(stat = "summary", fun = "sum") +
  geom_point(stat = "summary", fun = "sum") +
  labs(title = "Aggregate Tradeflow_IMF_O Over Time by Income Group
(2010-2020)",
       x = "Year",
       y = "Tradeflow_IMF_O",
       color = "Income Group") +
  theme_minimal()
```




```
ggplot(dataL, aes(x = year, y = tradeflow_imf_o, color =
income_category)) +
  geom_line(stat = "summary", fun = "sum") +
  geom_point(stat = "summary", fun = "sum") +
  labs(title = "Aggregate Tradeflow_IMF_O Over Time of low income
group (2010-2020)",
       x = "Year",
```

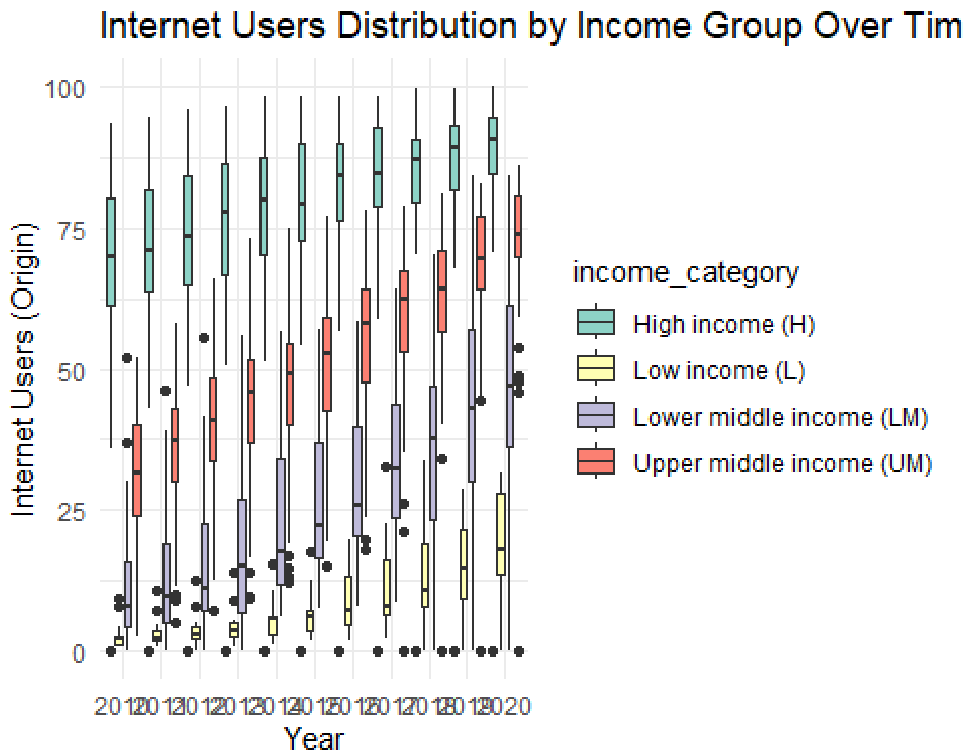
```
        y = "Tradeflow_IMF_O",
        color = "Income Group") +
  theme_minimal()
```



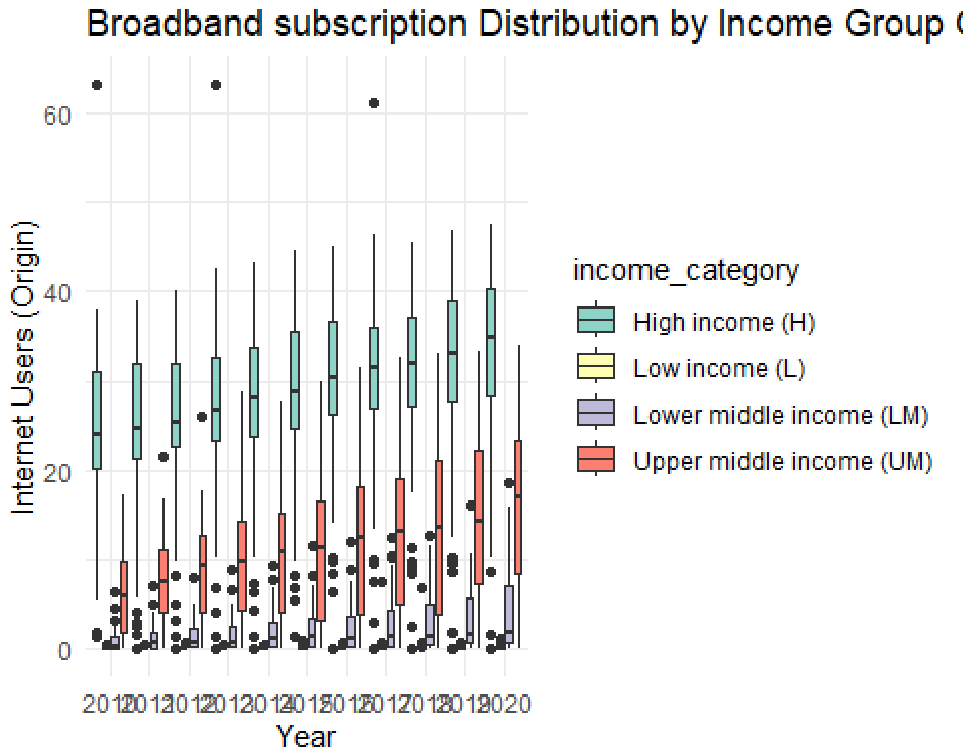
4.Distribution of ICT Variables by Income Group

We analyzed the distribution of various ICT variables, such as Internet users and broadband subscriptions, across different income groups.

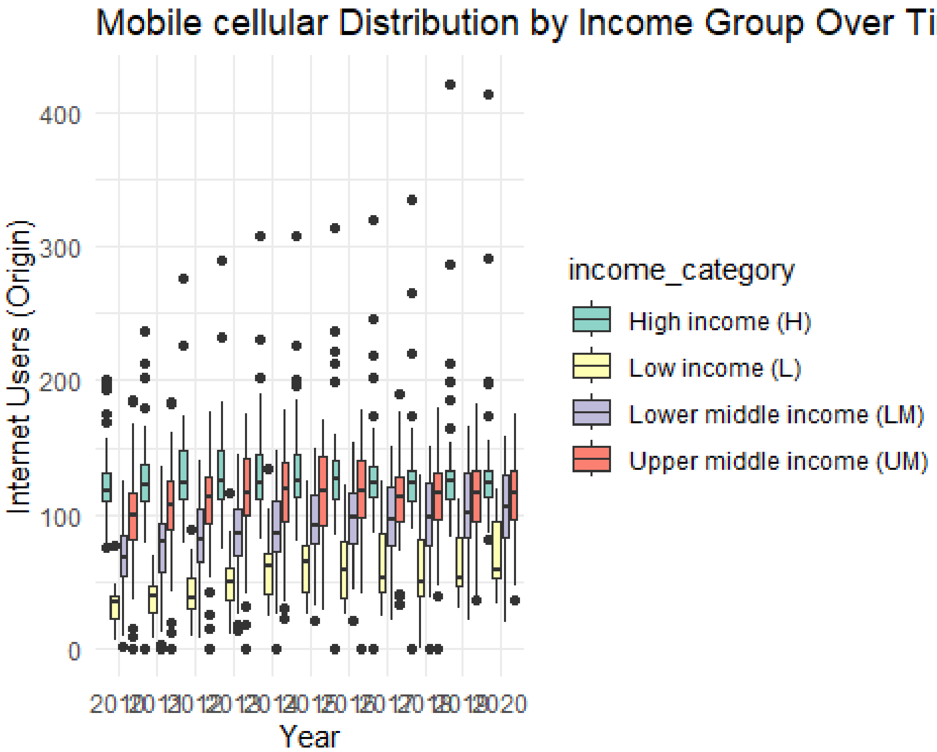
```
library(ggplot2)

