Erasmus University Rotterdam
Erasmus School of Economics
Bachelor Thesis Double Degree Econometrics and Economics

Assessing the Robustness of the Exchangeable Estimator in Times of Economic Instability

Hélène Friso (503908)

Supervisor:	P. Wan
Second assessor:	A. Archimbaud
Date final version:	2nd July 2023

Abstract

This thesis examines the robustness of the exchangeable estimator in predicting trade flows during periods of economic instability. The research investigates how the exchangeable estimator performs compared to a non-linear model when faced with economically disruptive events, focusing on the aftermath of the fall of the Berlin Wall in 1989. Simulation comparisons reveal the superiority of the exchangeable estimator over the dyadic clustering estimator, especially when combined with the exchangeable error model. Trade flow predictions using a linear model demonstrate better performance with the exchangeable estimator compared to ordinary least squares estimation, as supported by prior research. The study evaluates the robustness of the exchangeable estimator by analyzing a structural break for the year 1989, corresponding to the fall of the Berlin Wall. Results confirm a significant impact on trade flows, supporting the hypothesis. Additionally, the exchangeable estimator's performance during economic instability is compared to the non-linear Random Forest model. Both, the Random Forest model and the linear model using the exchangeable estimator experience a slight drop in accuracy during the aftermath of the Berlin Wall, but quickly return to their previous level. In conclusion, this research provides insights into the exchangeable estimator's performance. It demonstrates the estimator's resilience in adapting to sudden changes in relational data. Overall, the exchangeable estimator remains a robust choice for predicting trade flows.

Contents

1 Introduction **3**

2 Literature Review **5**

2.1 Economic Instability and International Trade 5
2.2 Forecasting Models Trade Flow 7
2.3 Robustness Testing 8
2.4 The Fall of the Berlin Wall 9
2.5 Non-linear Prediction Models 10

3 Data **11**

4 Methodology **12**

4.1 Estimators for Linear Regression 12
4.1.1 Ordinary Least Squares 12
4.1.2 Dyadic Clustering Estimator 13
4.1.3 Exchangeable Estimator 14
4.2 Simulation 14
4.3 Forecasting Trade Flow 16
4.4 Structural Break Analysis 17
4.5 Random Forest 18

5 Results **20**

5.1 Simulation 20
5.2 Trade 21
5.3 Structural Break 22
5.4 Random Forest 23
5.5 Comparison 23

6 Discussion **24**

7 Conclusion **26**

References **29**

A Growth trade volume **34**

B List of countries **34**

C Exchangeable estimator **35**

1 Introduction

The forecasting of international trade flow holds a crucial task for governmental entities, businesses, and international organizations, as it lays the foundation for informed decision-making regarding trade policies, investment strategies, and economic development initiatives. Nevertheless, the accurate prediction of trade flow can be arduous, particularly during periods of economic instability, such as recessions or financial crises. The exchangeable estimator, introduced as a statistical model by Marrs, Fosdick and Mccormick (2023), has exhibited promising performance in achieving highly accurate trade flow predictions. However, the robustness of the exchangeable estimator in forecasting trade flow during times of economic instability remains uncertain. An exemplary economic event with significant international impact is the fall of the Berlin Wall, which took place in 1989, profoundly impacting the global economy and international trade. This event signified the conclusion of the Cold War and the emergence of new trade markets. Hence, the main research question for this thesis is: *"Is the exchangeable estimator robust in predicting trade flow during periods of economic instability, such as the aftermath of the fall of the Berlin Wall in 1989, and how does the linear predictive model using the exchangeable estimator perform compared to the non-linear Random Forest model when economically disturbing events occur?"*

To address this question comprehensively, several sub-questions necessitate examination. These inquiries include addressing the distinctions between the exchangeable estimator and the conventional ordinary least squares estimator. Furthermore, a comparative analysis between the exchangeable estimator and the recognized dyadic clustering estimator, another estimator for relational arrays, should be conducted. Lastly, the data should be tested on a structural break and the non-linear Random Forest (RF) model should be evaluated. Consequently, the ensuing sub-questions arise:

- What is the exchangeable estimator and how does it differ from the ordinary least squares estimator?
- How does exchangeable estimator performs against the dyadic clustering estimator?
- How does the performance of the exchangeable estimator compare to the regular ordinary least squares estimator in terms of model fit?
- Did the fall of the Berlin Wall significantly change the trade flow?
- How does the exchangeable estimator perform in predicting trade flow during periods of economic instability, such as the fall of the Berlin Wall?
- What approach should be considered to capture the effect of economic events, and does it improve the model?

The examination of the robustness of the exchangeable estimator in forecasting trade flow amidst economic instability holds significance both academically and practically. From an academic standpoint, this research augments the existing literature pertaining to the efficacy of statistical models in forecasting economic outcomes within periods of instability. By shedding light on the performance of the exchangeable estimator, it contributes to the understanding of predictive modeling under such conditions. In practical terms, this research bears relevance for trade policies and strategies, with the potential to provide valuable insights for mitigating the negative effects of economic instability on trade flow. Governments, businesses, and international organizations, reliant upon accurate trade flow prognostications to facilitate informed decision-making, stand to benefit from the findings of this study. Predicting international trade helps governments and policymakers in formulating effective economic policies, trade agreements, and investment strategies. Moreover, it allows them to assess the potential impacts of trade on their domestic industries, employment levels, and overall economic growth. Businesses, especially those engaged in international trade, rely on trade predictions to make informed decisions about market expansion, product development, and supply chain management. Predicting trade patterns and trends helps businesses identify new market opportunities, assess competition, and optimize their operations accordingly. Furthermore, understanding and predicting international trade patterns assist governments and businesses in efficiently allocating resources, such as capital, labor, and infrastructure. It enables them to align their resources with the expected demand for goods and services, optimize production capabilities, and enhance overall competitiveness. Predicting international trade allows governments to develop and modify trade policies, tariffs, and regulations that are aligned with market trends and dynamics. By having accurate predictions, policymakers can create an environment that fosters trade, attracts investments, and strengthens international partnerships. Lastly, international organizations, such as the World Trade Organization (WTO), use trade predictions to facilitate dialogue, negotiations, and cooperation among member countries. Predicting trade helps identify areas of potential conflict, explore opportunities for trade liberalization, and foster international economic integration. Thus, the research holds practical implications for stakeholders engaged in the formulation and implementation of trade-related measures.

The theoretical framework underpinning this research encompasses the utilization of mixed effects models for analyzing relational data. This analytical approach accommodates the consideration of both actor-specific and relation-specific effects, while also incorporating random effects to address unobservable heterogeneity. Of particular relevance is the exchangeable estimator, positing equal covariance across relations for all actors. By assuming exchangeability, this estimator assumes that the underlying structure of covariance is uniform across actors.

The exchangeable estimator, while recognized as a potent statistical tool in trade flow prediction, lacks understanding regarding its performance amidst economic instability. The occurrence of global economic instability, encompassing recessions or financial crises, exerts a substantial influence on trade flow. However, the precise capacity of the exchangeable estimator to accurately capture this influence remains uncertain. Consequently, this research is imperative for bridging this knowledge gap and acquiring a more nuanced understanding of the exchangeable estimator's limitations and potential in the realm of trade flow prediction during periods of economic instability.

By addressing this research gap, the study not only contributes to knowledge but also facilitates the development of more precise and robust statistical models for trade flow prediction. The implications of such findings are of practical importance, as they can enable the formulation of more effective trade policies and strategies. Ultimately, this research strives to enhance the accuracy and reliability of trade flow predictions, thereby fostering informed decision-making processes within the field of international trade.

In Section 2, the existing literature is discussed and hypotheses are defined. Next, Section 3 describes the dataset of the trade flows and covariates. Then, Section 4 explains the methodology by first going over the different estimators, then explaining the method applied to perform a simulation, followed by the linear forecasting model for the trade flows. Afterwards, the RF algorithm is introduced, and comparisons measures are defined to evaluate the robustness of the estimator. Subsequently, in Section 5 the results are presented. Finally, the limitations and further research ideas are discussed, after which the conclusion is drawn.

2 Literature Review

2.1 Economic Instability and International Trade

International trade has increased steeply since the establishment of the WTO in 1995 , as shown in Appendix A (World Trade Organization, 2022). Trade plays a crucial role in maintaining a competitive global economy by fostering innovation, promoting market specialization, and facilitating lower prices for goods on an international scale (Qurban, 2021). The dynamics of international trade is highly reliant on the economic status of the countries involved in the commerce. Political stability, trade agreements, and tariffs and trade barriers play crucial roles in facilitating international trade (Acemoglu, Johnson & Robinson, 2001; Freund & Ornelas, 2010; Anderson & van Wincoop, 2004), as well as market size and demand, comparative and technological advantages, and infrastructure and logistics (Helpman, Melitz & Rubinstein, 2008a; Ricardo, 1817; Feenstra & Kee, 2008; Djankov, Freund & Pham, 2010). However, international trade is prone to exogenous factors, due to the numerous influential elements. While regional external events could impact the trade, such as natural disasters, global external events yield more profound repercussions. Considering some recent occurrences, the international trade has suffered noticeably.

For example, the global financial crisis of 2008 had a profound impact on international trade, leading to a significant decline in trade volumes and altering trade patterns (Baldwin, 2011). Weakened global demand and financial market turmoil resulted in a contraction of international trade, as shown in Appendix A, particularly in durable goods and capital-intensive industries. Banks' reluctance to lend and increased borrowing costs reduced access to trade finance, affecting trade transactions, particularly for small and medium-sized enterprises (Chor & Manova, 2012).

Another notable example is the COVID-19 pandemic. The pandemic had significant effect on international trade due to widespread disruptions and containment measures (International Monetary Fund, 2021; World Trade Organization, 2020). The global supply chain was disrupted due to the lockdowns, travel restriction, and factory closures (Baldwin & Tomiura, 2020). This led to reduced production and trade volumes, again demonstrated in Appendix A. Furthermore, decreased consumer spending and reduced business investments resulted in decreased demand for goods and services (McKinsey & Company, 2021). Lastly, countries implemented trade restrictions, export bans, and protective measures to ensure domestic supplies of essential goods (Federation of German Industries (BDI), 2021). All these factors contributed to a strong decline in international trade, and partially led to strong economic decline (World Trade Organization, 2021).

