Universitat Pompeu Fabra

Departament d'Economia i Empresa

DOCTORAT EN ECONOMIA, FINANCES I EMPRESA

Essays on Trade Agreements

Tesi Doctoral de:

Maria Ptashkina

Directors de la tesi:

Jaume Ventura

Manuel García-Santana

TESI DOCTORAL UPF / 2023

Acknowledgements

This thesis is the final chapter of my PhD studies in Barcelona, and many amazing people have contributed to it. First and foremost, I want to thank my advisor, Jaume Ventura, who has been a mentor and a source of motivation, curiosity, and a constant strive for the better. Jaume has always encouraged and supported me in matters related to research, as well as in life decisions. Jaume has taught me to think deeply about every question—no matter how 'simple' it seems—and has been the key person in bringing me on to the academic career path, opening new perspectives during the many conversations we had throughout these years. I want to thank my second advisor—Manuel García-Santana—for his guidance, support, and patience. Manuel has inspired me to work hard and keep the focus on the goals. Manuel's many detailed comments and suggestions to my work have helped me to pay attention to the detail, and substantially improved my writing and presentation skills.

Second, I want to thank the faculty at CREi for their feedback and ideas on how to improve my work. Fernando Broner, David Nagy, Giacomo Ponzetto, Alberto Martin, Joan Monras, Victoria Vanasco all have contributed to the development of my research. I am also grateful to the UPF faculty, and, in particular, to Geert Mesters and Albrecht Glitz. Antonio Penta has been a great mentor, and I thank him for the academic advice and guidance. I thank Marta Araque, who has always answered my numerous (not always smart) questions, and provided excellent administrative support, especially during the job market.

This journey would not have been the same, have I not encountered my amazing PhD peers who have long become more friends, rather than just colleagues. Zoel: we've made this journey shoulder to shoulder, from the beginning to the end. You have always been there for me, through the darkest hours of the advanced track, the highs and the lows of the PhD, and the life turmoils. Thank you for that, and for the time we spent together in the office and outside of it. Malachy: thank you

for the many hours discussing research and ideas, but even more—for your fantastic sense of humor and your ability to cheer me up. I hope we will keep 'writing' that 'paper' forever. Eva: thank you for the countless Tuesday bloody-mary classics, and for being supportive and fun. The 'Pan Tumacas,' Joaquim and Miquel: thank you for being close when I needed it most, and for the feeling of light-heartedness that you bring.

I also want to give special thanks to my friends outside of the PhD. Claudia: thank you for trusting me, and for being an immense source of support and care, through my lowest physical states included. You are also my number one fun partner: this 'stuff' is on! Andras, thank you for both the academic support, and for being a close and understanding friend—the numerous fun situations involving you will always make me smile, Jagodichki. Hector, Jairo, Mayte, Melissa, Florent, Barna, and Anna: thank you for sharing some of the best moments in Barcelona, Costa Brava and the Pyrenees with me. Adalberto: thank you for the many fantastic 'situazione' together, keep them coming! Tanya, Ben and Oge: thank you for being my role models in organization, patience, and maturity.

Всем в своей жизни я обязана своей семье. Мамочка: ты мой самый близкий и любимый человек. Ты поддерживала меня всегда, что бы я не творила. Ты всегда говоришь, что главное – чтобы я была счастлива, и я тебе безмерно благодарна за это! Спасибо, что ты поняла и приняла моё решение переехать и начать жизнь и карьеру за границей—несмотря на нашу физическую разлуку, ты всегда рядом. Папа: я знаю, что ты гордишься мной, и для меня это важно. Спасибо тебе за поддержку и любовь. Мои маленькие сестрёнки (уже не такие маленькие!) – Вика и Саша: вы для меня постоянный источник радости и гордости! Томочка и дядя Саша: спасибо вам за вашу искреннюю любовь и поддержку, я всегда чувствую вашу заботу, даже на расстоянии. Братик Саша, Алёна, Саша и Гоша: вы для меня источник вдохновения и пример семейной гармонии, спасибо, что всегда принимаете и понимаете меня. Лыс, спасибо тебе, что ты всегда рядом, ты и так сам всё знаешь! Алёнчки, я тебя благодарю здесь, в абзаце, посвященном семье, не просто так. Ты мне родной человек, моя барселонская семья, спасибо тебе за любовь, заботу и поддержку все эти годы. И это не конец, ведь теперь ты – семья.

Finally, I want to thank my life partner and family, Cristiano. Nothing would have been possible without you. Thank you for the deepest understanding and care, for the little Ciopino, and for giving a meaning to everything.

Abstract

This dissertation consists of two chapters studying trade agreements. The first chapter estimates the effects of preferential trade agreements on bilateral trade between their signatories. A novel empirical method used in the chapter reduces the self-selection bias in the estimates. The results uncover heterogeneous effects of trade agreements across different types of country pairs and time horizons. Equipped with the empirical estimates, this chapter builds a model to analyze the general equilibrium effects of a recent important agreement. The second chapter studies the long-run evolution of the content of commercial treaties. A unique dataset collects information about trade agreements dating back to 1815. The analysis shows that many features of the contemporary trade agreements were already present in the earlier treaties. Moreover, the usage of separate commercial provisions was adjusted and polished along with the political and economic developments of the different historical periods.

Resum

Aquesta tesi conté dos capítols estudiant acords comercials. El primer capítol estima l'efecte dels acords de comerç preferencial en el comerç bilateral entre els seus signataris. Un nou mètode empíric utilitzat en el capítol redueix el biaix d'autoselecció de les estimacions. Els resultats destapen efectes heterogenis dels acords comercials entre els diferents tipus de parelles de països i els diferents horitzons temporals. A partir d'aquests resultats, aquest capítol construeix un model per analitzar els efectes d'equilibri general d'un acord comercial recent. El segon capítol estudia l'evolució a llarg termini del contingut dels acords comercials. Una base de dades única remuntant a 1815 recull informació sobre acords comercials. L'anàlisi mostra que moltes de les característiques dels acords comercials contemporanis ja eren presents en tractats anteriors. Encara més, l'ús de disposicions comercials separades es va ajustar i polir juntament amb l'evolució política i econòmica dels diferents períodes històrics.

Preface

This dissertation investigates two questions related to trade agreements. In this study I employ novel empirical methods, and collect unique datasets to analyze the the effects of trade agreements on trade, and to understand the origins and the long-run evolution of the commercial provisions in the treaties.

The first chapter estimates the effects of preferential trade agreements (PTAs) on bilateral trade between their signatories. Identifying the effects of PTAs on trade is challenging due to self-selection bias: countries choose to become members. To address the selection issue, this chapter builds a comprehensive dataset, and uses the blocking estimator from the causal inference framework. The results show that, after accounting for selection, PTAs increase bilateral trade by 48% fifteen years after entry into force. Importantly, the effects phase in gradually, with one-third being realized five years prior to the agreement's entry into force. These anticipation effects are only present for non-natural trading partners—geographically distant countries that trade little and have a low probability of signing an agreement. Comparing the estimates obtained by using the blocking estimator with the estimates obtained using other methods, I show that the latter do not account for selection, and substantially overestimate the effects of trade agreements. Finally, equipped with the empirical estimates, I build a quantitative model to analyze the general equilibrium effects of the largest preferential trade agreement in the world—the Regional Comprehensive Economic Partnership.

The second chapter takes a long-run perspective and analyzes the evolution of trade agreements from the 'first wave of globalization'—the beginning of the 19th century—to modern times. I collect a novel dataset containing information about trade agreements dating back to 1815. I present the information from the texts of the treaties in two different ways. First, by reading the treaty texts, I identify and codify the key design features that were agreed upon by the signatories. Second, I analyze

the evolution of the design of the commercial treaties, the emergence and the prevalence of the key features in different time periods, and the geographical distribution of the treaties. The results show that many of the features of the contemporary trading system were developed long before the General Agreement on Tariffs and Trade (GATT)—the foundation of the modern trading system—was signed. Moreover, it emerges from the analysis that many of the commercial provisions in the treaties appeared, adjusted and developed along with the political and economic developments of the different historical periods.

Contents

List of Figures

List of Tables

Chapter 1

Revisiting the Effects of Preferential Trade Agreements

1.1 Introduction

Almost every country in the world is a member of at least one preferential trade agreement (PTA). As of 2021 a total of 354 trade agreements were in force, regulating trade between 16% of all country pairs in the world (in the 1970s, less than 1% of all pairs had an agreement). The provisions of these agreements cover 80–90% of bilateral trade between their signatories, and currently govern more than half of all world trade.