# Create the plot
ggplot(merged_data, aes(x = as.factor(year), y = Internetuser_o, fill
= income_category)) +
  geom_boxplot(position = position_dodge(width = 0.9)) +
  labs(title = "Internet Users Distribution by Income Group Over
Time",
        x = "Year",
        y = "Internet Users (Origin)") +
  theme_minimal() +
  scale_fill_brewer(palette = "Set3")
```



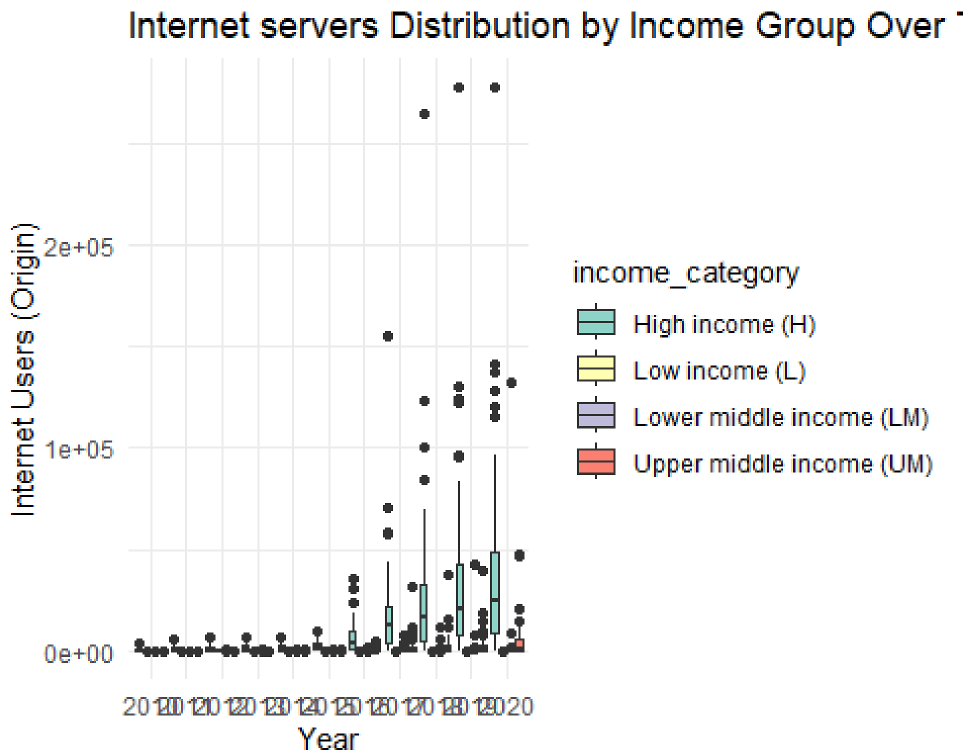

```r
ggplot(merged_data, aes(x = as.factor(year), y = broadband_o, fill =
income_category)) +
  geom_boxplot(position = position_dodge(width = 0.9)) +
  labs(title = "Broadband subscription Distribution by Income Group
Over Time",
       x = "Year",
       y = "Internet Users (Origin)") +
  theme_minimal() +
  scale_fill_brewer(palette = "Set3")
```



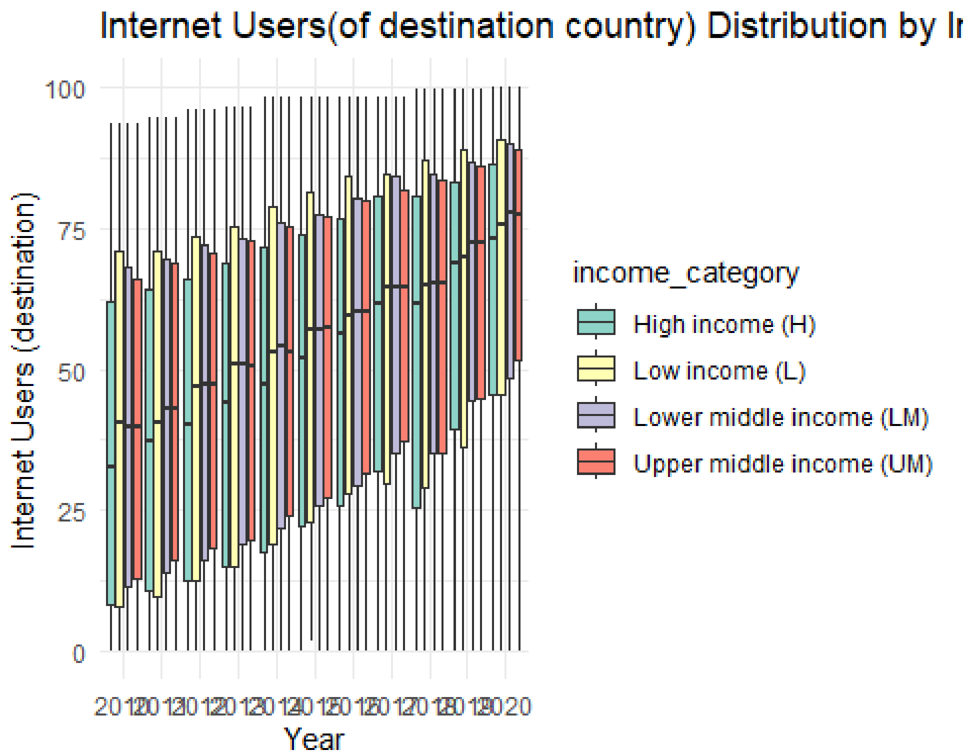
```
ggplot(merged_data, aes(x = as.factor(year), y = mobilecellular_o,
fill = income_category)) +
  geom_boxplot(position = position_dodge(width = 0.9)) +
  labs(title = "Mobile cellular Distribution by Income Group Over
Time",
       x = "Year",
       y = "Internet Users (Origin)") +
  theme_minimal() +
  scale_fill_brewer(palette = "Set3")
```



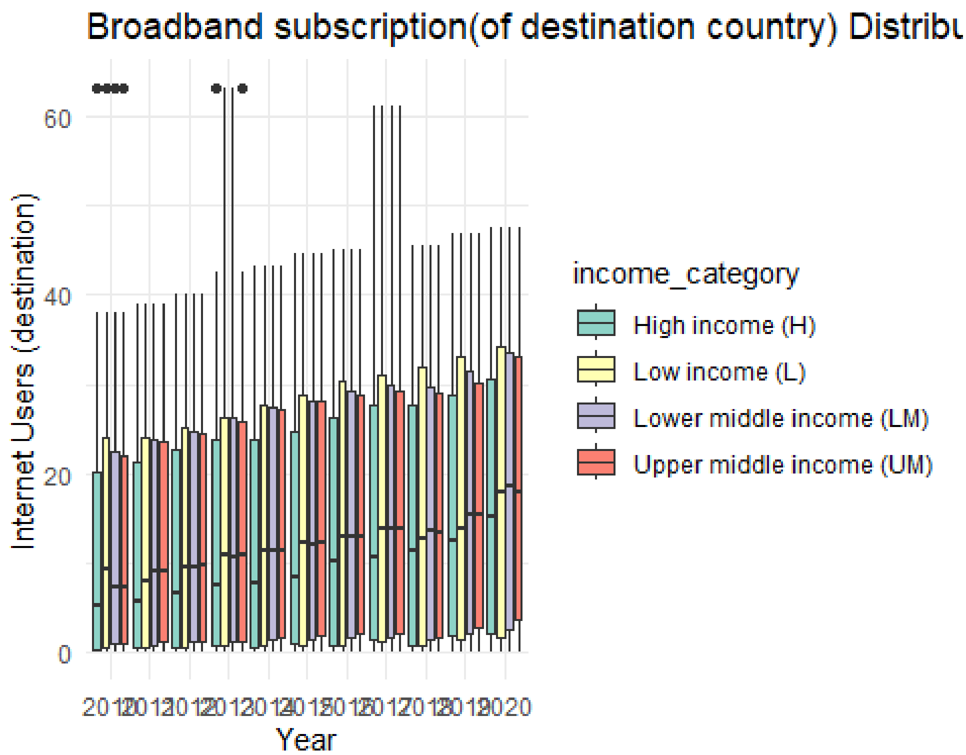