At last, the war in Ukraine exemplifies the effect of economic instability on international trade and led to socioeconomic consequences, both domestically and in exporting countries. The war led to conflicts involving gas and oil supplies, which especially affected the energy sector (Di Bella et al., 2022; International Energy Agency, 2022). The disruption in gas and oil supplies and geopolitical tensions in Ukraine led to gas and oil price spikes, affecting energy-dependent industries and trade costs for countries reliant on gas and oil imports. Uncertainty and instability in gas supplies raised concerns about energy security, prompting countries to diversify energy sources, and trade routes. Furthermore, the Russian-Ukrainian conflict poses significant risks to many African communities that rely heavily on agricultural imports. As globally recognized exporters of crucial foodstuffs and grains, Agence Française de Développement (2023) state that both Ukraine and Russia play a vital role in the food supply chain across Africa.

These events underscore the significant impact of economic instability on international trade, and evidently shape the livelihoods of individuals both within domestic economies and across exporting nations. Often unpredictable or unforeseen, such events lead to abrupt changes in the exchange of goods and services between countries. The examples mentioned evince how economic shocks propagate across borders partially through trade. This phenomenon is closely tied to the concepts and principles underlying the international business cycle theory. International business cycle theory examines the synchronization and transmission of economic fluctuations across countries. It explores how shocks, such as financial crises or changes in global demand, can propagate through international trade channels and affect economic stability in different countries (D. K. Backus, Kehoe

& Kydland, 1992; D. Backus, Kehoe & Kydland, 1993; Ravn, 1997). The theory has been developed and expanded upon by various economists over time through empirical research and theoretical contributions. While some have focused on the financial interconnections between countries, others have conducted research on the role international trade plays within this theory. Obstfeld and Rogoff (1996) thoroughly analyze the mechanisms through which shocks are transmitted between countries and highlight the crucial role played by international financial linkages in shaping global business cycles. Building upon these insights, Gourinchas and Obstfeld (2012) dive into the complexity of global business cycles. They place particular emphasis on financial factors such as capital flows and exchange rates, providing an in-depth analysis of how these elements can transmit economic shocks across borders. Adding to the discourse, Waugh (2014) conducted extensive empirical research on the synchronization of business cycles across countries. His work highlights the role of trade linkages in the transmission of shocks, advancing our comprehension of the substantial impact that trade can have on global business cycles. The influence of monetary policy is another crucial aspect of this theory. Corsetti and Pesenti (2001) explored how monetary policy decisions in one nation can echo through international channels and affect economies globally. These contributions show the affect that domestic economic shocks can have cross boarders, which results in a more complex framework to forecast trade flows.

The international business cycle theory suggests that shocks, such as financial crises or changes in global demand, can significantly impact economic stability in different countries. These shocks can disrupt trade patterns, affect market conditions, and influence the behavior of economic agents involved in international trade. As a result, accurate forecasting of trade flows requires econometric models that are capable of capturing and adapting to sudden exogenous changes, which is where robust estimators and structural break analysis come into play. Structural break analysis helps identifying whether certain events indeed impact the trade flow.

2.2 Forecasting Models Trade Flow

Existing research has extensively investigated the utilization of various statistical models to forecast trade flow, primarily emphasizing accuracy and predictive efficacy. Prominent models employed in prior studies include the gravity model, which considers the distance between trading partners, and the dyadic panel data model, incorporating time-varying variables (Anderson, 2011; Tinbergen, 1962). Additionally, nonlinear least squares, semiparametric, and nonparametric models have been applied to analyze trade flows (Helpman, Melitz & Rubinstein, 2008b). More recently, mixed effects models have garnered attention in the literature, enabling the analysis of relational data by accounting for actor-specific and relation-specific effects (Ghisletta, 2015; Brauer & Curtin, 2017; Simpson, Bahrami

& Laurienti, 2019). Relational data represent measures of association between pairs of actors. Since trade flows inherently involve two parties, it is considered relational data.

The study by Marrs et al. (2023) has introduced a new approach to estimate the regression with relational arrays, namely the exchangeable estimator. In their paper, they provided an in-depth analysis that shows superiority in accuracy of the exchangeable estimator compared to both ordinary least squares and dyadic clustering estimator. This research builds upon their findings. While previous research has explored the application of mixed effects models in analyzing relational data, limited studies have specifically examined the exchangeable estimator and its performance in the context of economic instability. The study by Marrs et al. (2023) investigated the use of the exchangeable estimator for trade flow prediction but did not explicitly evaluate its robustness during periods of economic instability.

2.3 Robustness Testing

Robustness testing is crucial for assessing the performance of an estimator or statistical method under different scenarios or assumptions. Conducting robustness tests enhances the validity and generalizability of research findings, ensuring that conclusions are not driven by unrealistic assumptions. Popular methods for robustness testing include evaluating the estimator under different conditions and benchmarking against alternative models. Changes in the data structure require careful examination of the performance of the estimator. Conducting structural break analysis in time series allows to identify these abrupt changes. While not a direct test of robustness, structural breaks analysis allows for assessing the sensitivity of a statistical model to changes in the underlying data generating process (Page, 1955). By examining structural breaks, researchers can determine whether estimates from a statistical model remain stable across different time periods or exhibit sensitivity to changes in the data. If estimates are found to be unstable or sensitive to structural breaks, it may suggest that the model is not robust to changes in the data generating process, necessitating additional tests or modifications. Therefore, the first crucial step to measure the robustness of the estimates includes testing for possible structural breaks using domain knowledge, as for this research the year 1989. Structural break analysis has been successfully applied in various fields (Casini & Perron, 2018; Andreou & Ghysels, 2009; Reeves, Chen, Wang, Lund & Lu, 2007). The Chow test, introduced by Chow (1960), has emerged as a prominent method for investigating structural breaks and remains extensively employed in various disciplines. This test is particularly effective in detecting singular structural breaks within time series data. Despite the widespread utilization of the Chow test in diverse domains, no academic research to date has specifically examined the testing of economically disturbing events as a structural break in the context of international trade forecasting.

2.4 The Fall of the Berlin Wall

The year 1989 is considered as structural break point in the trade flow data, due to the occurrence of a pivotal event in world history which marked the destruction of the Berlin Wall. Similar to the financial crisis in 2008, the COVID-19 pandemic, and the Russia-Ukraine conflict, the fall of the Berlin Wall was an exogenous event that had a significant impact on international trade. The Berlin Wall, which divided East and West Berlin, was a symbol of the Cold War and the division between the capitalist West and the communist East. Its fall on November 9, 1989, marked a turning point in history and had far-reaching effects on various aspects, including international trade (Djankov & Nikolova, 2016).

The international trade flows were impacted through various ways. Firstly, it led to market integration. Due to the fall, East and West Germany was reunified and created a single market (Wacziarg, Spolaore & Alesina, 2003). The integration created new opportunities for trade and investments in Germany, which led to an acceleration of economic activity and cross-border commerce (Redding & Sturm, 2008). Furthermore, this historic event contributed to the broader process of European integration, as it ultimately led to the expansion of the European Union to include former Eastern Bloc countries (Becker, Egger & von Ehrlich, 2010; Kaminski, 2001). The enlargement of the European Union facilitated trade liberalization. Besides, the Eastern Bloc countries underwent political transitions and changed from centrally planned economies to market-oriented economies (Wolf et al., 1999; Blanchard, Froot & Sachs, 1994). The countries dismantled their trade barriers and opened up their markets. As a result, international trade with these countries increased significantly as they became attractive destinations for foreign investment and trade partnerships. In addition, the subsequent collapse of the Soviet Union led to expansion of global supply chain (Djankov & Freund, 2002). The integration of Eastern European economies into the global market opened up new sourcing and production opportunities for businesses. Lastly, the change in trade relations with the former Soviet Union resulted in economic reforms (Djankov & Freund, 2002). Trade barriers were reduced and new opportunities for trade emerged. Subsequently, Western countries had increased access to vast markets of the former Soviet Union and trade relations were established or expanded.

Overall, the fall of the Berlin Wall played a pivotal role in reshaping the geopolitical landscape and had a profound impact on international trade. Consequently, the sum of the import and export as percentage of the gross domestic product increased significantly globally after the fall due to liberalization and globalization (Aiyar et al., 2023). This supports the following hypothesis.

Hypothesis 1: *The year 1989 marks a structural break in the trade flow.*

2.5 Non-linear Prediction Models

In addition, testing for robustness includes comparing different model specifications using the same data. Initially, the trade flows are predicted using a linear model. The robustness is examined by comparing the linear predicting model to a non-linear model. For this research, a machine-learning method is employed to capture the non-linear effect in the model.

Machine learning methods has gained their popularity to model nonlinear dynamics in data (Pathak, Hunt, Girvan, Lu & Ott, 2017; Lusch, Nair, Gao, Bettencourt & Buell, 2018; Sapsis, 2018; Takeishi, Fujii & Iwata, 2017; Pathak, Lu, Hunt, Girvan & Ott, 2018; Zlotnik & Khasanova, 2019; Zhou & Wu, 2020; Pathak, Wikner, Fussell, Chandra & Girvan, 2018). However, machine-learning algorithms have little applications with relational arrays. Schlichtkrull et al. (2018) showed significant improvement in modelling relational data using the convolutional networks as classifier.

RF has gained significant popularity as a widely utilized machine learning algorithm due to its effectiveness in addressing high-dimensional data (Guo, Hastie & Tibshirani, 2007), non-linear relationships, and noisy or missing values. It is known for its robustness, interpretability, and ability to handle a wide range of problem types, including classification and regression. RF predominantly serves as a classifier (D. Cutler et al., 2007), and outperforms other machine-learning methods (Prasad, Iverson & Liaw, 2006a).

Initially, RF is not applied to the field of forecasting. However, Ciner (2019) finds that RF is an efficient method to produce accurate forecasts, as it can model nonlinear dynamics in data. RF, with its ensemble of decision trees and inherent randomness, can capture and learn from non-linear and complex relationships between covariates. This flexibility allows RF to adapt to unexpected events and capture patterns that may not be evident in a linear model. Several research has proven the superior accuracy of RF over linear model (D. R. Cutler et al., 2007; Prasad, Iverson & Liaw, 2006b; Biau & Scornet, 2016). Therefore, the RF could outperform the linear model in times of economic instability. Nonetheless, no prior research has been conducted to apply RF to relational arrays. Therefore, the following hypothesis is formulated.

Hypothesis 2: *Random Forest outperforms the linear model due to its flexibility to adapt to unexpected events.*

To summarize, previous research has investigated statistical models for trade flow prediction, but limited attention has been devoted to the exchangeable estimator and its performance during times of economic instability. Furthermore, no specific investigations have been conducted to test for structural breaks associated with the fall of the Berlin Wall in the context of predicting international trade. While, multiple studies have proven the effectiveness of RF compared to linear model, the machine-learning model has not been applied to relation arrays. This research aims to address this gap in knowledge and contribute to the existing literature on the usage of statistical models in predicting economic outcomes.