What are the effects of preferential trade agreements on bilateral trade between their members? If trade agreements were randomly assigned, comparing the average trade of member country pairs with the average trade of the outsiders would provide an unbiased estimate of the causal effects. The main issue, however, is that PTAs are endogenous trade policy decisions of countries.

Since countries self-select into membership, trade and PTA assignment are intrinsically related: bigger and closer countries have larger trade volumes and are more likely to form PTAs (Baier and Bergstrand (2004); Magee (2003); Egger et al. (2011)). This paper uses the blocking estimator from the causal inference framework of Imbens and Rubin (2015) to account for the probability of signing an agreement given past trade and other country-pair characteristics, and to estimate the effects without relying on functional form assumptions.

The main finding is that PTAs have sizeable effects even after accounting for selec-

tion. In particular, bilateral trade outcomes are 48% higher for pairs with a PTA fifteen years after the agreement's entry into force. The effects phase in gradually, with one-third of the total increase observed in anticipation, five years prior to entry into force. These effects are heterogeneous across different types of country pairs. Natural trading partners—geographically close countries with high initial trade levels and higher probability to conclude a trade deal—do not react in anticipation. The entire anticipation effect is thus driven by country pairs with larger bilateral distances, lower pre-PTA trade volumes, and low probability of having a PTA (non-natural trading partners). In the long run, however, the percentage increase in trade of country pairs with PTAs relative to their non-member counterparts is similar for all types of country pairs. Additionally, the paper demonstrates that selection is important: the effects are halved when compared to the results of the alternative empirical research designs.

The second part of the paper builds a simple model to demonstrate how the estimates obtained in the empirical part can be used to study the general equilibrium effects of PTAs. In particular, the application focuses on the Regional Comprehensive Economic Partnership (RCEP) agreement—the largest preferential trade agreement in the world.

Self-selection into PTA membership generates a bias largely due to the fact that country pairs with certain characteristics have higher trade volumes and are more likely to become PTA members. The blocking estimator of Imbens and Rubin (2015) finds subsamples (blocks) of country pairs, such that within the blocks the pairs with PTAs (treated) and without PTAs (control) are similar in all characteristics, and also in the probability of signing an agreement.

The implementation of the estimator is very demanding in terms of data. One of the contributions of this paper is a comprehensive dataset of bilateral trade flows and country pair characteristics, which tracks virtually all pairs in the world over a period of 60 years. The dataset combines all existing trade data sources, and adds almost one million observations from 1960 to 2019, which would otherwise be considered missing. The dataset's extensive coverage makes it possible to analyze the dynamic effects of a large set of trade agreements concluded between different types of country pairs.

Some of the country pair characteristics, such as geographical, cultural and historical ties or past trade, predict the probability of signing a trade agreement, but are not affected by its presence. From an empirical viewpoint the probability of signing a PTA and trade outcome can be thus simply conditioned on these covariates. The

issue of economic size is more subtle. Since PTAs increase trade and economic size by reducing trade costs, controlling for size would not be a viable empirical strategy.[1] To deal with this issue, the outcome variable is defined as a 'size-free' measure of bilateral trade following Santamaría et al. (2020): the normalized market share is calculated as a market share of an origin i in a destination j normalized by the average share of i in all markets. The advantage of using such normalized market shares is that they are not mechanically affected by the size of origin or destination countries.

The identification strategy consists of three distinct stages. In the first stage—design (Rubin (2005))—no outcome data is used, the focus is solely on the PTA indicators and covariates. The probability of entering a PTA is modeled as a function of geographical, cultural, and historical characteristics of country pairs, as well as their past trade. Country pairs with values of covariates such that they have no counterparts in the other treatment group are trimmed away from the sample. The reason is that no general estimation procedure would give robust estimates in a sample with 'incomparable' units (Imbens (2014)).[2] After trimming the dataset, the blocking algorithm constructs blocks of country pairs such that, within each block, the conditional probability of receiving a PTA in the future is similar, and country pairs with and without PTAs, to the extent possible, do not differ in observable covariates. This design treats PTA assignment within blocks as random.

The second stage is diagnostics and robustness analysis. First, the focus is on the differences in covariate distributions by treatment status across and within blocks. This diagnostic highlights the importance of separately estimating the effects of PTAs for different types of trading partners. The second check is on the extent of treatment heterogeneity: PTAs differ across multiple dimensions (Dür et al. (2014); Hofmann et al. (2019)), and different types of PTAs might have different effects (Magee (2008)). The paper tests whether the characteristics of PTAs, such as, for example, the type (free trade area or customs union) or composition (bilateral or agreements with multiple members), can be predicted by the types of country pairs signing them. While there is some variation across the blocks, the covariates within each block cannot predict the features of a PTA that a pair will sign. Therefore, PTA types are treated as random within each block. Third, the paper examines the extent of the missing

[1] Since trade volumes and size have a positive association with a PTA, controlling for size would lead to overestimating the effects of PTAs.

[2] The intuition is similar to extrapolation bias in a linear regression: if covariate distributions of treated and control units are substantially different, conventional regression methods can be very sensitive to minor changes in the specification.

values problem. Conditioning the analysis on positive trade flows might induce a downward bias in the estimates. To understand the severity of the issue, an interpolation exercise recovers some low-trade observations. Since the proportion of missing values within the blocks is small, the final estimates appear to be robust to the reconstruction of the trade matrix.

The next stage—analysis—involves estimating the PTA effects and their sampling variances. Regression adjustment within each block accounts for the residual differences in covariate distributions. The estimator of the average effects of PTAs on their members in the entire trimmed sample is then calculated as the weighted average of individual treatment effects within each block (Rosenbaum and Rubin (1984)). To estimate the time-varying responses to PTAs, the outcomes are defined to correspond to different time windows around agreements' entry into force: anticipation, short-run, medium-run and long-run. By doing so, the outcome variable takes into account the dynamic adjustment of trade in response to PTA shocks (Magee (2008); Egger et al. (2020)).[3]

The magnitude of the quantitative estimates obtained in this paper is lower compared to those obtained by applying alternative research designs. For example, a standard empirical specification—a linear regression with three-way fixed effects—yields estimates which are twice larger in anticipation.[4] A simple numerical simulation indicates that the self-selection bias might not be picked up by the alternative estimators. Since PTA members tend to have higher levels of trade (normalized market shares), be closer to each other in geographical and cultural dimensions, the bias in the full sample is expected to be positive. Causal inference framework applied in this setting is aimed at reducing the selection bias, and thus predicts smaller effects.

Importantly, the resulting empirical estimates should be interpreted as partial equilibrium effects, i.e. the effects of a PTA on its members under the assumption that "everything else" stays equal. In this setting, partial equilibrium effects represent increases in bilateral trade of a country pair with a PTA, *regardless* of whether this trade increase is driven by pure substitution from other destinations or from domestic trade. Another way to think about this assumption is that country pairs are

[3] Anticipation corresponds to the average trade outcomes in the five years prior a PTA's entry into force (approximately corresponding to a mean negotiation period across different agreements); short-run outcome measures a five-year average following a PTA's entry into force; medium-run and long-run outcomes are defined as averages of five to ten and ten to fifteen years respectively.

[4] As the paper shows, both trimming and blocking into subsamples play a role in reducing the size of the bias associated with self-selection into PTAs.

small, and endowing one of them with a PTA will not affect the trade outcomes for all other pairs. Here, the interest is not in predicting what happens in a world where all pairs get a PTA, but rather what would happen to a randomly selected country pair if it gets assigned a PTA.

Partial equilibrium interpretation may seem to contrast with a large body of existing literature using structural gravity models to study the effects of PTAs. Identifying partial equilibrium estimates, however, was a focus of many empirical studies (see, for example, Soloaga and Wintersb (2001); Baier and Bergstrand (2009); Egger and Tarlea (2021)), and plays an important role in determining the general equilibrium effects (Egger et al. (2011)).

In the second part of the paper this idea is developed further. The paper builds a general equilibrium model to make predictions about the changes in welfare (real consumption) and trade patterns following a shock to trade costs. The model is the simplest version of the quantitative structural gravity setup, defined in Costinot and Rodríguez-Clare (2014). To make counterfactual predictions given the decrease in trade costs for PTA members, the model is solved in changes using the 'exact hat algebra' approach.