```
ggplot(merged_data, aes(x = as.factor(year), y = internetservers_o,
fill = income_category)) +
  geom_boxplot(position = position_dodge(width = 0.9)) +
  labs(title = "Internet servers Distribution by Income Group Over
Time",
       x = "Year",
       y = "Internet Users (Origin)") +
  theme_minimal() +
  scale_fill_brewer(palette = "Set3")
```



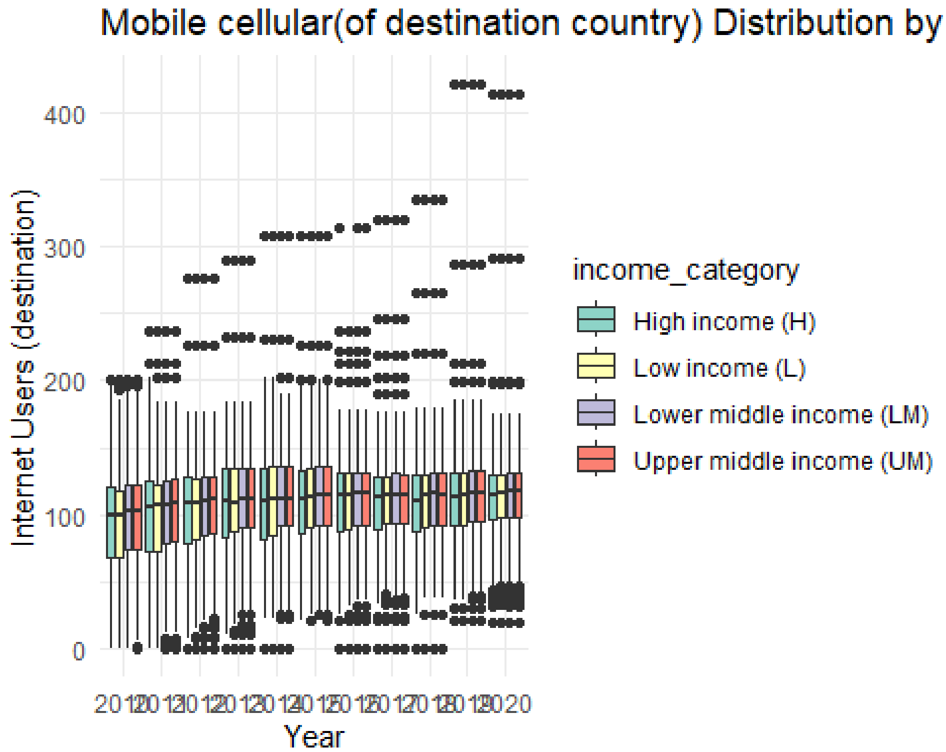

```
# Create the plot
ggplot(merged_data, aes(x = as.factor(year), y = Internetuser_d, fill
= income_category)) +
  geom_boxplot(position = position_dodge(width = 0.9)) +
  labs(title = "Internet Users(of destination country) Distribution by
Income Group Over Time",
       x = "Year",
       y = "Internet Users (destination)") +
  theme_minimal() +
  scale_fill_brewer(palette = "Set3")
```



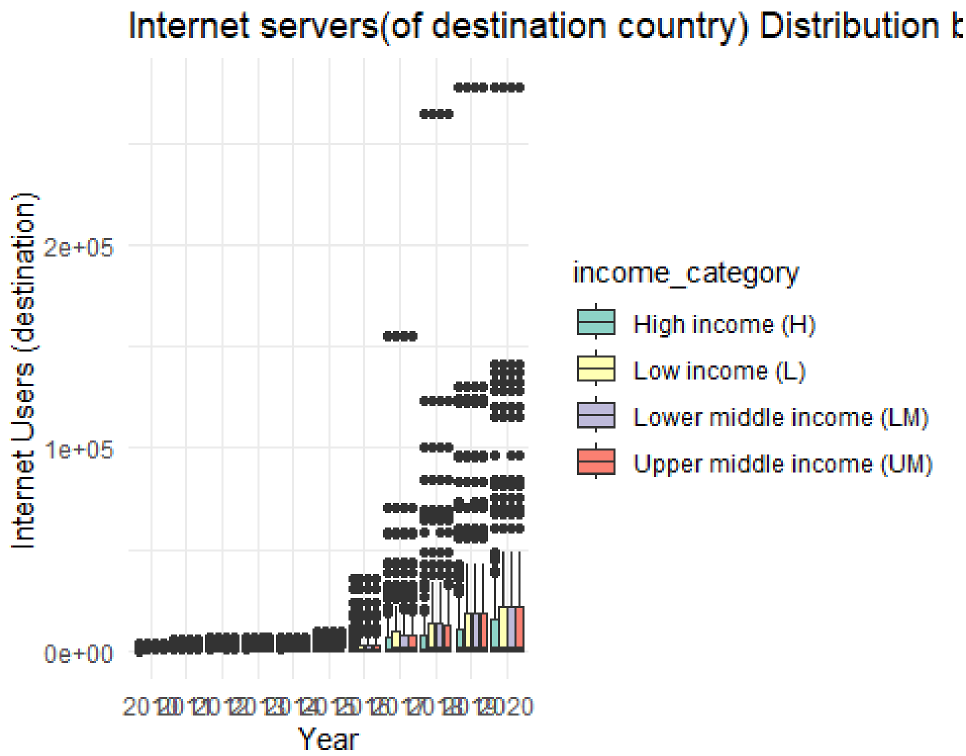

```
ggplot(merged_data, aes(x = as.factor(year), y = broadband_d, fill =
income_category)) +
  geom_boxplot(position = position_dodge(width = 0.9)) +
  labs(title = "Broadband subscription(of destination country)
Distribution by Income Group Over Time",
       x = "Year",
       y = "Internet Users (destination)") +
  theme_minimal() +
  scale_fill_brewer(palette = "Set3")
```



```
ggplot(merged_data, aes(x = as.factor(year), y = mobilecellular_d,
fill = income_category)) +
  geom_boxplot(position = position_dodge(width = 0.9)) +
  labs(title = "Mobile cellular(of destination country) Distribution
by Income Group Over Time",
       x = "Year",
       y = "Internet Users (destination)") +
  theme_minimal() +
  scale_fill_brewer(palette = "Set3")
```




```
ggplot(merged_data, aes(x = as.factor(year), y = internetservers_d,
fill = income_category)) +
  geom_boxplot(position = position_dodge(width = 0.9)) +
  labs(title = "Internet servers(of destination country) Distribution
by Income Group Over Time",
       x = "Year",
       y = "Internet Users (destination)") +
  theme_minimal() +
  scale_fill_brewer(palette = "Set3")
```



Econometric model

PPML model for all