3 Data

The data used for trade predictions in this paper is like Marrs et al. (2023) extracted from the dataset provided by Westveld and Hoff (2011), due to its completeness and availability. This dataset contains international trade flows between 58 countries in the period 1981-2000. The countries included in the dataset are listed in Appendix B. The dataset consists of 66,120 data points. Each data point includes information about the countries involved in the trade that is of relevance for the predictions. This information consists of the gross domestic product of each country and the distance between the countries. Also, the polity is giving which measures the nation's level of democracy. The level ranged on a scale of 0 to 20 from highly authoritarian regimes to highly democratic. Furthermore, the cooperation in conflict of each country is presented which measures active military cooperation. The value is a positive one if the two countries cooperated on a particular dispute, and a negative one if they were on opposite sides. In Table 1, the descriptive statistics are given. The descriptive statistics provide an overview of the central tendency, variability, and range of values for each variable. Obviously, the descriptive statistics of log GDP of the export countries is equivalent to the log GDP of the import countries. Similarly, the descriptive statistics of polity is equal.

Table 1: Descriptive statistics of international trade flows dataset

	Mean	St. Dev	Min.	Max.
Log trade	15.72	5.73	0.00	26.17
Log GDP export/import country	25.02	1.89	21.21	29.91
Polity export/import country	16.08	5.92	0.00	20.00
Cooperation in conflict	0.04	0.26	-3.00	4.00

Previous research only included the GDP of the corresponding countries and the distance (Westveld, 2007). However, including both the polity and the cooperation in conflict provides crucial information on the relation between the export and import country. As Westveld and Hoff (2011) have proven, including these variables results in a more accurate prediction of the trade flows.

4 Methodology

This section provides an explanation of the methodologies employed in this study. First, the most competent estimators for regressions with relational arrays are compared. The comparison of estimator performance begins by outlining the specifications of the estimators used. Subsequently, the simulation methodology is described, which is applied to generate synthetic data for the comparison. Following this, the application of linear regression for trade flow prediction is defined, including the relevant variables. Lastly, the specification of the RF algorithm for trade flow forecasting is presented, elucidating the specific parameters, and considerations utilized in the analysis.

4.1 Estimators for Linear Regression

To begin, the estimates for the coefficients, in case of a relational array as dependent variable, are calculated following two approaches. These approaches entail the dyadic clustering estimator and the exchangeable estimator as formulated by Marrs et al. (2023). Initially, the ordinary least squares estimator is presented as it serves as the fundamental basis for the subsequent estimators. The estimates obtained through the exchangeable estimator are compared to those derived from the dyadic clustering estimator. Likewise to the exchangeable estimator, the dyadic clustering enables the analysis of relation data. The difference lays in the allowance of overlapping pairs for the exchangeable estimator.

4.1.1 Ordinary Least Squares

The primary goal of the estimators is to capture the effects of the exogenous covariates on the values of the relation array, the dependent variable, in a linear regression,

$$y_{ijr} = \beta^\top x_{ijr} + \xi_{ijr}, \quad (i, j = 1, ..., m; i \neq j; r = 1, ..., R), \tag{1}$$

where y_{ijr} is the measure of the rth relation between actor i to actor j, x_{ijr} is a vector of covariates, and ξ_{ijr} is the unobserved random error. The β should be estimated for which multiple functions are formulated. An unbiased estimator for β is the ordinary least squares estimator,

$$\widehat{\beta} = (X^\top X)^{-1} X^\top y, \tag{2}$$

where the variables have the same specification as in (1). The least squares estimator is the best linear unbiased estimator for β when the covariance matrix $\Omega = \text{var}(y \mid X)$ is proportional to the identity matrix. If Ω is known, the best unbiased estimator is the generalized least squares estimator,

$$\widehat{\beta}_{GLS} = (X^\top \Omega^{-1} X)^{-1} X^\top \Omega^{-1} y. \tag{3}$$

Usually, Ω is unknown and should be estimated. The distribution of the estimator β is approximated as the multivariate normal random variable and the confidence intervals is determined by estimating the variance. For the ordinary least squares, the estimator for the variance is

$$\mathrm{var}(\widetilde{\beta} \mid X) = (X^\top X)^{-1} X^\top \Omega X (X^\top X)^{-1}, \tag{4}$$

in which $\widetilde{\beta}$ is a vector of regression coefficients, X is a design matrix, Ω is a covariance matrix, and $\mathrm{var}(\widetilde{\beta} \mid X)$ is the variance of the coefficients conditional on the design matrix. Furthermore, the variance of the generalized least squares is formulated as

$$\mathrm{var}(\widetilde{\beta}_{GLS} \mid X) = (X^\top \widetilde{\Omega}^{-1} X)^{-1} X^\top \widetilde{\Omega}^{-1} \Omega \widetilde{\Omega}^{-1} X (X^\top \widetilde{\Omega}^{-1} X)^{-1}, \tag{5}$$

where $\widetilde{\Omega}$ is the final estimate of Ω from the procedure of generalized least squares.

4.1.2 Dyadic Clustering Estimator

Moving on with the dyadic clustering estimator, the method described by Marrs et al. (2023) is applied. As mentioned earlier, the exchangeable estimator is closely related to the dyadic clustering estimator, as this estimator is also applied to relational arrays. Therefore, the performance of the exchangeable estimator will be compared to this estimator. The estimate for the variance of the dyadic clustering estimator is

$$\widehat{\mathrm{V}}_{DC} = (X^\top X)^{-1} X^\top \widehat{\Omega}_{DC} X (X^\top X)^{-1}, \tag{6}$$

in which $\widehat{\Omega}_{DC}$ denote the covariance matrix of ξ, subject to non-overlapping pair independence assumption. Their methodology proposed a flexible standard error estimator for relational regression. The method makes the assumption that two relations (i, j, r) and (k, l, s) are independent if (i, j) and (k, l) do not share an actor. This assumption implies that the covariance between the outcome variables for non-overlapping relation pairs is zero, but places no restrictions on the covariance elements for pairs of relations that share an actor. Fafchamps and Gubert (2007) propose estimating each non-zero entry of Ω_{DC} with a product of residuals. Specifically, they use $e_{ijr}e_{iks}$ to estimate $cov(\xi_{ijr}, \xi_{iks})$, where $e_{ijr} = y_{ijr} - \beta^T x_{ijr}$ is the residual associated with the outcome variable y_{ijr} and x_{ijr} is the corresponding row of the design matrix. The estimator $\widehat{\Omega}_{DC}$ can be seen as taking the empirical covariance of the residuals defined by ee^T, where e is a vector of the set of residuals e_{ijr}, and introducing zeros to enforce the non-overlapping pair independence assumption. This yields a covariance matrix estimator that is flexible and can be used for various types of relational data.

4.1.3 Exchangeable Estimator

Lastly, the prime estimator of this research is defined. The paper of Marrs et al. (2023) proposes an exchangeable estimator of the variance-covariance matrix of the regression coefficients in relational regression models with Ω_E of the exchangeable form. The estimator is denoted as $\hat{V}_E$, and can be expressed as

$$\widehat{\mathrm{V}}_E = (X^T X)^{-1} X^T \widehat{\Omega}_E X (X^T X)^{-1}, \tag{7}$$

where $\hat{\Omega}_E$ is the estimate of the exchangeable covariance matrix, defined as the sum of $\phi(\eta)$ times a binary matrix S^η, where $\eta = 1, 2$ and S_u^η denotes the Rn(n1)×Rn(n1) binary matrix with 1's in the entries corresponding to relation pairs of type (u = 0, a, b, c, d; η = 1, 2) as defined by Marrs et al. (2023).

The proposed method estimates the ten parameters in Ω_E by averaging the residual products that share the same index configurations. Specifically, the estimate of $\mathrm{cov}(\xi_k ls, \xi_i jr)$, corresponding to $u = b$ and $\eta = 2$, is given by

$$\hat{\phi}_b^{(2)} = \binom{R}{2}^{-1} \frac{1}{n(n-1)(n-2)} \sum_{r \neq s} \sum_{i=1}^{n} \sum_{j \neq i} e_{ijr} \left(\sum_{k \neq j} e_{iks} - e_{ijs} \right), \tag{8}$$

where $e_{ijr} = y_{ijr} - \beta^T x_{ijr}$ is the residual for the rth relation between actors i and j. The remaining nine estimators for $(s = 0, a, ..., e; \eta = 1, 2)$ are defined analogously. The remaining covariance estimates are given in Appendix C. Finally, $\hat{\Omega}_E$ can be interpreted as the projection of $\hat{\Omega}_{DC}$ into the vector space over symmetric matrices of the form of Ω_E.

4.2 Simulation

The performance of the exchangeable estimator will be compared to the dyadic clustering estimator. In order to generate data, a simulation will be executed. For the purpose of conducting the simulation, the methodology outlined in the specification of Marrs et al. (2023) is adopted. Their code serves as a primary reference and is meticulously adhered to, forming the fundamental framework for this research. The data used for the simulation are based on the formula

$$y_{ij} = \beta_1 + \beta_2 1_{(x_{2i} \in C)} 1_{(x_{2j} \in C)} + \beta_3 |x_{3i} - x_{3j}| + \beta_4 x_{4ij} + \xi_{ij}, \tag{9}$$

where the covariate matrix is generated using the total number of members in the simulation. The covariate matrix consists of x_{1ij}, which is set to 1, x_{2ij}, that follows Bernoulli(1/2) independently, and x_{3ij} and x_{4ij}, that are drawn from standard normal distribution. Note that the matrix is created using x_{3ij} which equal $|x_{3i} - x_{3j}|$. All betas are fixed to 1 in the simulation. The error term is calculated with an independent and identically distributed errors, a non-exchangeable error model, and an exchangeable error model. For every error model, the random noise follows a normal distribution and is used to calculate the total error term, which differs per model. The variance of the error terms are set equal to

$$\sum_{ij} \mathrm{var}(\xi_{ij}) = 3n(n-1). \tag{10}$$

For the independent and identically distributed errors setting, it follows $\xi_{ij} \sim_{iid} \mathcal{N}(0,3)$ for all (i,j). The errors for the non-exchangeable error setting may be written as

$$\xi_{ij} = \tau 1_{(i \leq \lfloor n/2 \rfloor)} 1_{(j \leq \lfloor n/2 \rfloor)} + \epsilon_{ij}, \quad \tau \sim \mathcal{N}(0, \frac{9n}{4\lfloor n/2 \rfloor}), \quad \epsilon_{ij} \sim_{iid} \mathcal{N}(0, 3/4), \tag{11}$$

where $1_{(j \leq \lfloor n/2 \rfloor)}$ is an indicator of index i less than or equal to the floor of $n/2$.