The model uses the partial equilibrium estimates from the empirical part to study the effects of the Regional Comprehensive Economic Partnership agreement. RCEP is one of the most important and largest recent trade agreements involving fifteen countries in Asia.[5] It was signed in 2020, and started entering into force from the beginning 2022. The agreement is set to reduce trade barriers on 90% of goods trade. Moreover, it influences trade more broadly by covering multiple regulatory aspects relating to goods, services, investment, economic and technical cooperation, and creating new rules for electronic commerce, intellectual property, government procurement, competition, and small and medium sized enterprises. The contents and contemporaneity of RCEP makes it a relevant policy question to study.

The model is used to conduct two counterfactual exercises. The first one endows RCEP countries with a trade agreement using the long run average estimates obtained in the empirical part of the paper. The second one exploits the full heterogeneity of the empirical estimates across blocks and time since entry into force. Although the model does not feature any dynamics, a series of static exercises predicts the counterfactual changes in real consumption and trade reallocation in the antici-

[5]China, Japan, South Korea, Australia, New Zealand, and ten Southeast Asian economies (Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar, the Philippines, Singapore, Thailand, and Vietnam.

pation period and long run.

Both exercises predict increases in welfare for the RCEP members. Smaller countries (like Myanmar and Cambodia) experience the largest gains, while the larger countries (like China, Japan and Korea) experience negligible increases in welfare. The changes in the average welfare in the rest of the world (weighted by the countries' initial size) are small, but positive. The changes in trade are much larger in magnitude. The model predicts a substantial amount of trade creation, i.e. a disproportionate increase in trade of RCEP economies within the PTA. At the same time, trade diversion effects are small on average: RCEP countries reduce their trade with some outsides, while increasing it with others.

The general equilibrium exercise highlights two important implications. First, it demonstrates how well-identified partial equilibrium estimates can be used in conjunction with the model to study the consequences of the real-world PTA formation. Such estimates are crucial for policy makers to take informed decisions about entering PTA negotiations. Second, the model shows that the magnitudes of the partial equilibrium estimates matter for the general equilibrium predictions. In particular, the larger estimates that do not account for selection substantially amplify the predictions regarding trade diversion and trade creation.

Literature Review. Estimating the effects of PTAs has been a central question in trade literature for decades. The dominant paradigm to approach this question is using an empirical form of a gravity equation, where the volumes of bilateral trade flows are regressed on PTAs and covariates, or sets of fixed effects. Head and Mayer (2014) provide an extensive overview of the gravity literature, and note that typically studies find large point estimates (a 60% increase in bilateral trade). The estimates also vary greatly across studies, predicting increases in trade from zero to more than 200%. In the same vein, Ghosh and Yamarik (2004) and Baier and Bergstrand (2007) note that estimates of PTA effects in a gravity setting are highly unstable.

One reason is that empirical implementations of the structural gravity model are susceptible to changes in the estimation methodology.[6] However, even when using

[6]Yotov et al. (2016) summarize the best practices. Among other recommendations, they recommend applied researchers to estimate the gravity equation accounting for multilateral resistance terms (Anderson and van Wincoop (2003); Feenstra (2004); Olivero and Yotov (2012)); to use Poisson Pseudo Maximum Likelihood estimator to include zero trade flows (Santos Silva and Tenreyro (2006)); and to account for trade policy endogeneity by adding country-pair fixed effects (Baier and Bergstrand (2007).)

similar methods, there is little consensus on the magnitude of the point estimates. For example, when adding dyadic fixed effects, Baier and Bergstrand (2007) find that the PTA estimate is multiplied by more than a factor of two, while Head et al. (2010) find that the coefficient is halved.

Another reason gravity estimates for PTAs 'are not reliable', as noted by Head and Mayer (2014), is that they fail to correctly address the endogeneity of trade policy.[7] To understand the PTA formation mechanism, Baier and Bergstrand (2004) explore the role of the economic determinants. Magee (2003) additionally finds that, empirically, past trade is an important predictor of PTAs. Egger and Larch (2008) conclude that interdependence is positively correlated with the formation and enlargement of PTAs. Given these insights, Egger et al. (2011) model the selection into PTA membership and use the predicted probability as a regressor in a gravity model. They find that the estimated PTA coefficient increases by more than a factor of 5 (from 42% to 220%). My paper relies on these earlier studies highlighting the determinants of PTAs – such as geographical and cultural characteristics, past trade, and the number of PTAs already concluded – in calculating the conditional probability of PTA membership. This paper, however, departs from the empirical gravity framework to address selection.

This paper is closest to the literature using non-parametric estimation techniques to evaluate the effect of endogenous PTAs on trade. Egger et al. (2008) look at effects of PTAs on trade volumes and intra-industry trade in the subsample of OECD member countries. Using matching estimators, they conclude that a simple difference-in-difference estimator without accounting for self-selection into new PTA membership is downward-biased by 62-86% depending on the type of matching. Baier and Bergstrand (2009) explore cross-sections of data for 96 countries in different years using a matching estimator. They report that the estimates of the average treatment effects are between 0.68 for the year 2000 (implying an effect of about 97%) and 2.36 for the year 1990 (around 900%). Their preferred estimate – the average treatment effect on the treated – is more economically plausible, implying a 132% increase in bilateral trade. Egger and Tarlea (2021) employ entropy balancing to "compare apples to apples," i.e. PTA members with the outsiders with the same (re-weighted) values of

[7] While dyadic fixed effects forces identification to come from the within dimension of the data, the estimate cannot be interpreted as causal, since there are may be other factors, along with PTAs, that vary at the country-pair-time level, which will be picked up by the coefficient.

observable covariates.[8] They show, in contrast to the earlier non-parametric studies – and similarly to this paper – that enforcing covariate balance actually reduces the estimates of PTA effects.

The above-cited papers represent important conceptual and methodological advances in studying the effects of endogenous policy decisions. This paper builds on their insights in the following ways. First, the blocking estimator requires modeling the probability of signing a trade agreement. The advantage of such approach is that the subsequent estimation of the effects on trade takes into account not only the covariate distributions, but also the estimated probability. Thus, conditioning on the propensity score separates the influence of the covariates on the PTA assignment from their direct influence on trade, ultimately reducing the selection bias.

Another methodological improvement relates to using balancing instead of matching (like in Egger et al. (2008) and Baier and Bergstrand (2009)). Unlike propensity score matching, blocking allows to balance covariate distributions. As demonstrated by King and Nielsen (2019), the estimated propensity score should not be used for matching, since it implies matching on a uni-dimensional vector, thus potentially increasing covariate imbalance, inefficiency, model dependence, and bias.

Finally, in setups with substantial heterogeneity, such as international trade across different countries in the world, the estimates of the average treatment effects will depend on the sample composition. For example, given the results of this paper, studying samples which include more natural trading partners might lead to the underestimation of the anticipation effects. Similarly, timing matters: for example, if the data covers only the first five years after PTAs enter into force, the estimate would be mute on the long run effects for such agreements. In this setting, the blocking estimator takes into account the heterogeneity relating to different types of pairs, and the design in this paper makes it possible to study the dynamic effects.

This paper is structured as follows. Section 2 describes the sources and construction of the data. Section 3 explains the study's empirical design and the identification strategy. Section 4 presents and discusses the results of the empirical framework. Section 5 builds a general equilibrium model, and, using the empirical estimates obtained earlier, applies it to study the effects of RCEP formation. Section 6 concludes and discusses avenues for future research.

[8] Entropy balancing is equivalent to estimating the weights as a log-linear model of the covariate functions (Hainmueller (2012)), and involves minimizing divergence from a set of baseline weights chosen by researchers, i.e. the method might be inconsistent unless the correct functions are specified.

1.2 Data

One of the contributions of this paper is the construction of a comprehensive dataset containing most complete data on bilateral trade flows and domestic trade. By assembling data from virtually all existing data sources, more than one million bilateral trade observations over the period from 1960 to 2019 are additionally recovered. If, instead, only one of these data sources was used, these observations would be considered missing. The dataset also includes extensive information on the characteristics of country pairs, as well as the features of the preferential trade agreements.