```
library(fixest)
## Warning: package 'fixest' was built under R version 4.3.3
ppml_fixest_model <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno  | country_id_o + country_id_d + year,
  family = poisson(),
  data = data
)
summary(ppml_fixest_model)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 208,376
## Fixed-effects: country_id_o: 188,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                       Estimate   Std. Error  z value   Pr(>|z|)
## gdp_o             4.380000e-11 7.410000e-12  5.90095 3.6142e-09 ***
## gdp_d             3.980000e-11 5.500000e-12  7.22804 4.9001e-13 ***
## distw_arithmetic -1.570920e-04 1.714966e-05 -9.16007  < 2.2e-16 ***
## contig1           1.044032e+00 9.982229e-02 10.45891  < 2.2e-16 ***
## comlang_ethno1    2.200414e-01 9.108074e-02  2.41589 1.5697e-02 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -4.246e+10   Adj. Pseudo R2: 0.908243
##             BIC:  8.492e+10     Squared Cor.: 0.858812
ppml_fixest_model_IS <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic+ contig +
comlang_ethno + internetservers_o + internetservers_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = data
```

```
)
summary(ppml_fixest_model_IS)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 208,376
## Fixed-effects: country_id_o: 188,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                          Estimate   Std. Error   z value   Pr(>|z|)
## gdp_o              4.340000e-11 8.230000e-12  5.275776 1.3220e-07
***
## gdp_d              3.970000e-11 5.930000e-12  6.701630 2.0611e-11
***
## distw_arithmetic  -1.571015e-04 1.715552e-05 -9.157487  < 2.2e-16
***
## contig1            1.044022e+00 9.980435e-02 10.460682  < 2.2e-16
***
## comlang_ethno1     2.200088e-01 9.107837e-02  2.415599 1.5709e-02 *
## internetservers_o  3.321509e-07 5.890943e-07  0.563833 5.7287e-01
## internetservers_d  8.409180e-08 2.912365e-07  0.288741 7.7278e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -4.246e+10   Adj. Pseudo R2: 0.908249
##            BIC:  8.491e+10     Squared Cor.: 0.858801
ppml_fixest_model_MC <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic+ contig +
comlang_ethno  + mobilecellular_o + mobilecellular_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = data
)
summary(ppml_fixest_model_MC)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 208,376
## Fixed-effects: country_id_o: 188,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                         Estimate   Std. Error  z value   Pr(>|z|)
## gdp_o             3.950000e-11 7.970000e-12  4.96052 7.0306e-07 ***
## gdp_d             3.950000e-11 6.140000e-12  6.42403 1.3272e-10 ***
## distw_arithmetic -1.570894e-04 1.714231e-05 -9.16384  < 2.2e-16 ***
## contig1           1.044014e+00 9.979124e-02 10.46199  < 2.2e-16 ***
## comlang_ethno1    2.199545e-01 9.103465e-02  2.41616 1.5685e-02 *
## mobilecellular_o  1.327387e-03 1.299064e-03  1.02180 3.0687e-01
## mobilecellular_d -6.344137e-04 5.129710e-04 -1.23674 2.1618e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -4.244e+10   Adj. Pseudo R2: 0.908282
##            BIC:  8.488e+10     Squared Cor.: 0.859167
ppml_fixest_model_BB <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic+ contig +
comlang_ethno  + broadband_o + broadband_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = data
)
summary(ppml_fixest_model_BB)
```

```
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 208,376
## Fixed-effects: country_id_o: 188,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                        Estimate   Std. Error   z value   Pr(>|z|)
## gdp_o               4.720000e-11 1.443000e-11  3.268471 1.0813e-03 **
## gdp_d               4.030000e-11 6.370000e-12  6.318811 2.6358e-10
***
## distw_arithmetic -1.570945e-04 1.714758e-05 -9.161324  < 2.2e-16
***
## contig1             1.043969e+00 9.987180e-02 10.453095  < 2.2e-16
***
## comlang_ethno1      2.200337e-01 9.108533e-02  2.415688 1.5706e-02 *
## broadband_o        -2.701783e-03 6.721805e-03 -0.401943 6.8773e-01
## broadband_d        -6.517990e-04 2.490045e-03 -0.261762 7.9350e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -4.246e+10   Adj. Pseudo R2: 0.908249
##            BIC:  8.491e+10     Squared Cor.: 0.858956
ppml_fixest_model_IU <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic+ contig +
comlang_ethno  + Internetuser_o + Internetuser_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = data
)

summary(ppml_fixest_model_IU)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 208,376
## Fixed-effects: country_id_o: 188,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                        Estimate   Std. Error   z value   Pr(>|z|)
## gdp_o               4.590000e-11 8.650000e-12  5.308188 1.1072e-07
***
## gdp_d               3.940000e-11 5.390000e-12  7.316725 2.5410e-13
***
## distw_arithmetic -1.570935e-04 1.714995e-05 -9.159996  < 2.2e-16
***
## contig1             1.044067e+00 9.979403e-02 10.462220  < 2.2e-16
***
## comlang_ethno1      2.200598e-01 9.107096e-02  2.416355 1.5677e-02 *
## Internetuser_o     -2.706411e-03 2.845681e-03 -0.951059 3.4157e-01
## Internetuser_d      9.389141e-04 9.513611e-04  0.986917 3.2368e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -4.244e+10   Adj. Pseudo R2: 0.908274
##            BIC:  8.489e+10     Squared Cor.: 0.858539
```

PPML FE model for low income group

```
ppml_fixest_model_L <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataL
```

```
)
summary(ppml_fixest_model_L)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 14,381
## Fixed-effects: country_id_o: 22,  country_id_d: 187,  year: 11
## Standard-errors: Clustered (country_id_o)
##                         Estimate    Std. Error   z value   Pr(>|z|)
## gdp_o                1.762791e-08 1.246290e-08  1.414427 1.5724e-01
## gdp_d               -1.360000e-12 7.620000e-11 -0.017871 9.8574e-01
## distw_arithmetic -6.774364e-05 8.231953e-05 -0.822935 4.1054e-01
## contig1              1.527848e+00 3.902795e-01  3.914754 9.0496e-05
***
## comlang_ethno1       1.882133e-01 3.519172e-01  0.534823 5.9277e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -346,172,274.3   Adj. Pseudo R2: 0.640429
##            BIC:  692,346,683.5     Squared Cor.: 0.403897
ppml_fixest_model_IS_L <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic+ contig +
comlang_ethno + internetservers_o + internetservers_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataL
)
summary(ppml_fixest_model_IS_L)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 14,381
## Fixed-effects: country_id_o: 22,  country_id_d: 187,  year: 11
## Standard-errors: Clustered (country_id_o)
##                          Estimate    Std. Error   z value   Pr(>|z|)
## gdp_o                 2.170102e-08 1.083350e-08  2.003132 4.5163e-02 *
## gdp_d                -3.470000e-12 6.510000e-11 -0.053369 9.5744e-01
## distw_arithmetic  -6.516643e-05 8.095099e-05 -0.805011 4.2081e-01
## contig1               1.539487e+00 3.831084e-01  4.018411 5.8592e-05
***
## comlang_ethno1        1.917403e-01 3.520824e-01  0.544589 5.8604e-01
## internetservers_o  1.383009e-02 4.413212e-03  3.133792 1.7256e-03
**
## internetservers_d -5.878466e-06 2.996776e-06 -1.961597 4.9809e-02 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -342,791,052.0   Adj. Pseudo R2: 0.643941
##            BIC:  685,584,258.0     Squared Cor.: 0.417837
ppml_fixest_model_MC_L <- feglm(
    tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic+ contig +
comlang_ethno +  mobilecellular_o + mobilecellular_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataL
)
summary(ppml_fixest_model_MC_L)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 14,381
## Fixed-effects: country_id_o: 22,  country_id_d: 187,  year: 11
```