Lastly, the specification for exchangeable error setting is described in Marrs et al. (2023) and follows

$$\begin{gathered} \xi_{ij} = a_i + b_j + z_i^T z_j + \gamma_{ij} + \epsilon_{ij}, \quad (a_i, b_i) \sim \mathcal{N}(0_2, \Sigma_{ab}), \\ \Sigma_{ab} = \begin{bmatrix} \sigma_a^2 & \rho_{ab}\sigma_a\sigma_b \\ \rho_{ab}\sigma_a\sigma_b & \sigma_b^2 \end{bmatrix}, \\ z_i, z_j \sim \mathcal{N}_d(0, \sigma_z^2 I_d), \quad \gamma_{ij} = \gamma_{ji} \sim \mathcal{N}(0, \sigma_\gamma^2), \quad \epsilon_{ij} \sim \mathcal{N}(0, \sigma_\epsilon^2). \end{gathered} \tag{12}$$

The dimension of the latent space, d, is set to 2, the correlation between sender and receiver effects set to $\rho_{ab} = 1/2$, and the sender variance set to be twice that of receiver variance, i.e. $\sigma_a^2 = 2\sigma_b^2$. Furthermore, $\sigma_z = \sigma_\gamma = \sigma_b$ and $\sigma_\epsilon^2 = 3/4$.

In total, 500 sets of random covariates are generated for relational data, with varying numbers of actors (n = 10, 20, 30, 40). For each set of covariates, we generated 1,000 random realizations of errors under three different data generators: independent and identically distributed, exchangeable, and non-exchangeable. For each combination of covariates and error realizations, a new dataset is simulated using model (9).

Next, a regression model is fitted using the list of nodes to each simulated dataset using ordinary least squares. The list of nodes consists of dyad indices for shared membership without overlaps. The standard errors are measures using both exchangeable and dyadic clustering sandwich variance estimators, explained in 4.1.2 and 4.1.3. To assess the performance of the estimators, the coverage of confidence intervals for each covariate realization is estimated by calculating the fraction of confidence intervals that contained the true coefficient.

Additionally, the bias and variance is estimated of the standard error estimators for each covariate realization relative to the known true standard errors of the ordinary least squares estimator, given the specific covariate realization.

4.3 Forecasting Trade Flow

Subsequently, the performance of the exchangeable estimator is measured by comparing the coefficient of determination, resulting from one-step ahead forecasting of trade flows, to ordinary least squares. To execute this comparison, the data from Westveld and Hoff (2011) is used. To forecast the trade flow, the logarithm of yearly international trade flow data are used in a modified gravity mean model as linear function of seven covariates, defined in a linear regression proposed by Westveld and Hoff (2011). The linear regression follows

$$\begin{aligned}\ln(Trade_{i,j,t}) =& \beta_{0,t} + \beta_{1,t}\ln(GDP_{i,t}) + \beta_{2,t}\ln(GDP_{j,t}) + \beta_{3,t}\ln(D_{i,j,t}) + \beta_{4,t}Poli_{i,t} \\ &+ \beta_{5,t}Poli_{j,t} + \beta_{6,t}CC_{i,j,t} + \beta_{7,t}Poli_{i,t} \times Poli_{j,t} + \Xi_t,\end{aligned} \tag{13}$$

where $GDP_{i,t}$ denotes the gross domestic product, $Pol_{i,t}$ denotes the polity, $CC_{i,t}$ the cooperation of conflict, and $D_{i,t}$ the distance between the two countries all of country i at time t. The error term Ξ_t is a vector of errors for year t, which specification varies for using ordinary least squares estimates and exchangeable estimates. Ξ_t and Ξ_{t+h} are assumed to be independent.

First, the model is estimated using the feasible generalized least squares procedure. The β is initialized using ordinary least squares and the Ω is estimated using the residuals under the assumption of joint exchangeability. Dissimilar to the simulation, where the data was fitted using ordinary least squares and the variance was estimated afterwards using the sandwich method with the exchangeable estimator specification. Now, the exchangeable estimator for the error model is applied in combination with generalized least squared procedure, for which the formula in (7) is used. Then, iteratively β_{GLS} in (3) is re-estimated by setting Ω equal to $\hat{\Omega}$ until convergence. Specifically, convergence is under the condition that

$$\begin{gathered}|Q^\gamma - Q^{\gamma-1}| < \epsilon, \\ Q^\gamma = (y - X\widehat{\beta})^T(\widehat{\Omega}^\gamma)^{-1}(y - X\widehat{\beta}), (\gamma = 1, 2, ...),\end{gathered} \tag{14}$$

where $\widehat{\beta}^\gamma$ and $\widehat{\Omega}^\gamma$ are estimators of the regression coefficient and error matrix at the γth iteration. Furthermore, $\epsilon = 10^{-6}$. The final Ω is set to $\widetilde{\Omega}$ in (5) to obtain the standard errors.

The one-step ahead trade is predicted using the first 4 through 11 years of data, with a rolling window approach. The conditional expectation $E(y_T|\{y_t\}_{t=1}^{T-1})$ is calculated to generate predictions from the model, based on the assumption that y_T and (y_r), for $r = 1, ..., T-1$, are jointly normal.

Using the ordinary least squares estimator this results in

$$E(y_T|\{y_t\}_{t=1}^{T-1})_{OLS} = X_T\tilde{\beta}_{T-1}. \tag{15}$$

For the exchangeable procedure, we likewise set $\beta_T = \beta_{T-1}$. The prediction for the exchangeable procedure is defined as

$$E(y_T|\{y_t\}_{t=1}^{T-1})_E = X_T\tilde{\beta}_{T-1} + \Omega_2\{\Psi_1 + (T-2)\Psi_2\}\sum_{t=1}^{T-1}(y_t - X_t\tilde{\beta}_T), \tag{16}$$

where $\Omega_1 = \text{var}(y_t)$ and $\Omega_2 = \text{cov}(y_t, y_{t+h})$ for all h. Taking $z_{T-1}^T = (y_1^T, y_2^T, ..., y_{T1}^T)$ as the concatenated vector, it is of the same patterns as the variance $\text{var}(z_{T-1})^{-1}$. The variance of z has Ω_1 along the diagonal blocks and Ω_2 on the off-diagonal blocks. In the equation Ψ_1 defines the diagonal blocks and Ψ_2 the off-diagonal blocks.

The R-squared, also known as the coefficient of determination, is used to compare the performance of the estimators. It is a statistical measure used to assess the goodness-of-fit of a regression model. It provides an indication of how well the model's predictions fit the observed data points. The R-squared value ranges from 0 to 1, where 0 indicates that the model explains none of the variability in the data, and 1 indicates that the model perfectly predicts the observed outcomes. However, a R-squared of 1 is not desired as it could signal overfitting. A higher R-squared value suggests a better fit of the model to the data. The formula for R-squared is given by

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \tag{17}$$

where n represents the number of data points, y_i represents the observed values of the dependent variable, $\hat{y}_i$ represents the predicted values of the dependent variable from the regression model, and $\bar{y}$ represents the mean of the observed values of the dependent variable.

4.4 Structural Break Analysis

The robustness of the exchangeable estimator is tested based on evaluating the estimates during instability in the data. Accordingly, structural break analysis is used to identify abrupt change in the data in 1989, after the fall of the Berlin Wall.

To detect structural breaks in the relationship between time and international trade flows, the Chow test is used. The Chow test is a widely-used test for detecting structural breaks in time series data (Chow, 1960). The Chow test begins with the formulation of null and alternative hypotheses. The null hypothesis posits that there is no structural break or difference in the relationship between the variables of interest across sub-samples or time periods. The alternative hypothesis suggests the presence of a structural break. The dataset is divided into distinct sub-samples based on year before or after the fall of the Berlin Wall. Separate regression models are estimated for each sub-sample, capturing the relationship between the variables within each segment. This division enables the examination of potential structural breaks in the relationship across different sub-samples. The sum of squared residuals is calculated for each sub-sample, representing the discrepancies between the observed values and the predicted values obtained from the respective regression models. This step quantifies the overall fit of the models within each sub-sample. A combined regression model is constructed by pooling together all the observations from the sub-samples. The sum of squared residuals is computed for this combined model, which provides an overall measure of the model's fit when considering the entire dataset. The Chow test statistic is calculated by comparing the sum of squared residuals from the separate sub-sample models to the sum of squared residuals from the combined model. The test statistic follows an F-distribution under the assumption of no structural break. The calculated Chow test statistic is compared to the critical value from the F-distribution at a chosen significance level. If the test statistic exceeds the critical value, the null hypothesis of no structural break is rejected in favor of the alternative hypothesis, indicating evidence of a significant difference in the relationship between the variables across sub-samples or time periods. The `strucchange` library in R is used to perform the test.

4.5 Random Forest

Next, the RF machine-learning algorithm introduced by Breiman (2001) is incorporated to predict the trade flows using a non-linear model. The model fit is compared to the estimates resulting from the linear model described in (13), to measure the robustness. RF model is an ensemble learning algorithm that combines multiple decision trees to make predictions. Each decision tree in the RF is constructed independently by using a subset of the training data and a random subset of the input features. The predictions from individual trees are then combined to reach the final result.

The RF model follows the following steps. First, at each node m of a decision tree, a random subset of features is selected from the available input features. This helps in reducing the correlation between trees and introduces randomness in the model. Next, the RF model uses a technique called bootstrapping, where a random subset of the training data is selected with replacement. This means that some samples may be repeated in the subset, while others may be left out. This technique helps create diverse subsets of the training data for each decision tree. Then, using the selected subset of features and the bootstrapped subset of training data, a decision tree T_m is built by recursively splitting the data based on the selected features. The splitting criterion used mostly in each decision tree is based on minimizing the mean squared error (MSE) at each node. The MSE is calculated as the following

$$\text{MSE}_{RF} = \frac{1}{n_m} \sum_{i \in D_m} (y_i - \bar{y}_m)^2, \tag{18}$$

where n_m is the number of samples in node m, D_m is the set of samples in node m, y_i is the target value of the ith sample, and $\bar{y}_m$ is the mean target value of the samples in node m.

Once all the decision trees are constructed, predictions from each tree are combined to obtain the final prediction. For regression tasks, the final prediction $\hat{y}$ is calculated as the average of the predictions from all the trees by

$$\hat{y} = \frac{1}{T} \sum_{i=1}^{T} \hat{y}_i, \tag{19}$$

where $\hat{y}_i$ represents the prediction from the ith tree, and T is the total number of trees.