Data on bilateral exports is constructed by combining several data sources: UN Comtrade Database, CEPII Gravity Database, World Trade Flows (WTF) bilateral cross-sectional data, and IMF Direction of Trade Database. These trade flows are complemented by data on international trade and domestic trade from WTO Structural Gravity Database, USITC International Trade and Production Database for Estimation (ITPD-E), and UNIDO Industrial Statistics. The values reported by the destination are used as a default, and are complemented with values reported by origin, whenever available. As most sources have varying number of missing trade flows, the addition of different data sources helps to fill in many of the missing values. Appendix A details the exact procedure used to construct the dataset.

Even after combining different sources to get a fuller matrix of bilateral trade flows, many missing values remain. In particular, over the entire sample period 61% of international trade flows are missing. One approach to deal with the missing flows in trade literature is to declare them as zeros. It is, however, virtually impossible to distinguish missing values from zero trade flows. In fact, the mere combination of different trade data sources helped to recover a substantial amount of missing trade flows, suggesting that many observations are not zeros after all. In addition, there are data patterns that suggest that some flows might indeed be missing. For example, when we observe large trade flows at time t and $t+2$, but a missing value at $t+1$.

In order to deal with the missing data problem, this paper employs imputation to predict trade flows. The main purpose of imputing the missing values is to gain statistical power for the subsequent analysis, and to carefully deal with the participation bias (see Section 3.4 for further discussion). Appendix B lays out a detailed procedure to impute the missing trade flows. To summarize, the imputation procedure uses a flexible form of the empirical gravity model to impute values of trade for those pairs that have all the necessary covariate data available. This procedure

leads to imputing additional 428,267 observations, decreasing the number of missing values in the full sample to 45%. Later on, Section 3.4 additionally reports the results using the data obtained by applying an interpolation procedure. Interpolation additionally recovers 97,618 bilateral trade flows, reducing the number of missing values to 35%.

Importantly, all the data exercises never use the imputed *values* of trade flows in the analysis directly. Trade volumes are used to construct the normalized market shares, which depend not only on the bilateral flows, but on the whole matrix of flows. In this sense, the imputation helps to recover more precise shares, but does not bias the results. Appendix F shows that the differences in the distributions of imputed and raw normalized market shares are small, with the imputed shares having a slightly lower mean and variance. Appendix F implements the whole procedure without the imputation, and demonstrates that the main conceptual results remain unchanged, however, the statistical power is reduced when using the outcomes without the impultation.

Having obtained the matrix of bilateral trade flows, I construct the normalized market shares following Santamaría et al. (2020):

$$s_{ij} = \frac{V_{ij}/E_j}{Y_i/E} \tag{1.1}$$

where V_{ij} are the sales from origin i to destination j; $E_j = \sum_i V_{ij}$ is the total expenditure of j; $Y_i = \sum_j V_{ij}$ is the total income of i; and $E = \sum_j E_j$ is the world's total expenditure. If market j has above average importance for i, i.e. $V_{ij}/E_j > Y_i/E$, the normalized market share is above one. The important feature of the normalized market shares is that the economic size of origin and destination does not mechanically affect them.

To construct theory-consistent normalized market shares, s_{ij} should measure i's share in j, normalized by i's share in all markets, including itself. Unfortunately, before 1980 the data for domestic trade (or production data used to construct it) is available only for a very limited set of countries. To overcome this issue, this paper constructs the vector of domestic trade flows by combining multiple data sources (see detailed procedure in Appendix A), and checks whether the normalized market shares with and without domestic trade differ in the sample after year 1980, when domestic trade and production data becomes available. Appendix C discusses in detail the various checks, but here it suffices to say that the differences between the

two methods of calculating the outcome variable are small. The paper thus proceeds to construct the normalized market shares without domestic trade for all country pairs and years.

To get more intuition about this measure, consider as an example the bilateral trade flows from China to Vietnam and Germany (Table 1.1). In 2017 China exported 71.6 billion USD worth of goods to Vietnam, and a similar value – 71.1 billion USD – to Germany. While the value of the exported goods is very similar for the two destinations, the bilateral relationship of the two country pairs is not the same. In particular, 28% of all the Vietnamese expenditure on imports goes to China's goods, while only 8% of the Germany's imports come from China. China is a very large exporter in the world, accounting for almost 15% of the total world's exports. Normalizing China's market shares in each market by its share in the world's exports would predict that China and Vietnam are almost four times more important trading partners to each other, than China and Germany. This example demonstrates that although trade volumes are the same, the bilateral importance of the trading partners can be very different, once normalized by the sizes of origin and destination countries.[9]

Table 1.1: Trade volumes and normalized market shares for China's trade with Vietnam and Germany.

Year 2017	Volume of exports (billion USD)	Share of origin in the destination (%)	Share of origin in the world (%)	Normalized market share
China - Vietnam	71.6	28.64	14.99	1.91
China - Germany	71.1	7.99	14.99	0.53

The treatment dummy and the dataset on the characteristics of PTAs are constructed using the Design of Trade Agreements Dataset (Dür et al. (2014)). It contains the information on both the agreements notified to the WTO, as well as those that were not notified. Partial scope agreements are deleted from the sample, and the final treatment only includes fully enforced deals (free trade areas and customs unions). For each given treated county pair, the date of agreement's entry into force is coded as the earliest agreement. This way, a balanced panel is created, without superseding or overlapping PTAs, amendment protocols, or revisions. Appendix A provides more detail about the precise steps and examples of how to clean the dataset. Table 1.D.1 in Appendix D provides the descriptive statistics of all the PTAs in the full sample.

[9]The World Bank uses this measure, called the Trade Intensity Index, to describe trade relationships between countries.

Geographic and cultural characteristics come from CEPII's Gravity Dataset. The resulting set of variables is then complemented with other geographical variables using NASA's Earth Observing System Data and Information System (EOSDIS).

The full dataset includes 210 unique customs territories, with 319 PTAs, over the period 1960-2019. There are a total of 43,890 country pairs in cross-section, 16.13% of which are treated by year 2019. In comparison, the number of pairs with a PTA in 1970 was less than one percent. In a panel setting, only 6.37% have a PTA out of more than 2.5 million observations (Table 1.2).

Table 1.2: Full sample characteristics.

	Cross-section (2019)	Percent (2019)	Panel (1960-2019)	Percent (1960-2019)	Mean Share (1960-2019)
No PTA	36,812	83.87	2,465,521	93.63	2.55
PTA	7,078	16.13	167,879	6.37	17.69
Both	43,890		2,633,400		3.51

Figure 1.1 plots the average normalized market shares for pairs which had a PTA at any point in time, and those that never had a PTA. The treated country pairs have always had higher bilateral trade, and the gap with the control pairs has been increasing over the entire period of time. The question is how much of this increase can be attributed to the effects of PTAs, and how much is driven by other factors. The next section lays out the empirical design aimed at tackling the issue of self-selection into PTAs.

1.3 Empirical Strategy

The estimation of the causal effects of PTAs requires understanding the counterfactual outcomes of the treated units had they not received the treatment. The following empirical setup is defined using the causal inference framework of Imbens and Rubin (2015).

1.3.1 Setup, Assumptions and the Blocking Estimator

For each country pair with origin i and destination j there are two potential normalized market shares at a given time $T = \{A, S, M, L\}$ (anticipation, short, medium

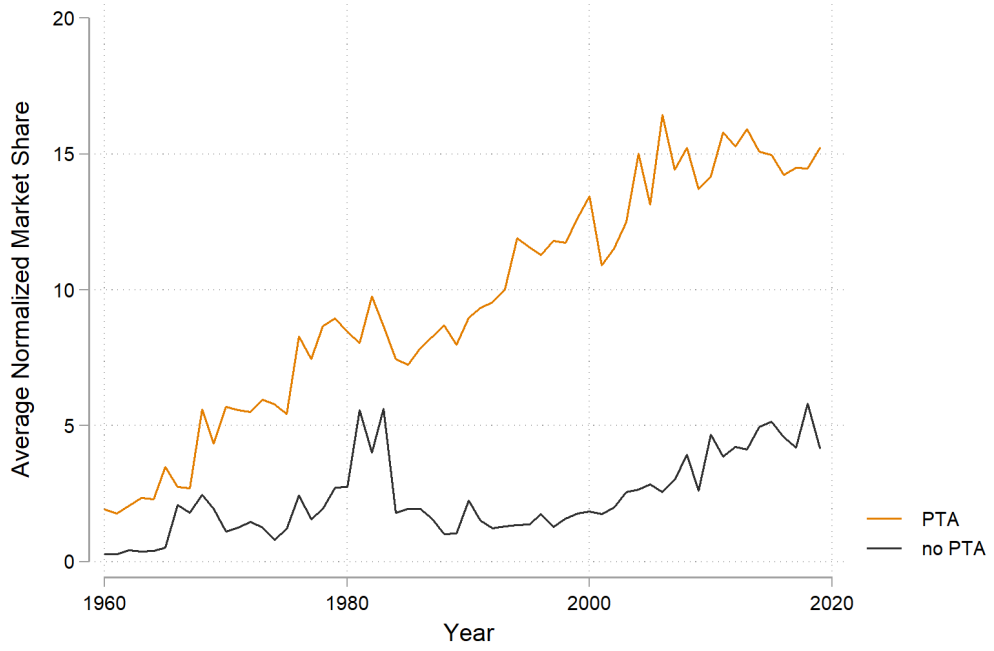

Figure 1.1: Average normalized market shares for pairs with and without PTAs, 1960-2019.