```
## Standard-errors: Clustered (country_id_o)
##                       Estimate   Std. Error   z value   Pr(>|z|)
## gdp_o               1.831810e-08 1.236240e-08  1.481768 1.3840e-01
## gdp_d              -1.530000e-11 7.530000e-11 -0.203447 8.3879e-01
## distw_arithmetic -6.716769e-05 8.196187e-05 -0.819499 4.1250e-01
## contig1             1.535415e+00 3.900706e-01  3.936248 8.2765e-05
***
## comlang_ethno1      1.902884e-01 3.511762e-01  0.541860 5.8791e-01
## mobilecellular_o  2.193338e-03 4.150392e-03  0.528465 5.9718e-01
## mobilecellular_d  6.978275e-03 2.965089e-03  2.353479 1.8599e-02 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -345,101,036.3   Adj. Pseudo R2: 0.641542
##             BIC:  690,204,226.7     Squared Cor.: 0.411825
ppml_fixest_model_BB_L <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic+ contig +
comlang_ethno +  broadband_o + broadband_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataL
)
summary(ppml_fixest_model_BB_L)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 14,381
## Fixed-effects: country_id_o: 22,  country_id_d: 187,  year: 11
## Standard-errors: Clustered (country_id_o)
##                       Estimate   Std. Error   z value   Pr(>|z|)
## gdp_o               1.990470e-08 1.171290e-08  1.699383 8.9247e-02 .
## gdp_d              -3.770000e-11 8.370000e-11 -0.450024 6.5269e-01
## distw_arithmetic -6.760456e-05 8.218506e-05 -0.822589 4.1074e-01
## contig1             1.523600e+00 3.923315e-01  3.883450 1.0298e-04
***
## comlang_ethno1      1.901335e-01 3.523516e-01  0.539613 5.8946e-01
## broadband_o         7.040758e-01 1.804443e-01  3.901900 9.5441e-05
***
## broadband_d         2.167006e-02 1.658716e-02  1.306435 1.9140e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -344,090,931.3   Adj. Pseudo R2: 0.642591
##             BIC:  688,184,016.7     Squared Cor.: 0.407949
ppml_fixest_model_IU_L <- feglm(
    tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic+ contig +
comlang_ethno +  Internetuser_o + Internetuser_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataL
)
summary(ppml_fixest_model_IU_L)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 14,381
## Fixed-effects: country_id_o: 22,  country_id_d: 187,  year: 11
## Standard-errors: Clustered (country_id_o)
##                       Estimate   Std. Error   z value   Pr(>|z|)
## gdp_o               1.880336e-08 1.232410e-08  1.525739 1.2707e-01
```

```
## gdp_d              -8.490000e-12 7.740000e-11 -0.109697 9.1265e-01
## distw_arithmetic -6.730404e-05 8.203917e-05 -0.820389 4.1199e-01
## contig1            1.530386e+00 3.900319e-01  3.923744 8.7183e-05
***
## comlang_ethno1     1.895187e-01 3.514811e-01  0.539200 5.8975e-01
## Internetuser_o    -3.233680e-03 1.339528e-02 -0.241404 8.0924e-01
## Internetuser_d     8.127415e-03 3.973797e-03  2.045252 4.0830e-02 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -345,891,769.0   Adj. Pseudo R2: 0.64072
##            BIC:  691,785,692.1     Squared Cor.: 0.406708
```

PPML FE model for lower-middle income group

```
ppml_fixest_model_LM <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataLM
)
summary(ppml_fixest_model_LM)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 48,620
## Fixed-effects: country_id_o: 51,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                      Estimate   Std. Error  z value   Pr(>|z|)
## gdp_o               4.465000e-10 1.474000e-10  3.02880 2.4553e-03 **
## gdp_d               1.660000e-11 1.530000e-11  1.08346 2.7860e-01
## distw_arithmetic -1.566479e-04 3.102429e-05 -5.04920 4.4366e-07 ***
## contig1             5.367365e-01 4.317247e-01  1.24324 2.1378e-01
## comlang_ethno1      6.668321e-01 1.620090e-01  4.11602 3.8547e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -5.008e+9    Adj. Pseudo R2: 0.797589
##            BIC:  1.002e+10     Squared Cor.: 0.481466
ppml_fixest_model_IS_LM <- feglm(
    tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic+ contig +
comlang_ethno +  internetservers_o + internetservers_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataLM
)
summary(ppml_fixest_model_IS_LM)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 48,620
## Fixed-effects: country_id_o: 51,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                       Estimate   Std. Error   z value   Pr(>|z|)
## gdp_o                4.464000e-10 1.502000e-10  2.971687 2.9617e-03
**
## gdp_d                1.400000e-11 1.360000e-11  1.026463 3.0467e-01
## distw_arithmetic  -1.567116e-04 3.100942e-05 -5.053677 4.3338e-07
***
## contig1              5.376568e-01 4.312885e-01  1.246629 2.1253e-01
## comlang_ethno1       6.661435e-01 1.618534e-01  4.115721 3.8597e-05
***
```

```
## internetservers_o -5.507176e-06 1.487407e-05 -0.370254 7.1119e-01
## internetservers_d  7.284620e-07 7.158227e-07  1.017657 3.0884e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -5.007e+9    Adj. Pseudo R2: 0.797625
##             BIC:  1.001e+10     Squared Cor.: 0.482662
ppml_fixest_model_MC_LM <- feglm(
    tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic+ contig +
comlang_ethno +  mobilecellular_o + mobilecellular_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataLM
)
summary(ppml_fixest_model_MC_LM)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 48,620
## Fixed-effects: country_id_o: 51,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                        Estimate   Std. Error   z value   Pr(>|z|)
## gdp_o               4.428000e-10 1.417000e-10  3.124268 1.7825e-03 **
## gdp_d               1.120000e-11 1.870000e-11  0.597993 5.4984e-01
## distw_arithmetic -1.565091e-04 3.100751e-05 -5.047459 4.4773e-07
***
## contig1             5.365391e-01 4.318541e-01  1.242408 2.1409e-01
## comlang_ethno1      6.669858e-01 1.621172e-01  4.114220 3.8849e-05
***
## mobilecellular_o  2.269233e-03 3.571205e-03  0.635425 5.2515e-01
## mobilecellular_d  2.207813e-03 1.431754e-03  1.542033 1.2307e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood:    -5.002e+9   Adj. Pseudo R2: 0.797816
##             BIC: 1e+10          Squared Cor.: 0.480466
ppml_fixest_model_BB_LM <- feglm(
    tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic+ contig +
comlang_ethno +  broadband_o + broadband_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataLM
)
summary(ppml_fixest_model_BB_LM)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 48,620
## Fixed-effects: country_id_o: 51,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                        Estimate   Std. Error   z value   Pr(>|z|)
## gdp_o               4.487000e-10 1.782000e-10  2.518560 1.1784e-02 *
## gdp_d               2.030000e-11 2.210000e-11  0.918145 3.5854e-01
## distw_arithmetic -1.566968e-04 3.101190e-05 -5.052796 4.3539e-07
***
## contig1             5.373350e-01 4.308681e-01  1.247099 2.1236e-01
## comlang_ethno1      6.661475e-01 1.617847e-01  4.117495 3.8301e-05
***
## broadband_o      -5.409394e-04 2.807060e-02 -0.019271 9.8463e-01
## broadband_d      -3.418796e-03 8.969741e-03 -0.381148 7.0309e-01
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -5.007e+9    Adj. Pseudo R2: 0.797605
##            BIC:  1.001e+10     Squared Cor.: 0.482134
ppml_fixest_model_IU_LM <- feglm(
    tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic+ contig +
comlang_ethno +  Internetuser_o + Internetuser_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataLM
)
summary(ppml_fixest_model_IU_LM)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 48,620
## Fixed-effects: country_id_o: 51,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                       Estimate    Std. Error   z value   Pr(>|z|)
## gdp_o               3.478000e-10 1.524000e-10  2.282053 2.2486e-02 *
## gdp_d               1.560000e-11 1.560000e-11  1.004830 3.1498e-01
## distw_arithmetic -1.566247e-04 3.101186e-05 -5.050477 4.4071e-07
***
## contig1             5.370526e-01 4.315978e-01  1.244336 2.1338e-01
## comlang_ethno1      6.670095e-01 1.619493e-01  4.118631 3.8113e-05
***
## Internetuser_o    -7.336838e-03 4.275683e-03 -1.715946 8.6172e-02 .
## Internetuser_d      4.709358e-04 2.302158e-03  0.204563 8.3791e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood:     -5.002e+9   Adj. Pseudo R2: 0.797813
##            BIC: 1e+10             Squared Cor.: 0.481893
```

PPML FE model for upper-middle income group