For the implementation, the `randomForest` package in R is used. Furthermore, several model parameters are decided on before fitting the model. Most importantly, number of trees and the splitting criterion. For this problem, the number of trees is 250. The model is tested on 500 trees as well which did not significantly improved the results, but it did increase the run time. For that reason, 250 trees is chosen. Regarding the splitting creation in the `randomForest` package, the algorithm minimizes the MSE as introduced earlier. One-step ahead predictions for trade are forecasted in a rolling window for each combination of countries, in which the previous years are used to train the model.

Equal dataset as employed for the linear model (13) is used. RF cannot handle categorical predictors with more than 53 categories. Therefore, one-hot encoding is used to control for the 58 export/import countries. One-hot encoding is a technique used to convert categorical variables into a numeric representation that can be used by machine learning algorithms (Dinesh, 2019). It creates binary variables to represent each category of the original variable. Furthermore, it is important to note that the algorithm has to make the predictions for the all the combinations of actor i and j with $i \neq j$, referring to the countries. Therefore, the combinations are prespecified and used to prepare the test data.

The results from the RF are compared on predictive power to the regular linear model in (13), based on the R-squared (17) and the mean squared error,

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2, \tag{20}$$

where n is the total number of predictions, y_i is the real value, and $\hat{y}_i$ is the predicted value.

5 Results

5.1 Simulation

In this section, the results are presented of the simulation formulated in 4.2. Figure 1 presents the estimated probability that true coefficient is in 95% confidence interval for each of three covariates.

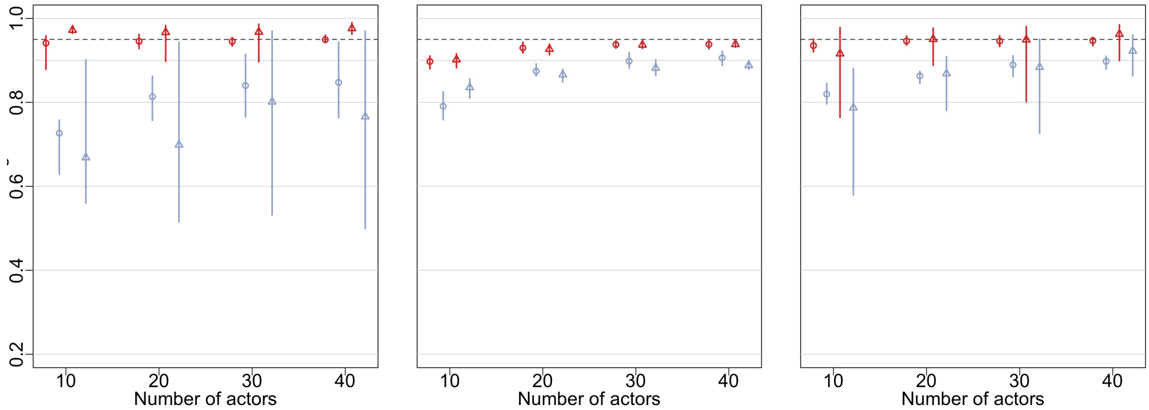

Figure 1: Estimated probability that true coefficient is in 95% confidence interval for each of three covariates (binary, positive, and real-valued). The circles represent the errors that are generated from exchangeable and the triangles represent the non-exchangeable error models. Points denote mean estimated coverage and lines represent the middle 95% of coverages for exchangeable and dyadic clustering estimators, in red and blue subsequently.

As shown in Figure 1, it is observed that the estimated probability of the true coefficient falling within the 95% confidence interval is examined for each of the three covariates, namely binary, positive, and real-valued, under two different error models, the exchangeable model and the non-exchangeable model. The mean estimated coverage is indicated by the points, while the lines illustrate the middle 95% of coverages for both the exchangeable and dyadic clustering estimators.

The findings reveal that the exchangeable estimator exhibits superior performance compared to the dyadic clustering estimator. The estimated probabilities of the true coefficients lying within the 95% confidence interval are consistently higher for the exchangeable model, as indicated by the points denoting the mean estimated coverage. Moreover, the middle 95% range of coverages, represented by the lines, consistently demonstrates better performance for the exchangeable estimator in contrast to the dyadic clustering estimator. Furthermore, for both models the middle 95% of coverages are smaller when using the exchangeable error model compared to the non-exchangeable error model. This shows the effectiveness of incorporating exchangeability in the error term if the dependent variable is a relational array.

These results collectively suggest that the exchangeable estimator outperforms the dyadic clustering estimator in accurately capturing the true coefficients' probabilities within the confidence interval for both the binary, positive, and real-valued covariates.

These results are in line with those from Marrs et al. (2023). They ran the simulation for $n = 20, 40, 80, 160, 320$. However, due to time limitation the node sizes were reduced. As expected, the results from $n = 10, 30$ are in line with the outcomes presented in Marrs et al. (2023). In contrary to their results, the middle 95% of coverages for the binary covariate in this simulation is noticeable smaller when error are generated using exchangeable model in combination with the exchangeable estimator. This could be due to a difference in seed specification.

5.2 Trade

Next, the trade is predicted as explained in 4.3. This section presents the findings regarding the prediction of trade using the exchangeable estimator and the ordinary least squares (OLS) estimator. The performance of these two models is evaluated based on the R-squared statistic, which provides an indication of the proportion of variance in the dependent variable explained by the independent variables.

The analysis reveals that the exchangeable estimator demonstrates superior predictive performance compared to the OLS estimator, as shown in Graph 2. The R-squared values obtained from the exchangeable estimator consistently outperform those obtained from the OLS estimator across different trade scenarios. Specifically, the exchangeable estimator exhibits higher R-squared values, indicating a better fit of the model to the observed trade data.

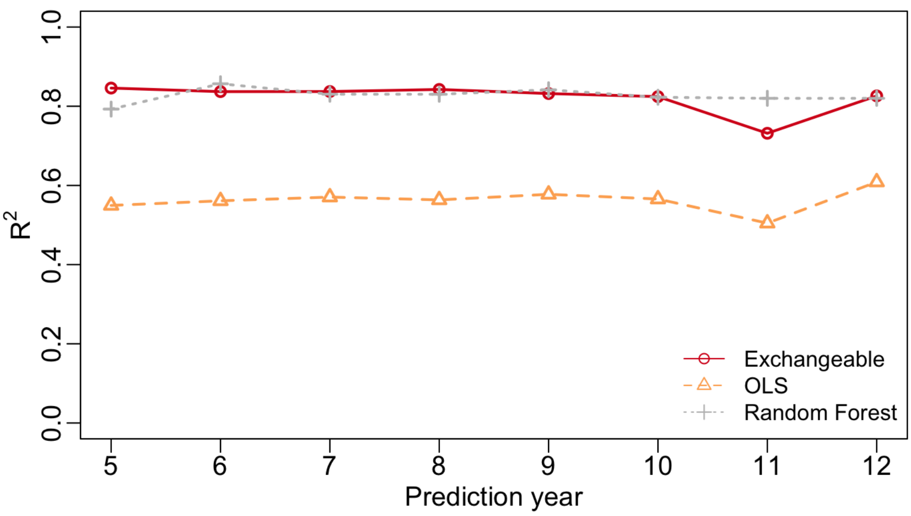

Figure 2: R-squared using exchangeable and ordinary least squares approaches when predicting one-year-ahead trade flows.

For instance, when predicting year 6 the exchangeable estimator yields an R-squared over .84, while the OLS estimator achieves an R-squared of .56. Similarly, as more years are included in predicting, i.e. year 12, the exchangeable estimator obtains an R-squared of .83, surpassing the OLS estimator's R-squared of .61. The R-squared values obtained from the exchangeable estimator remain stable and relatively high across various time periods. These results consistently demonstrate that the exchangeable estimator captures a larger proportion of the variation in trade compared to the OLS estimator. It is worth noting that the disparity of the estimators in performance diminishes with increasing the number of time periods observed.

Noticeable, is the drop in both R-squares in year 11. The R-squared of the OLS estimator dropped from .57 to .51, and increased to .61 again. Similarly, the R-squared of the exchangeable estimator dropped from .82 to .73, and increased to .83 again. The percentage change of the R-squared of the exchangeable estimator is larger than of the OLS, namely -11.17% and -10.78% subsequently. Also, the rise afterwards was stronger for the OLS estimator, namely 20.59% versus 12.84%.

Although, the true reasoning behind the drop in model fit is unknown, sudden changes in the data could be a cause. Since the fall of the Berlin Wall happened in year 9, the impact on international trade will most likely be evident later and could very well be the cause for the drop in model fit accuracy in year 11.

These findings emphasize the superiority of the exchangeable estimator in predicting trade compared to the OLS estimator. The results are completely in line with the results from Marrs et al. (2023). It clearly shows the better fit of the model observed when using the exchangeable estimation. The higher R-squared values obtained from the exchangeable estimator indicate a better fit of the model to the observed trade data, suggesting that it captures a larger proportion of the underlying trade dynamics. This highlights the importance of considering the relational structure and actor-specific effects, which the exchangeable estimator incorporates, in accurately predicting trade.

5.3 Structural Break

The Chow test was conducted to examine the presence of a structural break in the data at the time of the fall of the Berlin. The test aimed to determine if there was a significant change in the relationship between the trade, and the independent variables specified in section 4.3, at a specific point in time. The Chow test statistic was calculated to be 34.38 with associated p-value $< 2.2e\text{-}16$. This p-value indicates extremely strong evidence against the null hypothesis, suggesting that a structural break exists in the data.

The significant result suggests a substantial change in the relationship between the variables after 1989. This finding implies that there are different relationships and dynamics influencing the dependent variable before and after the break point, indicating a structural change in the underlying factors driving the observed patterns. The underlying factor for this change could be the fall of the Berlin Wall as has been argued in Section (2).

5.4 Random Forest

As explained earlier, RF is implemented to compare the linear model with a non-linear model. As the fall of the Berlin Wall marks a significant change in the dataset, only predictions afterwards are compared. However, first the overall performance of the model is evaluated for $t = 4, ..., 11$, to predict $t + 1$.

In Graph 2, the resulting R-squared of the one-year ahead forecasting are presented. The R-squared results are measures of the goodness of fit for the one-year ahead forecasts. These values suggest that the RF model used for the forecasts generally explain a substantial portion of the variance in the dependent variable. The R-squared values range from .79 to .86, indicating a good fit overall. In addition, the model better compared to OLS estimator and mostly equal to the exchangeable estimator. As for the drop in the 11th prediction estimator, a more in-depth analysis is given in section (5.5).