and long run), denoted as $s_{ij}^T(0)$ and $s_{ij}^T(1)$ – without and with a PTA respectively. The effect of a PTA at a given time is defined as the percentage change in average normalized market shares in a period around PTA's entry into force:

$$\tau_{ij}^T = \ln \frac{s_{ij}^T(1)}{s_{ij}^T(0)} \tag{1.2}$$

Each pair, however, is observed to either receive or not receive a binary treatment, $\text{PTA}_{ij} = 1$ or $\text{PTA}_{ij} = 0$. The realized (and observed) outcome for a pair is denoted with a subscript "obs" to distinguish it from the potential outcome which is not always observed:

$$s_{ij}^{T,obs} = \begin{cases} s_{ij}^T(0), & \text{if } \text{PTA}_{ij} = 0 \\ s_{ij}^T(1), & \text{if } \text{PTA}_{ij} = 1 \end{cases}$$

For each country pair there is also a K-component covariate Z_{ij}. The key characteristic of these covariates is that they are known not to be affected by the treatment: these are geographical, cultural and historical characteristics of country pairs, as well

as past trade (the next section will discuss the covariate selection in more detail). A triple $(s_{ij}^{T,obs}, \text{PTA}_{ij}, Z_{ij})$ is thus observed for all pairs in the sample.

In order to define an estimator for the average treatment effect which can be expressed in terms of the joint distribution of the observed data $(s_{ij}^{T,obs}, \text{PTA}_{ij}, Z_{ij})$, a few assumptions are necessary. The first key assumption is unconfoundedness (Rubin (1990)) or conditional independence (Dawid (1979)):

$$\text{PTA}_{ij} \perp \left(s_{ij}^{T}(0), s_{ij}^{T}(1)\right) | Z_{ij}$$

Intuitively, this assumption states that, conditional on the set of covariates, potential outcomes are independent of the treatment. In this setting it means that after conditioning on geographical, cultural and historical characteristics of country pairs, there are no such qualities on which trade outcomes depend that also relate to selection into PTAs. Being an identification assumption, unconfoudedness cannot be directly tested.

The second key assumption is overlap (Rosenbaum and Rubin (1983)):

$$0 < e(z) < 1$$

where $e(z) = \mathbb{E}\left(\text{PTA}_{ij}|Z_{ij} = z\right) = \Pr\left(\text{PTA}_{ij} = 1|Z_{ij} = z\right)$ is the propensity score. This assumption means that all country pairs have a non-zero probability of assignment to each treatment condition (either having or not having a PTA). The probability of concluding a PTA between two countries may be very small, but not zero, especially given the dynamic nature of trade.

The combination of these two assumptions implies that the average effects can be estimated by adjusting for differences in covariates between treated and control pairs. The main statistical challenge is now to understand how to estimate objects such as:

$$\tau^T = \mathbb{E}(\ln s_{ij}^T|\text{PTA}_{ij} = 1, Z_{ij} = z) - \mathbb{E}(\ln s_{ij}^T|\text{PTA}_{ij} = 0, Z_{ij} = z) \tag{1.3}$$

The goal is to provide an estimate $\hat{\tau}^T$ without relying on strong functional form assumptions for the conditional distributions. The estimator should also be robust to minor changes in the implementation.

The estimator used in this paper is the blocking estimator. It relies on the initial estimate of the propensity score and uses sub-classification (Rosenbaum and Rubin

(1983), Rosenbaum and Rubin (1984)), combined with regression adjustment within the blocks.

Conceptually, the advantage of the blocking estimator is its flexibility compared to a single weighted regression. In this setting, the blocking estimator serves several important purposes. First, it approximately averages the propensity score and ensures the balance in covariate distributions across treatment groups within the blocks. The implication is that the comparison is made for similar pairs which have similar probability of signing an agreement. Second, as Section 3.3 shows, there are large differences across blocks. In this setting the blocking estimator allows to perform inference within blocks without relying on functional form assumptions and heavy extrapolation. Third, dividing the sample into blocks also uncovers additional heterogeneity across different types of country pairs, which would not be possible to analyze with a simple average effects estimator.

Implementing the estimator requires the estimation of the propensity score, $\hat{e}(z)$. The range of the propensity score is then partitioned into B intervals of the form $[m_{b-1}, m_b)$ for $b = 1 \dots B$. Let $B_{ij}(b) \in \{0, 1\}$ be a binary indicator for the event that the estimated propensity score for a country pair ij satisfies $m_{b-1} < \hat{e}(z) \leq m_b$. Within each block the average treatment effect in each time period is estimated using linear regression with covariates, and the indicator for the treatment (the time period T superscripts are omitted for simplicity):

$$\left(\hat{\alpha}_b, \hat{\tau}_b, \hat{\beta}_b\right) = \text{argmin}_{\alpha,\tau,\beta} \sum_{ij=1}^{N} B_{ij}(b) \left(s_{ij} - \alpha - \tau \text{PTA}_{ij} - \beta' Z_{ij}\right)^2 \tag{1.4}$$

This leads to B estimates $\hat{\tau}_b$ for each $T = \{A, S, M, L\}$, one for each block. To obtain the average estimate over the B blocks, the proportion of treated units in each block, $N_{\text{treat},b}$, is used as weights:

$$\tau_{\text{block, treat}} = \sum_{b=1}^{B} \frac{N_{\text{treat},b}}{N_{\text{treat}}} \hat{\tau}_b \tag{1.5}$$

The next sections show exactly how to implement the estimator in the setting of interest. Section 3.2 explains the procedures to estimate the propensity score and to find the right number of blocks to perform inference. Section 3.5 discusses the regression adjustment and the standard error estimation.

1.3.2 Design: PTA Assignment and Blocking

Understanding the assignment of preferential trade agreements is central to the empirical strategy. This section builds an empirical model of selection to estimate the probability of concluding a PTA for different types of country pairs.

The treatment period runs from 1970 to 2005 in order to estimate both anticipation and long term effects of PTAs.[10] In this setup, the treated country pairs are those that had a PTA entering into force in this period, while the pool of potential control country pairs is comprised of those pairs which never had a PTA. Country pairs which had a PTA before 1970 or after 2005 are excluded from the sample.

To model PTA assignment it is crucial to understand how countries decide to enter a trade agreement. The existing literature on the topic is scarce. Baier and Bergstrand (2004) develop a simple theoretical model, which gives the predictions about the economic factors influencing the likelihood of PTA formation. In their setting, a pair is more likely to conclude a PTA if (1) countries are closer in terms of geographical distance; (2) a pair is more remote from the rest of the world; (3) countries are larger and more similar in size; (4) countries are different in capital-labor ratios; and (5) a pair's difference in capital-labor ratios is smaller with respect to the rest of the world. The economic factors highlighted by Baier and Bergstrand (2004) are informative, but the list of factors is not exhaustive. For example, the static model at the core of Baier and Bergstrand (2004) does not allow to incorporate another strong predictor of PTA formation – the past level of trade, as highlighted by Magee (2008). Additionally, there are numerous other geographical, cultural and political characteristics affecting the likelihood of a PTA formation.

The approach to understand the formation of PTAs in this paper is informed by the literature, but is ultimately data driven.[11] The dataset constructed in this paper contains comprehensive information on country pair characteristics which are relevant for the PTA assignment, and are also correlated with the trade outcomes. A step-wise procedure suggested by Imbens and Rubin (2015) selects a set of covariates that maximizes the predictive power of the empirical model.