```
ppml_fixest_model_UM <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataUM
)
summary(ppml_fixest_model_UM)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 58,400
## Fixed-effects: country_id_o: 52,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                       Estimate    Std. Error  z value   Pr(>|z|)
## gdp_o               5.020000e-11 9.410000e-12  5.34026 9.2812e-08 ***
## gdp_d               5.430000e-11 1.405000e-11  3.86413 1.1149e-04 ***
## distw_arithmetic -1.102493e-04 2.856927e-05 -3.85902 1.1384e-04 ***
## contig1             1.006872e+00 1.359364e-01  7.40693 1.2925e-13 ***
## comlang_ethno1      6.625722e-01 2.398000e-01  2.76302 5.7269e-03 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -1.057e+10   Adj. Pseudo R2: 0.928692
##            BIC:  2.115e+10     Squared Cor.: 0.927496
ppml_fixest_model_IS_UM <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
```

```
comlang_ethno + internetservers_o + internetservers_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataUM
)
summary(ppml_fixest_model_IS_UM)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 58,400
## Fixed-effects: country_id_o: 52,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                        Estimate   Std. Error   z value   Pr(>|z|)
## gdp_o               5.340000e-11 1.050000e-11  5.085743 3.6619e-07
***
## gdp_d               6.930000e-11 1.330000e-11  5.214647 1.8417e-07
***
## distw_arithmetic   -1.102366e-04 2.856950e-05 -3.858542 1.1407e-04
***
## contig1             1.006719e+00 1.359294e-01  7.406194 1.2998e-13
***
## comlang_ethno1      6.631124e-01 2.398914e-01  2.764220 5.7059e-03
**
## internetservers_o  2.661179e-06 4.937395e-06  0.538984 5.8990e-01
## internetservers_d -1.450388e-06 4.756545e-07 -3.049246 2.2942e-03
**
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -1.056e+10   Adj. Pseudo R2: 0.928807
##            BIC:  2.111e+10     Squared Cor.: 0.927955
ppml_fixest_model_MC_UM <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno +  mobilecellular_o + mobilecellular_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataUM
)
summary(ppml_fixest_model_MC_UM)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 58,400
## Fixed-effects: country_id_o: 52,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                       Estimate   Std. Error   z value   Pr(>|z|)
## gdp_o              4.450000e-11 1.080000e-11  4.137637 3.5090e-05
***
## gdp_d              5.390000e-11 1.300000e-11  4.155895 3.2402e-05
***
## distw_arithmetic  -1.102118e-04 2.854663e-05 -3.860764 1.1303e-04
***
## contig1            1.006979e+00 1.359317e-01  7.407976 1.2824e-13
***
## comlang_ethno1     6.626390e-01 2.397317e-01  2.764085 5.7083e-03 **
## mobilecellular_o  1.039461e-03 1.494767e-03  0.695400 4.8680e-01
## mobilecellular_d  2.573748e-04 7.609928e-04  0.338209 7.3521e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Log-Likelihood: -1.057e+10   Adj. Pseudo R2: 0.928707
##             BIC:  2.114e+10     Squared Cor.: 0.927133
ppml_fixest_model_BB_UM <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno +  broadband_o + broadband_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataUM
)
summary(ppml_fixest_model_BB_UM)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 58,400
## Fixed-effects: country_id_o: 52,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                      Estimate   Std. Error   z value   Pr(>|z|)
## gdp_o             3.810000e-11 1.200000e-11  3.168456 1.5325e-03 **
## gdp_d             5.540000e-11 1.330000e-11  4.162309 3.1505e-05
***
## distw_arithmetic -1.102521e-04 2.857092e-05 -3.858894 1.1390e-04
***
## contig1           1.006804e+00 1.359751e-01  7.404326 1.3182e-13
***
## comlang_ethno1    6.629326e-01 2.396663e-01  2.766066 5.6737e-03 **
## broadband_o       6.423758e-03 6.701716e-03  0.958524 3.3780e-01
## broadband_d      -1.245918e-03 4.682642e-03 -0.266072 7.9018e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -1.057e+10   Adj. Pseudo R2: 0.928704
##             BIC:  2.115e+10     Squared Cor.: 0.927222
ppml_fixest_model_IU_UM <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno +  Internetuser_o + Internetuser_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataUM
)
summary(ppml_fixest_model_IU_UM)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 58,400
## Fixed-effects: country_id_o: 52,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                      Estimate   Std. Error  z value   Pr(>|z|)
## gdp_o             5.810000e-11 8.720000e-12  6.66783 2.5962e-11 ***
## gdp_d             5.320000e-11 1.265000e-11  4.20875 2.5679e-05 ***
## distw_arithmetic -1.101779e-04 2.851049e-05 -3.86447 1.1133e-04 ***
## contig1           1.007851e+00 1.357236e-01  7.42576 1.1213e-13 ***
## comlang_ethno1    6.630343e-01 2.396930e-01  2.76618 5.6717e-03 **
## Internetuser_o    1.257851e-02 4.369523e-03  2.87869 3.9933e-03 **
## Internetuser_d    5.348971e-03 1.773747e-03  3.01563 2.5644e-03 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -1.052e+10   Adj. Pseudo R2: 0.929038
##             BIC:  2.105e+10     Squared Cor.: 0.92759
```

PPML FE model for high income group

```r
ppml_fixest_model_H <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataH
)
summary(ppml_fixest_model_H)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 86,678
## Fixed-effects: country_id_o: 62,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                        Estimate    Std. Error    z value    Pr(>|z|)
## gdp_o              2.700000e-11 8.280000e-12   3.255080 1.1336e-03
**
## gdp_d              3.420000e-11 6.060000e-12   5.640014 1.7004e-08
***
## distw_arithmetic -1.801711e-04 1.568653e-05 -11.485720  < 2.2e-16
***
## contig1            1.280414e+00 1.105280e-01  11.584528  < 2.2e-16
***
## comlang_ethno1     1.262163e-02 8.338840e-02   0.151360 8.7969e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -2.039e+10   Adj. Pseudo R2: 0.921134
##             BIC:  4.078e+10     Squared Cor.: 0.922021
ppml_fixest_model_IS_H <- feglm(
   tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno+ internetservers_o + internetservers_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataH
)
summary(ppml_fixest_model_IS_H)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 86,678
## Fixed-effects: country_id_o: 62,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                         Estimate    Std. Error    z value    Pr(>|z|)
## gdp_o               1.730000e-11 7.300000e-12   2.369457 1.7814e-02
*
## gdp_d               3.350000e-11 6.230000e-12   5.385909 7.2080e-08
***
## distw_arithmetic   -1.801811e-04 1.568694e-05 -11.486054  < 2.2e-16
***
## contig1             1.280314e+00 1.105423e-01  11.582116  < 2.2e-16
***
## comlang_ethno1      1.256278e-02 8.338625e-02   0.150658 8.8025e-01
## internetservers_o   8.891775e-07 5.796581e-07   1.533969 1.2504e-01
## internetservers_d   4.590065e-07 3.092404e-07   1.484303 1.3773e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -2.037e+10   Adj. Pseudo R2: 0.921191
##             BIC:  4.075e+10     Squared Cor.: 0.921435
```