5.5 Comparison

Moving on, the results from the exchangeable estimator approach are compared to RF algorithm. To begin, the R-squared are analyzed focusing on the years around the fall of the Berlin Wall. The chosen timeframe extends up to five years after the fall, specifically spanning from $t = 8$ to $t = 14$.

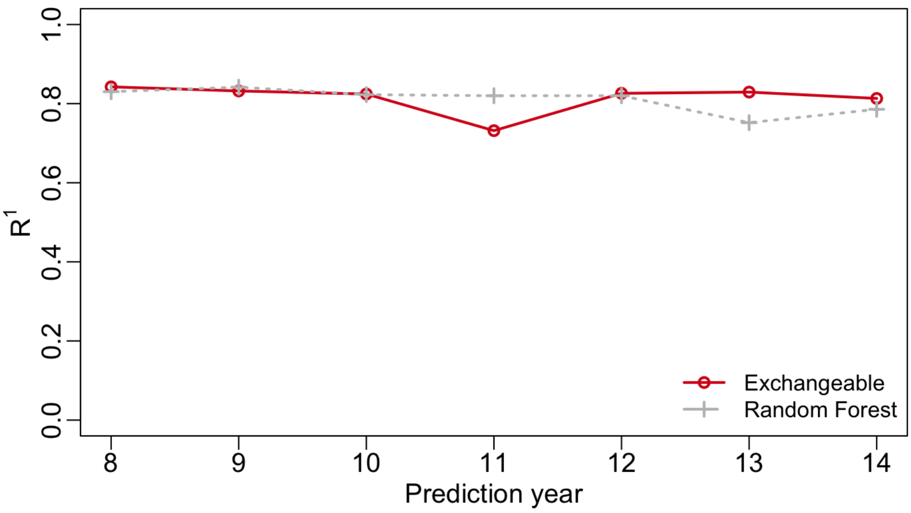

Figure 3: R-squared using exchangeable estimator and Random Forest approaches when predicting one-year-ahead trade flows.

In Graph 3, the results are presented of the model fit using the exchangeable estimator and RF approaches. Overall, the fit is similar in year 10 and 12. However, the R-squared remains stable for the RF approach while there is a slight drop in prediction year 11 using the exchangeable estimator, as was evaluated in 5.2. In addition, the model fit of the RF regression exhibits a modest decline in year 13, whereas the exchangeable estimator is unaffected.

Due to the fact that the fall of the Berlin Wall is a structural break in the time series data, instead of a transient shock, it takes more time for the RF model to adjust. Linear models might be quicker to adapt to a structural break. While tree-based models like RF might take more time to adjust to the new data structure. Furthermore, RF models are prone to overfitting especially when there are high dimensional inputs or noise. The effect of the fall of the Berlin Wall introduced significant noise and complexity into the data, the RF model might have initially overfit to this, maintaining its performance. Nonetheless, as more data came in, it might have struggled to generalize, causing the drop in R-squared. Noticeable, the drop in R-squared is smaller for the RF model compared to the linear model, namely .75 versus .73.

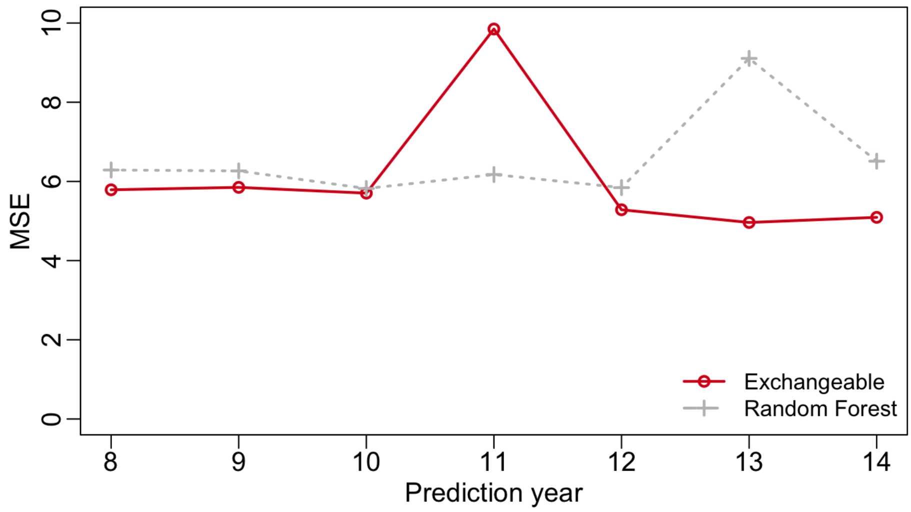

Figure 4: MSE using exchangeable estimator and Random Forest approaches when predicting one-year-ahead trade flows.

Furthermore, the MSE is compared for $t = 8, ..., 14$ to clearly analyze the difference between before and after the fall of the Berlin Wall. Based on these observations, it seems that the model's performance, much like the R-squared performance, varies across different time points. The exchangeable estimator performs poorly in year 10 while the RF model performs poorly in year 12. This suggests that the model's accuracy is influenced by specific factors that affect the trade flows, like the fall of the Berlin Wall. Similar analysis as employed for the R-squared results can be applied to interpret these findings.

6 Discussion

The trade prediction outcomes showed the superior results for the exchangeable estimator compared to the ordinary least squares, as was reveal earlier by Marrs et al. (2023). However, this raises the discussion in what way the exchangeable estimator is generalizable for alternative use cases. The estimator relies strongly on the assumption of exchangeability, which assumes that the observations are identically distributed. However, this assumption may not hold true in all real-life scenarios. Potential violating of this assumption may affect the accuracy and especially the reliability of the estimator.

The data used for this research was extensive and of high quality. It is questionable if this is reasonable for further application of the model. The model relies heavily on the quality of the data and completeness. It is of importance that the data include trade flows from both ways of the exchange. This relates back to the assumption of exchangeability. However, this might be challenging to gather and keep up to date to measure accurate prediction. This calls for further research to apply the method to more recent data.

For this research, the Chow test was conducted to detect a structural break. Limitation of this method include the requirement for domain knowledge to decide on the point to test for in the dataset. The Chow test is sensitive to the assumed break location. If the break is misidentified or the timing is imprecise, the test may fail to detect the true structural change. Conducting sensitivity analyses by testing different break locations can help mitigate this limitation. Also, the Chow test relies on certain assumptions, including independence and identically distributed residuals. Violations of these assumptions, such as autocorrelation or non-normality in the residuals, can affect the test's validity and lead to inaccurate results. Furthermore, the Chow test is designed to detect a single structural break, assuming that the data exhibit a stable relationship before and after the break. However, in some cases, multiple structural breaks may occur. The Chow test may not be suitable for identifying and analyzing multiple breaks, requiring the use of alternative techniques such as the Bai-Perron test or the Quandt-Andrews test. Nonetheless, the Chow test has proven its success in various conditions (Casini & Perron, 2018; Andreou & Ghysels, 2009; Reeves et al., 2007).

Considering the evaluation of robustness, this has to be measured more extensively to draw a valid conclusion. Firstly, various non-linear models exist, especially considering machine-learning algorithms. More algorithms could be evaluated and implemented to obtain the best solution for predicting trade flows, limiting the scope of this comparison. The subjectivity in model selection introduces a degree of uncertainty and may lead to different conclusions depending on the chosen criteria or tests. Also, model comparisons assume that the models are correctly specified and capture the true data-generating process. However, in practice, it is challenging to know the true underlying model. If a model is misspecified or omit important variables or relationships, the comparisons may favor the wrong model. Researchers should carefully consider the model specifications and ensure that they adequately capture the relevant factors and dynamics. Nevertheless, the linear model to predict trade flows has been researched extensively and proven its performance (Westveld & Hoff, 2011). Furthermore, the exchangeable estimator is solely compared to the RF in case of predicting trade flows, which could lead to data-specific results. Therefore, caution should be exercised when extrapolating the results to different contexts. New scenarios involving relational arrays should be considered to have a better understanding which model has superior performance.

While RF provides accurate predictions, the interpretability of the model can be challenging due to the ensemble nature and multiple trees involved. Understanding the exact decision-making process of a RF model can be complex, especially when the number of trees is large. Furthermore, the tuning of the hyperparameter are crucial to achieve optimal performance and avoid overfitting. The hyperparameters that need to be set include the number of trees, maximum depth of trees, and the number of features considered at each split, and should be considered carefully. Future research could explore tuning the hyperparameters differently or consider using alternative packages of programming languages.

Lastly, the MSE is used to compare the predictive power of the linear and non-linear power. However, this holds some limitation. Firstly, the MSE is sensitive towards outliers as it squares the error term, potentially leading to biased comparisons. Alternatively, the median absolute error (MAE) could be considered. By using the median instead of the mean, the MAE is less affected by outliers. However, the dataset used in this research does not contain extreme outliers which mitigates the limitation. For future research in the field of model performance evaluation, robust cross-validation techniques can be employed. Techniques like k-fold cross-validation can be adapted to use robust estimation methods within each fold, helping to mitigate the influence of outliers. This approach provides a more robust estimate of a model's predictive power by considering multiple subsets of the data and reducing the impact of outliers on the overall evaluation.

7 Conclusion

This paper has investigated the robustness of the exchangeable estimator applied to trade flows during time of economic instability. Accordingly, the research question for this paper was: *"Is the exchangeable estimator robust in predicting trade flow during periods of economic instability, such as the aftermath of the fall of the Berlin Wall in 1989, and how does the linear predictive model using the exchangeable estimator perform compared to a non-linear model when economically disturbing events occur?"*

To answer this research question, the exchangeable estimator was first compared to the dyadic clustering estimator by means of a simulation, for which the method of Marrs et al. (2023) was followed. Likewise, the results showed superior performance for the exchangeable estimator, especially in combination with the exchangeable error model versus the non-exchangeable error model. Afterwards, the trade flows were predicted using a linear model formulated as stated in Westveld and Hoff (2011). The generalized least squares method with exchangeable error model was used to estimate the linear model, and was compared to ordinary least squares estimator. By means of the R-squared, the model fit was better for one-year ahead prediction when using the exchangeable estimator, which was also demonstrated in Marrs et al. (2023).