The first set of covariates relates to geographical characteristics of country pairs,

[10] The data collected specifically on negotiation and implementation periods of PTAs, shows that the mean negotiation period is about four years, while the mean implementation period is around eight years.

[11] While the theory of PTA formation is outside the scope of this paper, the conclusions discuss some avenues for future research on this front.

and includes variables such as bilateral distances, remoteness,[12] indicators for contiguity,[13] being an island, and being landlocked. Since larger geographical barriers increase trade costs, the expected sign for the coefficients of these variables is negative. The second set of covariates relates to cultural and historical characteristics. Here the following variables are included: an indicator for common language, common colonizer, an existence of colonial relationship in the past, and a common type of the legal system. These characteristics relate to the closeness of two countries, and their expected contribution to the probability of forming a trade agreement is expected to be positive. Finally, there is a set of variables related to trade regulatory environment: membership in the GATT, membership in the EU, and the total number of preferential trade agreements concluded by 1965. The latter intends to capture the increasing likelihood of concluding more agreements in the future in case the countries had PTA experience in the past. Finally, past trade is included as a robust predictor of future PTAs. The idea is that natural trading partners would be more likely to form preferential trade agreements. An important implication of including past trade is that the entire analysis is conducted conditional on positive trade. The next sub-section discusses the question of participation in trade in more detail.

The probability of having a PTA is estimated using a logit regression. The left two columns in Table 1.3 show the results of the estimation in the full sample: the coefficients, the standard errors and the marginal effects. The marginal effects are computed at the means for the continuous variables, and as a switch from zero to one for the binary ones (keeping all the other variables at their means). For example, for a country pair with an average distance, the marginal effect of the distance is a 16% reduction in the probability of signing a PTA (holding other variables fixed at the means). Having a common language, on the other hand, increases the probability of signing and agreement by 6% (again, fixing all the other variables at their means). Table 1.3 shows that the biggest factors contributing negatively to the estimated probability are distance and remoteness. Variables such as having a common language and a common colonizer increase the probability of having a PTA.

The lower panel of Table 1.3 shows the predictive properties of the model. The

[12]The remoteness of a country is calculated as the sum of the bilateral distance from that country to every other country in the sample. The country-pair remoteness is the average remoteness of the two countries.

[13]Contiguity was not selected by the step-wise covariate selection procedure into the final estimation equation.

Table 1.3: Results of the logit estimation of the probability of having a PTA in 1970-2005.

	Raw Sample		**Trimmed Sample**	
	Coefficient	Marginal Effect	Coefficient	Marginal Effect
	(Std. Err.)	(Std. Err.)	(Std. Err.)	(Std. Err.)
Distance	-1.96**	-0.16	-2.07***	-0.45
	(0.05)	(0.004)	(0.07)	(0.014)
Remoteness	-5.26***	-0.42	-5.23***	-1.16
	(0.30)	(0.02)	(0.35)	(0.07)
Small Island	-0.94***	-0.06	-0.96***	-0.18
	(0.08)	(0.004)	(0.09)	(0.015)
Landlocked	0.46***	0.04	0.55***	0.12
	(0.05)	(0.005)	(0.06)	(0.014)
Common Language	0.64***	0.06	0.67***	0.15
	(0.07)	(0.008)	(0.07)	(0.017)
Common Colonizer	0.58***	0.06	0.69***	0.16
	(0.09)	(0.01)	(0.09)	(0.022)
Colonial Relationship	-0.63**	-0.04	-0.81***	-0.15
	(0.19)	(0.1)	(0.21)	(0.03)
Legal System	0.14*	0.01	0.13*	0.03
	(0.05)	(0.004)	(0.06)	(0.01)
GATT Membership	0.22***	0.02	0.12	0.03
	(0.06)	(0.005)	(0.07)	(0.016)
EU Membership	0.91***	0.09	0.90***	0.21
	(0.06)	(0.01)	(0.09)	(0.02)
Pre-treatment Share	0.08***	0.006	0.07***	0.014
	(0.02)	(0.001)	(0.02)	(0.004)
Pre-treatment PTAs	0.11	0.008	0.09	0.02
	(0.07)	(0.006)	(0.07)	(0.02)
Intercept	62.02***		62.72***	
	(2.69)		(3.37)	
N treated	3,200		2,612	
N control	13,392		4,673	
N Total	16,592		7,285	
Pseudo R-squared	0.39		0.19	
Area under ROC	0.89		0.78	
Correctly classified (0.5)	87.4		74.7	

Note: Standard errors in parenthesis. Levels of statistical significance correspond to: * p<0.05, ** p<0.01, *** p<0.001.

pseudo R-squared is equal to 0.39, representing a good fit.[14] The next indicator calculates the area under the Receiver Operating Characteristics (ROC) curve. Since the ROC curve is a probability curve, the area under it indicates how capable the model is to distinguish between the treatment groups: the closer is the value to one, the better are predictive properties of the model. The value of 0.89 means that there is a 89% chance that the model will be able to distinguish the two treatment groups. Finally, assuming that all the pairs with the predicted probability higher than 0.5 are treated, the model is able to correctly classify 87.4% of all the pairs.

The correct estimation of the object in Equation 1.3 requires finding units that would be similar in terms of overlap in their covariate distributions. At the extremes of the propensity score support (close to zero or one) such overlap is lacking, and thus these pairs should be dropped, since they have no counterparts in the other treatment group. A data-driven trimming procedure suggested by Crump et al. (2009) would result in a more robust estimation. The optimal cutoff of the propensity score distribution deletes 8.3% of the support on both sides. The last two columns of Table 1.3 present the coefficients, the standard errors and the marginal effects after trimming.

Figure 1.2 plots the distribution of the predicted probabilities for different treatment groups before and after trimming. The majority of observations without a PTA are concentrated on the lower end of the propensity score, and those are the ones being trimmed. Trimming procedure noticeably improves the overlap in the propensity score and covariate distributions. Table 1.D.2 in Appendix D shows the results of the t-test for balance in covariates and the standardized differences in covariate distributions before and after the trimming.

After the trimming, however, there still remain substantial differences in the distributions of covariates for the observations at the opposite spectrums of the propensity score. The presence of substantial heterogeneity suggests using the blocking estimator proposed by Imbens and Rubin (2015), and described earlier in Section 3.1. The blocking procedure partitions the sample into subsamples (blocks), based on the values of the estimated propensity scores, so that within the blocks, the estimated probabilities are approximately constant. This way the systematic biases in the comparisons of outcomes for treated and control pairs associated with the observed covariates can be eliminated. The causal effect can be estimated within each block as if the PTA assignment was at random. Regression adjustment within each block elim-

[14] Pseudo R-square represents an improvement from a model without any independent variables to a full model. Typically values from 0.2 are considered to indicate a good fit.

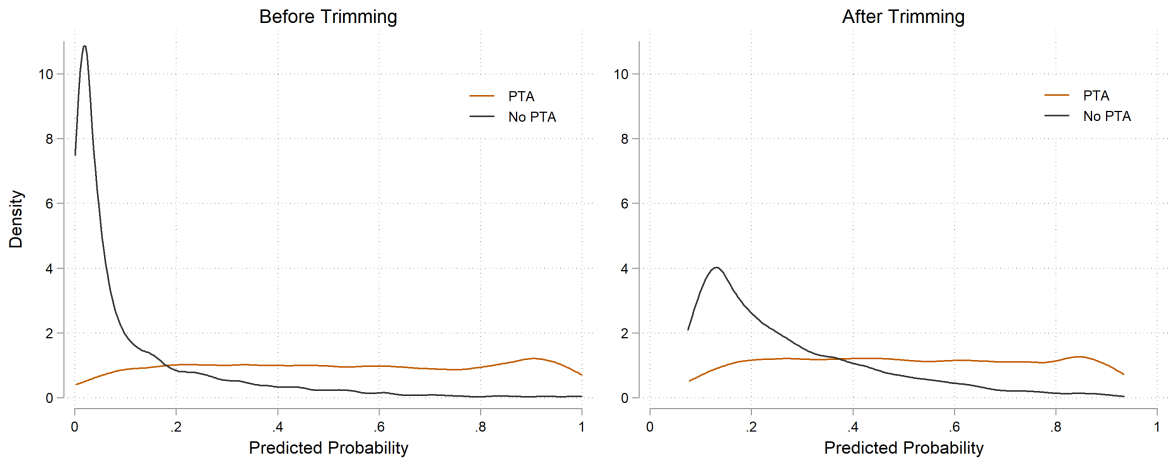

Figure 1.2: Distribution of the propensity score by treatment group before and after trimming.