```r
ppml_fixest_model_MC_H <- feglm(
   tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno+ mobilecellular_o + mobilecellular_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataH
)
summary(ppml_fixest_model_MC_H)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 86,678
## Fixed-effects: country_id_o: 62,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                        Estimate   Std. Error    z value   Pr(>|z|)
## gdp_o               2.590000e-11 8.840000e-12   2.925913 3.4345e-03
**
## gdp_d               3.420000e-11 7.770000e-12   4.404957 1.0581e-05
***
## distw_arithmetic -1.801579e-04 1.569147e-05 -11.481267  < 2.2e-16
***
## contig1             1.280163e+00 1.105163e-01  11.583472  < 2.2e-16
***
## comlang_ethno1      1.252790e-02 8.340225e-02   0.150211 8.8060e-01
## mobilecellular_o    1.549522e-03 1.912909e-03   0.810035 4.1792e-01
## mobilecellular_d -1.044631e-03 6.938598e-04  -1.505537 1.3219e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -2.037e+10   Adj. Pseudo R2: 0.921203
##             BIC:  4.074e+10     Squared Cor.: 0.922104
ppml_fixest_model_BB_H <- feglm(
   tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno+ broadband_o + broadband_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataH
)
summary(ppml_fixest_model_BB_H)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 86,678
## Fixed-effects: country_id_o: 62,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                        Estimate   Std. Error    z value   Pr(>|z|)
## gdp_o               2.360000e-11 7.900000e-12   2.991279 2.7781e-03
**
## gdp_d               3.260000e-11 8.170000e-12   3.987096 6.6887e-05
***
## distw_arithmetic -1.801617e-04 1.568953e-05 -11.482926  < 2.2e-16
***
## contig1             1.280485e+00 1.105485e-01  11.583016  < 2.2e-16
***
## comlang_ethno1      1.248099e-02 8.341239e-02   0.149630 8.8106e-01
## broadband_o        -1.247083e-02 1.025290e-02  -1.216322 2.2386e-01
## broadband_d         3.690141e-04 2.898463e-03   0.127314 8.9869e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Log-Likelihood: -2.036e+10   Adj. Pseudo R2: 0.921242
##            BIC:  4.072e+10     Squared Cor.: 0.922222
ppml_fixest_model_IU_H <- feglm(
   tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno+ Internetuser_o + Internetuser_d
  | country_id_o + country_id_d + year,
  family = poisson(),
  data = dataH
)
summary(ppml_fixest_model_IU_H)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 86,678
## Fixed-effects: country_id_o: 62,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                       Estimate   Std. Error    z value   Pr(>|z|)
## gdp_o              3.350000e-11 9.690000e-12   3.457328 5.4556e-04
***
## gdp_d              3.530000e-11 6.620000e-12   5.336226 9.4901e-08
***
## distw_arithmetic  -1.801582e-04 1.569090e-05 -11.481698  < 2.2e-16
***
## contig1            1.280407e+00 1.105372e-01  11.583495  < 2.2e-16
***
## comlang_ethno1     1.264164e-02 8.339852e-02   0.151581 8.7952e-01
## Internetuser_o    -6.153388e-03 4.108505e-03  -1.497719 1.3421e-01
## Internetuser_d    -6.482137e-04 8.680621e-04  -0.746737 4.5522e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -2.035e+10   Adj. Pseudo R2: 0.921271
##            BIC:  4.07e+10      Squared Cor.: 0.921636
```

Robustness check

##Interaction between ICT variables at origin and destination

```
# Interaction between ICT variables at origin and destination
ppml_fixest_model_interaction_MC <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno +   mobilecellular_o * mobilecellular_d|
  country_id_o + country_id_d + year,
  family = poisson(),
  data = data
)
summary(ppml_fixest_model_interaction_MC)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 208,376
## Fixed-effects: country_id_o: 188,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                                         Estimate   Std. Error    z
value
## gdp_o                                4.130000e-11 8.160000e-12
5.054632
## gdp_d                                4.130000e-11 6.900000e-12
5.981113
## distw_arithmetic                    -1.565746e-04 1.711021e-05 -
9.150943
## contig1                              1.043034e+00 1.012066e-01
```

```
10.305983
## comlang_ethno1                     2.314694e-01 9.024076e-02
2.565021
## mobilecellular_o                  -2.111384e-03 3.465470e-03 -
0.609263
## mobilecellular_d                  -4.197381e-03 2.584848e-03 -
1.623840
## mobilecellular_o:mobilecellular_d  2.707069e-05 2.017354e-05
1.341891
##                                    Pr(>|z|)
## gdp_o                             4.3122e-07 ***
## gdp_d                             2.2162e-09 ***
## distw_arithmetic                   < 2.2e-16 ***
## contig1                            < 2.2e-16 ***
## comlang_ethno1                    1.0317e-02 *
## mobilecellular_o                  5.4235e-01
## mobilecellular_d                  1.0441e-01
## mobilecellular_o:mobilecellular_d 1.7963e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -4.238e+10   Adj. Pseudo R2: 0.908406
##            BIC:  8.477e+10     Squared Cor.: 0.858928
ppml_fixest_model_interaction_BB <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno +   broadband_o * broadband_d|
  country_id_o + country_id_d + year,
  family = poisson(),
  data = data
)
summary(ppml_fixest_model_interaction_BB)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 208,376
## Fixed-effects: country_id_o: 188,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                                Estimate   Std. Error  z value
Pr(>|z|)
## gdp_o                       4.730000e-11 1.467000e-11  3.22571
1.2566e-03 **
## gdp_d                       4.040000e-11 6.750000e-12  5.98455
2.1699e-09 ***
## distw_arithmetic           -1.556637e-04 1.672989e-05 -9.30453  <
2.2e-16 ***
## contig1                     1.025227e+00 1.013315e-01 10.11756  <
2.2e-16 ***
## comlang_ethno1              2.196785e-01 9.228336e-02  2.38048
1.7290e-02 *
## broadband_o                -1.235310e-02 7.537862e-03 -1.63881
1.0125e-01
## broadband_d                -1.095385e-02 3.672823e-03 -2.98241
2.8599e-03 **
## broadband_o:broadband_d     3.957494e-04 1.285622e-04  3.07827
2.0821e-03 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Log-Likelihood: -4.223e+10   Adj. Pseudo R2: 0.908745
##            BIC:  8.445e+10     Squared Cor.: 0.857797
ppml_fixest_model_interaction_IU <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno +   Internetuser_o * Internetuser_d|
  country_id_o + country_id_d + year,
  family = poisson(),
  data = data
)
summary(ppml_fixest_model_interaction_IU)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 208,376
## Fixed-effects: country_id_o: 188,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                                     Estimate   Std. Error  z value
Pr(>|z|)
## gdp_o                            4.810000e-11 1.014000e-11  4.74195
2.1167e-06
## gdp_d                            4.010000e-11 5.450000e-12  7.36028
1.8352e-13
## distw_arithmetic                -1.569490e-04 1.688069e-05 -9.29755
< 2.2e-16
## contig1                          1.021117e+00 1.000728e-01 10.20374
< 2.2e-16
## comlang_ethno1                   2.328258e-01 9.226672e-02  2.52340
1.1623e-02
## Internetuser_o                  -8.249360e-03 4.049133e-03 -2.03732
4.1618e-02
## Internetuser_d                  -4.380669e-03 2.951760e-03 -1.48409
1.3779e-01
## Internetuser_o:Internetuser_d    9.475551e-05 4.784561e-05  1.98044
4.7654e-02
##
## gdp_o                           ***
## gdp_d                           ***
## distw_arithmetic                ***
## contig1                         ***
## comlang_ethno1                  *
## Internetuser_o                  *
## Internetuser_d
## Internetuser_o:Internetuser_d *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -4.231e+10   Adj. Pseudo R2: 0.908563
##            BIC:  8.462e+10     Squared Cor.: 0.854944
ppml_fixest_model_interaction_IS <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno +   internetservers_o * internetservers_d|
  country_id_o + country_id_d + year,
  family = poisson(),
  data = data
)
summary(ppml_fixest_model_interaction_IS)
```