Afterwards, the robustness of the exchangeable estimator was evaluated. First of all, the dataset was tested on a structural break in year 1989, corresponding to the year of the fall of the Berlin Wall. Earlier, literature already stated the significant impact of the fall of the Berlin Wall on international trade flows (Djankov & Nikolova, 2016). The fall impacted the trade in various ways. It led to market integration and contributed to the broader process of European integration, as it ultimately led to the expansion of the European Union to include former Eastern Bloc countries. Additionally, the subsequent collapse of the Soviet Union led to expansion of global supply chain and the changed in trade relations with the former Soviet Union resulted in economic reforms. Consequently, the sum of the import and export as percentage of the gross domestic product increased significantly globally after the fall due to liberalization and globalization (Aiyar et al., 2023). This led to the following hypothesis:

Hypothesis 1: *The year 1989 marks a structural break in the trade flow.*

Structural break analysis was performed to explore this hypothesis. The Chow test, where year 1989 was considered as break point, resulted in a significant p-value. Implicating that their was a structural break in the trade flow data in the corresponding year. Therefore, the first hypothesis is not rejected.

Moving on, the robustness of the exchangeable estimator was evaluated. As of now, the results showed that the exchangeable estimator outperforms alternative estimators for fitting a linear model. Therefore, the performance of the linear model using the exchangeable estimator was compared to a non-linear model, to evaluate the robustness during times of economic instability. As non-linear model, the RF algorithm was introduced. RF has gained popularity as a widely used machine learning algorithm due to its effectiveness in addressing high-dimensional data, non-linear relationships, and noisy or missing values. This flexibility allows RF to adapt to unexpected events and capture patterns that may not be evident in a linear model. Therefore, the following hypothesis was formulated:

Hypothesis 2: *Random Forest outperforms the linear model due to its flexibility to adapt to unexpected events.*

Overall, the RF model performed well when applied to relational arrays, with the R-squared values range from .75 to .84. Furthermore, the linear model using the exchangeable estimator performs equally to the RF algorithm, based on model fit using the coefficient of determination. However, focusing on the aftermath of the fall of the

Berlin Wall, the model fit accuracy dropped slightly for both models, but with a delay for the RF model. Possible explanations for this delayed response could encompass an initial overfitting scenario in the RF model and a potentially longer adjustment period required by the RF to accommodate the altered data structure. Consequently, the second hypothesis is rejected.

To conclude, this research presents a deeper understanding of the performance of exchangeable estimator. The performance of the exchangeable estimator was not drastically impacted by sudden change in the relational data. The model fit dropped slightly, but it quickly recovered. Furthermore, the RF model demonstrates a lag in its response, with a small drop in model fit occurring at a later stage, and therefore does not outperform the linear model.

References

Acemoglu, D., Johnson, S. & Robinson, J. A. (2001, December). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, *91*(5), 1369-1401. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.91.5.1369 doi: 10.1257/aer.91.5.1369

Agence Française de Développement. (2023). *Ukrainian war worsens african food crisis.* Retrieved from https://www.afd.fr/en/actualites/ukrainian-war-worsens-african-food-crisis (Accessed: 2023-06-28)

Aiyar, S., Chen, J., Ebeke, C., Garcia-Saltos, R., Gudmundsson, T., Ilyina, A., ... Trevino, J. P. (2023). Geoeconomic fragmentation and the future of multilateralism.

Anderson, J. E. (2011). The gravity model. *Annual Review of Economics*, *3*, 133-160.

Anderson, J. E. & van Wincoop, E. (2004, September). Trade costs. *Journal of Economic Literature*, *42*(3), 691-751. Retrieved from https://www.aeaweb.org/articles?id=10.1257/0022051042177649 doi: 10.1257/0022051042177649

Andreou, E. & Ghysels, E. (2009). Structural breaks in financial time series. *Handbook of financial time series*, 839–870.

Backus, D., Kehoe, P. J. & Kydland, F. E. (1993, October). *International business cycles: Theory and evidence* (Working Paper No. 4493). National Bureau of Economic Research. Retrieved from https://doi.org/10.3386/w4493 doi: 10.3386/w4493

Backus, D. K., Kehoe, P. J. & Kydland, F. E. (1992). International real business cycles. *Journal of Political Economy*, *100*(4), 745–775. Retrieved from https://doi.org/10.1086/261838 doi: 10.1086/261838

Baldwin, R. (2011). *The great trade collapse: Causes, consequences and prospects.* Centre for Economic Policy Research. Retrieved from https://books.google.nl/books?id=MUfGuQAACAAJ

Baldwin, R. & Tomiura, E. (2020). Thinking ahead about the trade impact of covid-19. In R. Baldwin & B. Weder di Mauro (Eds.), *Economics in the time of covid-19* (p. 7-22). CEPR Press.

Becker, S. O., Egger, P. H. & von Ehrlich, M. (2010, October). Going NUTS: The effect of EU Structural Funds on regional performance. *Journal of Public Economics*, *94*(9-10), 578-590. Retrieved from https://ideas.repec.org/a/eee/pubeco/v94y2010i9-10p578-590.html

Biau, G. & Scornet, E. (2016). A random forest guided tour. *Test*, *25*(2), 197–227.

Blanchard, O., Froot, K. & Sachs, J. (1994, 01). The transition in eastern europe, volume 1. *Bibliovault OAI Repository, the University of Chicago Press*, *55*. doi: 10.2307/2501929

Brauer, M. & Curtin, J. (2017, 11). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, *23*. doi: 10.1037/met0000159

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5–32.

Casini, A. & Perron, P. (2018). Structural breaks in time series. *arXiv preprint arXiv:1805.03807*.

Chor, D. & Manova, K. (2012). Off the cliff and back? credit conditions and international trade during the global financial crisis. *Journal of International Economics*, *87*(1), 117-133. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0022199611000493` (Symposium on the Global Dimensions of the Financial Crisis) doi: https://doi.org/10.1016/j.jinteco.2011.04.001

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica: Journal of the Econometric Society*, 591–605.

Ciner, C. (2019). Do industry returns predict the stock market? a reprise using the random forest. *The Quarterly Review of Economics and Finance*, *72*, 152–158.

Corsetti, G. & Pesenti, P. (2001). Welfare and macroeconomic interdependence. *The Quarterly Journal of Economics*, *116*(2), 421-445.

Cutler, D., Edwards Jr, T., Beard, K., Cutler, A., Hess, K., Gibson, J. & Lawler, J. (2007). Random forests for classification in ecology. *Ecology*, *88*(11), 2783–2792.

Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J. & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, *88*(11), 2783–2792.

Di Bella, G., Flanagan, M., Foda, K., Maslova, S., Pienkowski, A., Stuermer, M. & Toscani, F. (2022, July). Natural gas in europe: The potential impact of disruptions to supply. (22/150).

Dinesh, Y. (2019). *Categorical encoding using label-encoding and one-hot-encoder.* `https://towardsdatascience.com/categorical-encoding-using-label-encoding-and-one-` (Accessed: 2023-06-04)

Djankov, S., Freund, C. & Pham, C. (2010). Trading on time. *The Review of Economics and Statistics*, *92*(1), 166-173. Retrieved from `https://EconPapers.repec.org/RePEc:tpr:restat:v:92:y:2010:i:1:p:166-173`

Djankov, S. & Freund, C. L. (2002, Jul). Trade flows in the former soviet union, 1987 to 1996. *SSRN*. Retrieved from `https://papers.ssrn.com/sol3/papers.cfm?abstract`$_i d = 319145$

Djankov, S. & Nikolova, E. (2016). The fall of the iron curtain and the culture of entrepreneurship. *Journal of Economic Behavior Organization*, *127*, 295-376.

Fafchamps, M. & Gubert, F. (2007, May). Risk sharing and network formation. *American Economic Review*, *97*(2), 75-79. Retrieved from `https://www.aeaweb.org/articles?id=10.1257/aer.97.2.75` doi: 10.1257/aer.97.2.75

Federation of German Industries (BDI). (2021). *Export controls and export bans over the course of the covid-19 pandemic* (Report). World Trade Organization. `https://www.wto.org/english/tratop_e/covid19_e/bdi_covid19_e.pdf`. (26-02-2023)

Feenstra, R. & Kee, H. L. (2008, 03). Export variety and country productivity: Estimating the monopolistic competition model with endogenous productivity. *Journal of International Economics*, *74*, 500-518. doi: 10.1016/j.jinteco.2006.11.006

Freund, C. & Ornelas, E. (2010, 02). Regional trade agreements. *Annual Review of Economics*, *2*. doi: 10.1146/annurev.economics.102308.124455

Ghisletta, R. O. J. N. C. D., P. (2015). Linear mixed-effects and latent curve models for longitudinal life course analyses. *A Life Course Perspective on Health Trajectories and Transitions*, *4*. doi: 10.1007/978-3-319-20484-0_8

Gourinchas, P. O. & Obstfeld, M. (2012). Global imbalances and the financial crisis: Products of common causes. *Central Banking, Analysis, and Economic Policies Book Series*, *15*, 63-114.

Guo, Y., Hastie, T. & Tibshirani, R. (2007). Regularized linear discriminant analysis and its application in microarrays. *Biostatistics*, *8*(1), 86–100.

Helpman, E., Melitz, M. & Rubinstein, Y. (2008b). Estimating trade flows: Trading partners and trading volumes. *The Quaterly Journal of Economics*(2), 441-487.

Helpman, E., Melitz, M. J. & Rubinstein, Y. (2008a). Estimating trade flows: Trading partners and trading volumes. *Quarterly Journal of Economics*, *123*(2), 441-487.

International Energy Agency. (2022). *Russia's war on ukraine.* `https://www.iea.org/topics/russias-war-on-ukraine`. (26-02-2023)

International Monetary Fund. (2021). *World economic outlook, april 2021: Managing divergent recoveries.* International Monetary Fund. Retrieved from `https://www.imf.org/en/Publications/WEO/Issues/2021/03/23/world-economic-outlook-april-2021`

Kaminski, B. (2001). How accession to the european union has affected external trade and foreign direct investment in central european economies. doi: 10.13140/RG.2.2.25825.81762

Lusch, B., Nair, A., Gao, Y., Bettencourt, J. & Buell, D. A. (2018). Deep learning for universal linear embeddings of nonlinear dynamics. *Nature Communications*, *9*(1), 4950.

Marrs, F. W., Fosdick, B. K. & Mccormick, T. H. (2023). Regression of exchangeable relational arrays. *Biometrika*, *110*(1), 265-272.

McKinsey & Company. (2021). *The consumer demand recovery and lasting effects of covid-19.* https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/the-consumer-demand-recovery-and-lasting-effects-of-covid-19. (26-06-2023)

Obstfeld, M. & Rogoff, K. (1996). *Foundations of international macroeconomics.* MIT Press.

Page, E. S. (1955). A test for a change in a parameter occurring at an unknown point. *Biometrika*, *42*(3/4), 523–527. Retrieved 2023-05-09, from http://www.jstor.org/stable/2333401

Pathak, J., Hunt, B., Girvan, M., Lu, Z. & Ott, E. (2017). Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach. *Physical Review Letters*, *120*(2), 024102.