Note: The figure plots the predicted probability of having a PTA for control and treated pairs before and after trimming. The propensity score is estimated using a logit regression. The trimming cutoff is determined by an optimal data-driven cutoff (Imbens and Rubin (2015)).

inates the remaining differences in covariate distributions across treatment groups. Because the covariates are approximately balanced within the blocks, the regression does not rely heavily on extrapolation as it might do if applied to the full sample.

The main decision in the implementation of the blocking estimator is the number of blocks to partition the data into. The data-dependent procedure for selecting both the number of blocks and their boundaries is proposed by Becker and Ichino (2002). The algorithm starts with the entire sample, and checks whether the average estimated propensity score and the observable covariate distributions between the treated and the control pairs differ. If the test fails, the algorithm splits the sample at the median value of the propensity score and tests again, continuing until the average propensity score and the covariate distributions do not differ between treated and control pairs within the interval (or until the resulting blocks contain too few units to perform inference). As a result of applying this algorithm, the data is endogenously split into nine blocks.

Table 1.4 shows the average value of the estimated propensity score and the number of treated and control pairs within each block. For example, in the first block, the average probability of concluding a PTA is 10%, and eventually, given this probability only 115 pairs sign a trade agreement, while more than one thousand pairs with the same probability do not sign a trade deal. The proportion of treated and control pairs

is reversed for block nine, where the average probability of signing an agreement is almost 90%. In this block, only 24 units do not eventually sign a trade agreement. The blocks in the middle are the most suitable for inference, since they have a probability of signing a PTA close to 50%, and have more or less equal number of pairs in each treatment group. The next subsection characterizes the resulting blocks in more detail.

Table 1.4: The average propensity score and the number of observations in each block.

	B1	B2	B3	B4	B5	B6	B7	B8	B9
Average propensity score	0.10	0.15	0.22	0.31	0.43	0.56	0.69	0.81	0.89
Number of control pairs	1,008	1,028	657	873	524	312	153	81	24
Number of treated pairs	115	186	180	387	405	380	352	360	247
Total number of pairs	1,123	1,214	837	1,260	929	692	505	441	271

1.3.3 Diagnostics: Understanding the Blocks

The blocking algorithm sorts country pairs into different subsamples according to their probability of having a PTA. Lower block numbers correspond to a lower probability of a PTA being concluded. This probability, in turn, is correlated to with country pair characteristics: lower block pairs are, for example, far away from each other, and trade less in the pre-treatment period, in 1960-1965. Such pairs can be labeled as non-natural trading partners. Higher ranked blocks contain pairs which can generally be labeled as natural trading partners: geographically close countries which trade a lot, and have a high probability of signing a trade agreement.

Figure 1.3 plots as examples the means and the confidence intervals for the two covariates – distance and pre-treatment normalized market shares – for each block. It demonstrates the substantial differences between pairs classified to lower and upper blocks. In this setting, using the entire sample to fit, for example, a linear regression, will not correctly account for the covariate imbalances.

One necessary diagnostic is to formally test the balance of each covariate between pairs with and without a PTA *within* each block. Such test is important since, even if the probability of a PTA within blocks is similar, the estimator may still fail to correctly estimate the treatment effects if the covariate distributions are very different across treated and control groups. Assessing the balance in covariate distributions is also indicative of the importance of applying the regressions adjustment at the analysis stage.

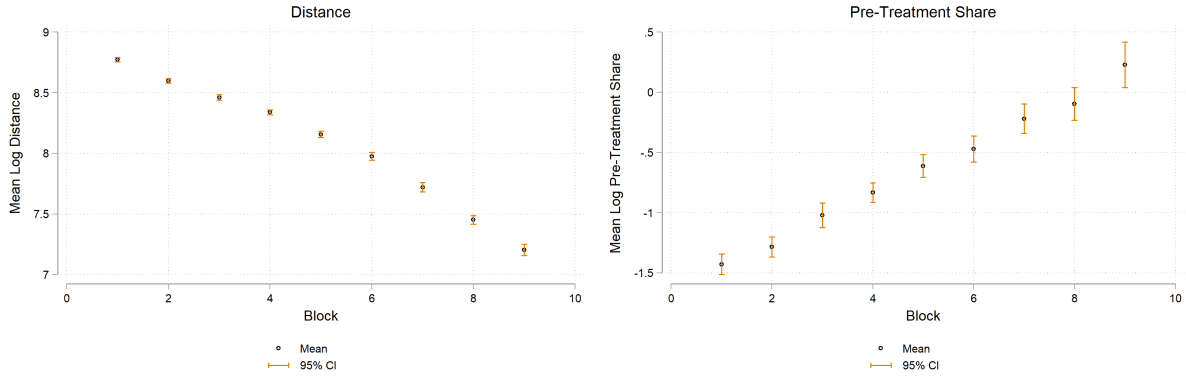

Figure 1.3: Mean and confidence intervals for distance and pre-treatment normalized market share, by block.

Table 1.5 presents the normalized differences between treated and control pairs for each block.[15] The normalized differences are more suitable to analyze covariate imbalances than the simple t-statistic, since they do not increase with the sample size (Imbens and Rubin (2015)).[16] To simplify the analysis of the insights from Table 1.5, the rule of thumb suggested by Austin (2009) states that an absolute normalized difference of 0.1 or more indicates that the covariates are not balanced between groups.

A few important conclusions emerge from the diagnostic analysis. First, the differences in covariate distributions within each block are substantially lower than in the full sample. The only exception is block nine, where, for some variables, the differences still remain large. Second, some differences remain only for a few covariates, suggesting to apply regression adjustment within blocks. To visually confirm the intuition that blocking procedure ensures a much better balance in covariate distributions Figure 1.D.1 and Figure 1.D.2 in Appendix D plot the distributions of the pre-PTA normalized market shares and bilateral distances in the treated and control groups by block. Again, with the exception of the last block, the general conclusion is that the distributions match well across treatment groups.

1.3.4 Robustness Analysis

This subsection explicitly discusses the plausibility of two elements of the empirical design. The first one is the assumption on the uniqueness of the potential outcome, or

[15]The normalized differences are calculated as the difference in average covariate values, normalized by the square root of the average of the two within-treatment group sample variances.

[16]For completeness, Table 1.D.3 in Appendix D presents the results of the t-test.

Table 1.5: The normalized differences by block.

	B1	B2	B3	B4	B5	B6	B7	B8	B9	All
Distance	-0.001	0.26	0.11	0.26	0.05	0.02	-0.51	-0.29	-1.11	0.82
Remoteness	0.02	-0.14	-0.18	-0.06	0.01	0.01	0.27	0.19	0.55	0.27
Small Island	-0.09	-0.12	0.05	-0.11	0.01	-0.003	0.32	-0.13	0.61	0.07
Landlocked	0.20	0.02	-0.02	0.27	-0.12	0.09	-0.41	-0.38	-0.92	-0.04
Common Language	-0.23	0.12	0.06	0.0005	-0.08	0.05	0.05	-0.06	0.03	-0.19
Common Colonizer	0.09	0.15	0.10	0.05	0.03	-0.02	-0.33	-0.09	-0.82	-0.14
Colonial Relationship	0.21	0.04	-0.02	0.08	-0.23	0.01	0.01	0.15	0.38	0.01
Legal System	-0.06	0.15	0.15	-0.01	-0.08	0.02	-0.32	0.13	0.11	-0.15
GATT Membership	0.17	0.06	0.15	0.18	0.05	-0.09	-0.33	-0.58	-0.45	0.03
EU Membership	-0.07	0.06	-0.21	0.11	0.14	-0.06	-0.11	-0.03	-0.15	-0.11
Pre-treatment Share	-0.17	-0.05	0.04	0.10	-0.02	-0.01	-0.02	0.14	-0.49	-0.30
Pre-treatment PTAs	0.14	0.15	0.14	0.01	0.05	-0.18	-0.34	-0.41	-0.79	-0.19

Note: The normalized differences are calculated using the method of Yang and Dalton (2012).

lack of treatment heterogeneity. The second relates to the conditioning of the analysis on positive trade flows, or the extent of the bias associated with the missing trade flows.