```
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 208,376
## Fixed-effects: country_id_o: 188,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                                        Estimate   Std. Error   z
value
## gdp_o                               4.386000e-11 7.920000e-12
5.536045
## gdp_d                               4.017000e-11 5.960000e-12
6.744043
## distw_arithmetic                   -1.570468e-04 1.714389e-05 -
9.160510
## contig1                             1.043415e+00 9.970311e-02
10.465221
## comlang_ethno1                      2.220249e-01 9.114980e-02
2.435824
## internetservers_o                   2.365417e-08 6.566864e-07
0.036020
## internetservers_d                  -1.963892e-07 3.521698e-07 -
0.557655
## internetservers_o:internetservers_d  8.350000e-12 9.650000e-12
0.865228
##                                      Pr(>|z|)
## gdp_o                               3.0938e-08 ***
## gdp_d                               1.5404e-11 ***
## distw_arithmetic                     < 2.2e-16 ***
## contig1                              < 2.2e-16 ***
## comlang_ethno1                      1.4858e-02 *
## internetservers_o                   9.7127e-01
## internetservers_d                   5.7708e-01
## internetservers_o:internetservers_d 3.8691e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -4.245e+10   Adj. Pseudo R2: 0.908263
##            BIC:  8.49e+10      Squared Cor.: 0.857974
```

##Quadratic Term for ICT Variables quadratic terms for the ICT variables to check for non-linear effects

```
ppml_fixest_model_quadratic_MC <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno +
  mobilecellular_o + mobilecellular_d + I(mobilecellular_o^2) +
I(mobilecellular_d^2)
 |  country_id_o + country_id_d + year,
  family = poisson(),
  data = data
)
summary(ppml_fixest_model_quadratic_MC)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 208,376
## Fixed-effects: country_id_o: 188,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                           Estimate   Std. Error   z value
Pr(>|z|)
## gdp_o                  4.210000e-11 1.024000e-11  4.111347 3.9336e-
05 ***
```

```
## gdp_d                       3.570000e-11 5.910000e-12  6.038617 1.5544e-
09 ***
## distw_arithmetic        -1.570755e-04 1.715759e-05 -9.154869  < 2.2e-
16 ***
## contig1                     1.044133e+00 9.970088e-02 10.472661  < 2.2e-
16 ***
## comlang_ethno1              2.200973e-01 9.099588e-02  2.418761 1.5573e-
02 *
## mobilecellular_o            3.616833e-04 2.212701e-03  0.163458 8.7016e-
01
## mobilecellular_d            1.515306e-03 1.682784e-03  0.900476 3.6787e-
01
## I(mobilecellular_o^2)  3.205915e-06 4.522611e-06  0.708864 4.7841e-
01
## I(mobilecellular_d^2) -6.819688e-06 5.143141e-06 -1.325977 1.8485e-
01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -4.243e+10   Adj. Pseudo R2: 0.908302
##            BIC:  8.486e+10     Squared Cor.: 0.859689
ppml_fixest_model_quadratic_BB <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno +
  broadband_o + broadband_d + I(broadband_o^2) + I(broadband_d^2)
  |  country_id_o + country_id_d + year,
  family = poisson(),
  data = data
)
summary(ppml_fixest_model_quadratic_BB)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 208,376
## Fixed-effects: country_id_o: 188,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                        Estimate    Std. Error   z value   Pr(>|z|)
## gdp_o               4.650000e-11 1.348000e-11  3.452714 5.5498e-04
***
## gdp_d               4.080000e-11 7.020000e-12  5.814522 6.0807e-09
***
## distw_arithmetic -1.570956e-04 1.714782e-05 -9.161263  < 2.2e-16
***
## contig1             1.044012e+00 9.975275e-02 10.465998  < 2.2e-16
***
## comlang_ethno1      2.200372e-01 9.105268e-02  2.416592 1.5667e-02 *
## broadband_o        -1.569549e-02 1.181636e-02 -1.328285 1.8408e-01
## broadband_d         2.129829e-04 5.162560e-03  0.041255 9.6709e-01
## I(broadband_o^2)  3.462202e-04 2.043403e-04  1.694331 9.0202e-02 .
## I(broadband_d^2) -2.022297e-05 1.088403e-04 -0.185804 8.5260e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -4.242e+10   Adj. Pseudo R2: 0.908327
##            BIC:  8.484e+10     Squared Cor.: 0.858329
ppml_fixest_model_quadratic_IU <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno +
```

```
  Internetuser_o + Internetuser_d + I(Internetuser_o^2) +
I(Internetuser_d^2)
  |  country_id_o + country_id_d + year,
  family = poisson(),
  data = data
)
summary(ppml_fixest_model_quadratic_IU)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 208,376
## Fixed-effects: country_id_o: 188,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                          Estimate   Std. Error   z value   Pr(>|z|)
## gdp_o                4.480000e-11 6.830000e-12  6.552475 5.6591e-11
***
## gdp_d                3.950000e-11 5.390000e-12  7.325409 2.3817e-13
***
## distw_arithmetic    -1.570920e-04 1.714985e-05 -9.159962  < 2.2e-16
***
## contig1              1.043835e+00 9.992530e-02 10.446150  < 2.2e-16
***
## comlang_ethno1       2.202584e-01 9.109495e-02  2.417899 1.5610e-02
*
## Internetuser_o       1.875136e-03 4.090094e-03  0.458458 6.4662e-01
## Internetuser_d       2.475767e-03 2.609300e-03  0.948824 3.4271e-01
## I(Internetuser_o^2) -4.334612e-05 3.867734e-05 -1.120711 2.6241e-01
## I(Internetuser_d^2) -1.536896e-05 1.937686e-05 -0.793160 4.2768e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -4.243e+10   Adj. Pseudo R2: 0.908304
##            BIC:  8.486e+10     Squared Cor.: 0.85856
ppml_fixest_model_quadratic_IS <- feglm(
  tradeflow_imf_o ~ gdp_o + gdp_d + distw_arithmetic + contig +
comlang_ethno +
  internetservers_o + internetservers_d + I(internetservers_o^2) +
I(internetservers_d^2)
  |  country_id_o + country_id_d + year,
  family = poisson(),
  data = data
)
summary(ppml_fixest_model_quadratic_IS)
## GLM estimation, family = poisson, Dep. Var.: tradeflow_imf_o
## Observations: 208,376
## Fixed-effects: country_id_o: 188,  country_id_d: 188,  year: 11
## Standard-errors: Clustered (country_id_o)
##                             Estimate   Std. Error   z value
Pr(>|z|)
## gdp_o                   4.408000e-11 9.170000e-12  4.806221
1.5381e-06 ***
## gdp_d                   3.967000e-11 5.930000e-12  6.693541
2.1783e-11 ***
## distw_arithmetic       -1.571030e-04 1.715659e-05 -9.157007  <
2.2e-16 ***
## contig1                 1.044020e+00 9.978316e-02 10.462891  <
2.2e-16 ***
```

```
## comlang_ethno1          2.199765e-01 9.107371e-02  2.415367
1.5719e-02 *
## internetservers_o       1.172391e-06 1.204486e-06  0.973354
3.3038e-01
## internetservers_d      -1.907359e-07 6.443748e-07 -0.296001
7.6723e-01
## I(internetservers_o^2) -5.450000e-12 5.600000e-12 -0.972255
3.3092e-01
## I(internetservers_d^2)  1.820000e-12 2.470000e-12  0.735043
4.6231e-01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -4.245e+10   Adj. Pseudo R2: 0.90826
##            BIC:  8.49e+10      Squared Cor.: 0.858833
```