Pathak, J., Lu, Z., Hunt, B. R., Girvan, M. & Ott, E. (2018). Using machine learning to replicate chaotic attractors and calculate lyapunov exponents from data. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *28*(4), 041101.

Pathak, J., Wikner, A., Fussell, R., Chandra, S. & Girvan, M. (2018). Hybrid forecasting of chaotic processes: Using machine learning in conjunction with a knowledge-based model. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *28*(4), 041102.

Prasad, A. M., Iverson, L. R. & Liaw, A. (2006a). Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems*, *9*(2), 181–199.

Prasad, A. M., Iverson, L. R. & Liaw, A. (2006b). Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems*, *9*(2), 181–199.

Qurban, M. (2021). *Why is trade important and how does it work?* https://www.tradefinanceglobal.com/posts/why-is-trade-important-how-does-it-work/. (26-06-2023)

Ravn, M. O. (1997). International business cycles in theory and in practice. *Journal of International Money and Finance*, *16*(2), 255-283. Retrieved from https://www.sciencedirect.com/science/article/pii/S0261560696000563 doi: https://doi.org/10.1016/S0261-5606(96)00056-3

Redding, S. J. & Sturm, D. M. (2008, December). The costs of remoteness: Evidence from german division and reunification. *American Economic Review*, *98*(5), 1766-97. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.98.5.1766 doi: 10.1257/aer.98.5.1766

Reeves, J., Chen, J., Wang, X. L., Lund, R. & Lu, Q. Q. (2007). A review and comparison of changepoint detection techniques for climate data. *Journal of applied meteorology* *and climatology*, *46*(6), 900–915.

Ricardo, D. (1817). *On the principles of political economy and taxation* (P. Sraffa, Ed.). London: John Murray.

Sapsis, T. P. (2018). Reduced-order modeling for nonlinear dynamical systems using deep learning techniques. *Journal of Computational Physics*, *366*, 201–216.

Schlichtkrull, M., Kipf, T. N., Bloem, P., van den Berg, R., Titov, I. & Welling, M. (2018). Modeling relational data with graph convolutional networks. In *European semantic web conference (eswc).*

Simpson, S. L., Bahrami, M. & Laurienti, P. J. (2019). A mixed-modeling framework for analyzing multitask whole-brain network data. (2), 307-324.

Takeishi, N., Fujii, K. & Iwata, T. (2017). Learning to describe unknown dynamic systems based on sparse modeling. *Science Advances*, *3*(10), e1700387.

Tinbergen, J. (1962). Shaping the world economy. suggestions for an international economic policy.

Wacziarg, R. T., Spolaore, E. & Alesina, A. F. (2003). Trade, growth and the size of countries. *Available at SSRN 367263*.

Waugh, M. E. (2014). International trade and income differences. *American Economic Review*, *104*(11), 3600-3633.

Westveld, A. H. (2007). *Statistical methodology for longitudinal social network data* (Unpublished doctoral dissertation). Department of Statistics, University of Washington, Seattle, WA.

Westveld, A. H. & Hoff, P. D. (2011). A mixed effects model for longitudinal relational and network data, with applications to international trade and conflict. *The Annals of Applied Statistics*, *5*(2A), 843-872.

Wolf, H. C. et al. (1999). *Transition strategies: choices and outcomes.* International Finance Section.

World Trade Organization. (2020). *Trade set to plummet as covid-19 pandemic upends global economy.* `https://www.wto.org/press/news443r.htm`.

World Trade Organization. (2021). *World trade statistical review 2021.* Retrieved from `https://www.wto.org/english/res`$_e/statis_e/wts2021_e/wts2021_e.pdf$

World Trade Organization. (2022). *Evolution of trade under the wto: handy statistics.* `https://www.wto.org/english/res_e/statis_e/trade_evolution_e.htm`. (26-06-2023)

Zhou, D. & Wu, Z. (2020). Deep learning for nonlinear dynamics in finance. *Journal of Financial Economics*, *137*(3), 811–829.

Zlotnik, A. & Khasanova, R. (2019). Nonlinear dynamic modeling with radial basis function neural networks. *Neural Networks*, *116*, 203–216.

A Growth trade volume

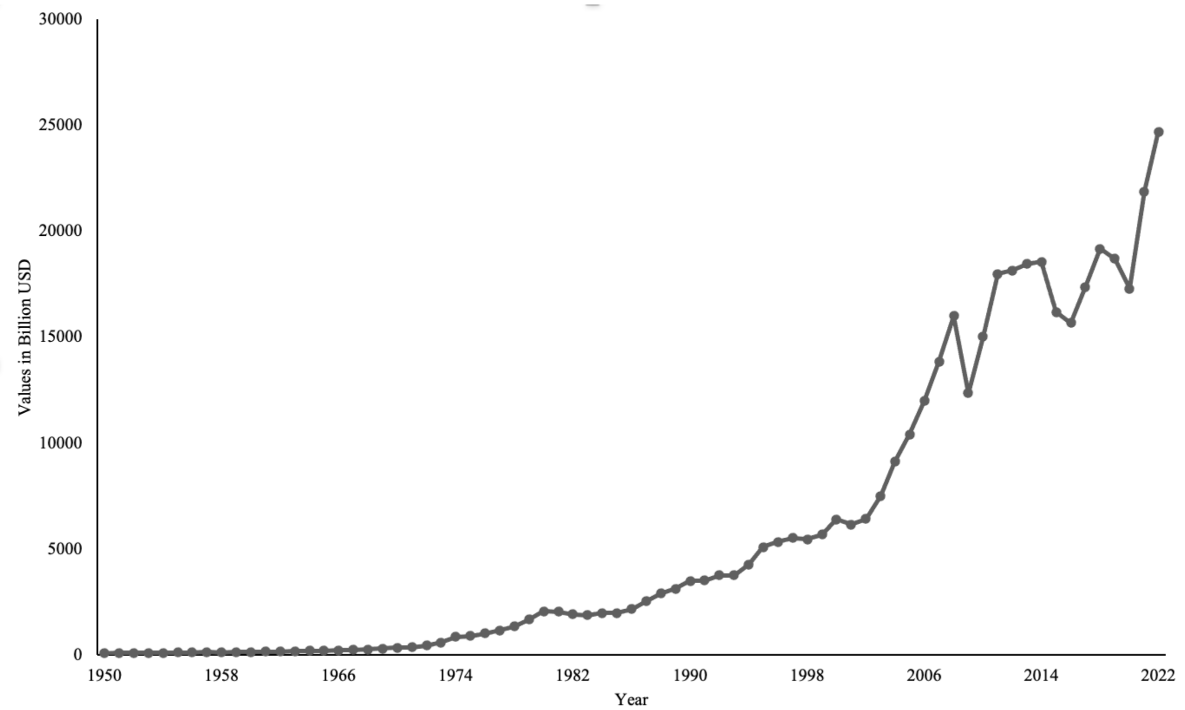

Figure 5: Growth of trade volume from 1950-2022.
Adapted source: World Trade Organisation (2022)

B List of countries

- Algeria
- United Arab Republic and Egypt
- Argentina
- Australia
- Austria
- Barbados
- Belgium
- Bolivia
- Brazil
- Canada
- Chili
- Colombia
- Costa Rica
- Cyprus
- Denmark
- Ecuador
- El Salvador
- Finland
- France
- Germany
- Greece
- Guatemala
- Honduras
- Iceland
- India
- Indonesia
- Ireland
- Israel
- Italy
- Jamaica
- Japan
- Malaysia

- Mauritius
- Mexico
- Marocco
- Nepal
- Netherlands
- New Zealand
- Norway
- Oman
- Panama
- Paraguay
- Peru
- Philippines
- Portugal
- Republic of Korea
- Singapore
- Spain
- Sweden
- Switzerland
- Thailand
- Trinidad and Tobago
- Tunisia
- Turkey
- United Kingdom
- United States
- Uruguay
- Venezuela

C Exchangeable estimator

$$
\begin{aligned}
\hat{\phi}_0^{(1)} &= \frac{1}{Rn(n-1)} \sum_r \sum_i \sum_{j \neq i} e_{ijr}^2, \\
\hat{\phi}_a^{(1)} &= \frac{1}{Rn(n-1)} \sum_r \sum_i \sum_{j \neq i} e_{ijr} e_{jir}, \\
\hat{\phi}_b^{(1)} &= \frac{1}{Rn(n-1)(n-2)} \sum_r \sum_i \sum_{j \neq i} e_{ijr} \left(\sum_{k \neq i} e_{ikr} - e_{ijr} \right), \\
\hat{\phi}_c^{(1)} &= \frac{1}{Rn(n-1)(n-2)} \sum_r \sum_i \sum_{j \neq i} e_{ijr} \left(\sum_{k \neq j} e_{kjr} - e_{ijr} \right), \\
\hat{\phi}_c^{(1)} &= \frac{1}{2Rn(n-1)(n-2)} \sum_r \sum_i \sum_{j \neq i} e_{ijr} \left(\sum_{k \neq i} e_{kir} + \sum_{k \neq j} e_{jkr} - 2e_{jir} \right), \\
\hat{\phi}_0^{(2)} &= \binom{R}{2}^{-1} \frac{1}{n(n-1)} \sum_{r \neq s} \sum_i \sum_{j \neq i} e_{ijr} e_{ijs}, \\
\hat{\phi}_a^{(2)} &= \binom{R}{2}^{-1} \frac{1}{n(n-1)} \sum_{r \neq s} \sum_i \sum_{j \neq i} e_{ijr} e_{jis}, \\
\hat{\phi}_b^{(2)} &= \binom{R}{2}^{-1} \frac{1}{n(n-1)(n-2)} \sum_{r \neq s} \sum_i \sum_{j \neq i} e_{ijr} \left(\sum_{k \neq i} e_{iks} - e_{ijs} \right), \\
\hat{\phi}_c^{(2)} &= \binom{R}{2}^{-1} \frac{1}{n(n-1)(n-2)} \sum_{r \neq s} \sum_i \sum_{j \neq i} e_{ijr} \left(\sum_{k \neq j} e_{kjs} - e_{ijs} \right), \\
\hat{\phi}_c^{(2)} &= \binom{R}{2}^{-1} \frac{1}{2n(n-1)(n-2)} \sum_{r \neq s} \sum_{i=1}^{n} \sum_{j \neq i} e_{ijr} \left(\sum_{k \neq i} e_{kis} + \sum_{k \neq j} e_{jks} - 2e_{jis} \right)
\end{aligned} \tag{21}
$$