Treatment Heterogeneity. One of the assumptions in Section 3.1 is that the unobserved potential outcome of a country pair, $s_{ij}^T(0)$ or $s_{ij}^T(1)$, is unique: with or without a PTA. However, it is clear that there are many different types of PTAs, so each PTA type could potentially have a distinct unobserved potential outcome. In what follows this section investigates the extent of treatment heterogeneity *across* and *within* blocks.

The constructed dataset contains information on various characteristics of PTAs: timing of entry into force, type (free trade area or customs union), composition (bilateral or with many members), notification in the WTO, presence of national treatment and third-party MFN provisions. As an example, Figure 1.4 shows the differences in selected PTA characteristics across blocks. The left panel shows the proportion of PTAs that entered into force after 1993 in each block. While among non-natural trading partners almost all agreements were concluded after 1993, for natural trading partners only around 60% of all treated pairs have later agreements. The right panel shows the differences in the proportion of WTO notifications across blocks. Again, natural trading partners seem to be more likely to notify their agreements, compared to the pairs in the lower-index blocks.

Figure 1.4 shows that there is some variation in the types of agreements that dif-

ferent pairs choose to sign. These differences, however, are not entirely defined by the types of pairs: in the case of the WTO notifications, for example, half of natural trading partners still choose to not notify their agreements. Therefore, the type of agreement and the type of pair signing it are confounded. The average effect of PTAs on trade across all blocks will inevitably reflect both the differences in types of agreements, and the types of country pairs. Similar patterns emerge for other characteristics of trade agreements: type, composition, national treatment provision, and third-party MFN provision.[17]

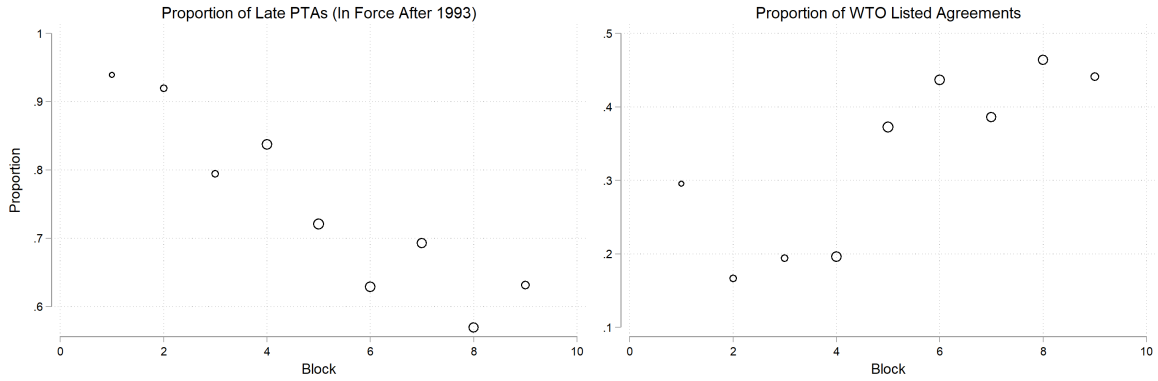

Figure 1.4: The proportion of late PTAs (entering into force after 1993), and the proportion of PTAs notified to the WTO, by block.

Note: The size of circles in each graph is proportional to the number of treated units within the block.

The key element of the design using the blocking estimator is that inference is performed within each block. Thus, the main assumption is that the unobserved potential outcome of a country pair is unique *within* the block. In settings where there is treatment heterogeneity, Imbens and Rubin (2015) recommend redefining the treatment such that the estimates are going to reflect the effects of a randomly selected treatment type. To provide evidence that the treatment types can be treated as random *within* the blocks, it is necessary to test whether the covariates can predict various treatment characteristics. Table 1.D.4 through Table 1.D.9 in Appendix D show the results of the regressions of PTA characteristics on covariates by block. Most of the coefficients appear to be not statistically significant, indicating that these characteristics are independent of the country pair characteristics.

[17]Ideally, the goal is to disentangle the effects of different types of treatment from the reactions of different pairs. Unfortunately, given the large number of characteristics, estimating the effects of each type separately is not possible due to the lack of statistical power.

To sum up, while it is not possible to disentangle the effects of different types of agreements, it is reasonable to assume that the types of PTAs within each block are random. Thus, the individual block estimates represent the effects of a randomly selected agreement, while the average estimate across all blocks represents a combination of the effects of the different types of agreements on the different types of pairs.

Missing Values. The second element of the design that deserves attention is conditioning the entire analysis on the positive trade flows. In the pre-treatment period, more than half of country pairs have missing trade flows. Since in the raw sample it is not possible to distinguish missing trade flows from zeros, the paper uses the imputation procedure discussed earlier in Section 2 and in Appendix B. Imputation recovered 10,804 bilateral trade observations in the pre-treatment period. 50% of the imputed trade flows correspond to low-trade pairs, with exports below 5 thousand USD per year. 90% of the normalized market shares calculated using the imputed data are below one. The majority of these low-trade observations are later cut away by the trimming procedure: trimming deletes 21% of the total sample corresponding to the low values of the estimated propensity score.

Modeling the treatment assignment requires using all the available data obtained after the imputation procedure. Since the level of trade in the pre-treatment period is one of the determinants of the treatment assignment, the probability of concluding a PTA is only defined for countries that trade in the pre-treatment period. Thus, pairs which have missing values in the pre-treatment period, and are not recovered by the imputation procedure, are dropped from the subsequent analysis.

The missing value problem, however, persists in the periods after the treatment. In particular, some pairs which were trading in 1960-1965 have missing values in the anticipation, short, medium, and long run. The missing trade flows could appear as a result of countries ceasing to trade, or as an artifact of the imputation procedure: there is enough data to impute their trade in the pre-treatment period, but there is limited data availability for later years. A simple diagnostic is aimed at testing the extent of the problem: Table 1.D.10 through Table 1.D.13 in Appendix D calculate the proportion of the missing values which were imputed in the pre-treatment period (in every block, for a given time period, and by the treatment status). For example, in the anticipation period in the first block there are a total of 13,104 country-pair-year observations without a PTA, 334 of which are missing. Out of these 334 observations,

trade values for 125 were imputed in 1960-1965. In the same block and time period, out of 115 treated units, there is only one missing value, and it was not missing (or imputed) in the pre-treatment period. The diagnostic is aimed at checking whether the problem is inherently related to the lack of data, or to the lack of balance within the blocks.

A few patterns of missing data emerge. First, there are many more missing values detected in every time period for lower-index blocks than for higher-index blocks. The *proportion* of missing values, however, is similar across blocks. Second, the share of missing values in total observations within the blocks does not exceed 3%: in the example above there are only 2.55% of total values missing in the first block. Additionally breaking by the treatment status, however, a difference is uncovered: there are on average 2.67% of values missing for the control pairs (across all time periods), and only 1.65% for the treated pairs. Third, the proportion of values that were initially imputed is roughly half for the control units in the lower-index blocks.

In sum, the results of the diagnostic reveal that the proportion of missing values differs by treatment status: there are more missing values in the control group than in the treatment group. Moreover, there are more non-traders corresponding to the lower index blocks. Finally, at least half of the missing data problem cannot be attributed to the lack of data to conduct imputation. These insights point to a problem: for the lower blocks the number of non-traders is not balanced between the treated and the control groups. If those pairs were instead trading, their normalized market shares would likely be small. Excluding these pairs from the analysis would thus produce a downward bias in the baseline estimates.

To understand how much the missing value problem could affect the results, an interpolation procedure is used to additionally recover some of the low-trade observations. Interpolation based on a simple linear regression additionally recovers 97,618 observations for the entire sample period: half of the zeros in anticipation, around 35% in the short run, and 20% in the medium and long run. The normalized market shares are then calculated using the newly obtained data. The majority of these observations – 75% of the total – are low-trade observations, with normalized market shares below one.

The paper repeats the full analysis using the average normalized market shares calculated after the interpolation. The final estimates of the PTA effects are indeed lower when using the fuller matrix of trade flows, but the differences with the baseline estimates are negligible: the changes appear to be in the second decimal of the

point estimates (see Figure 1.D.4 in Appendix D for comparison of the final estimates). Thus, while it is difficult to test directly the extent of the missing value problem, interpolation exercise suggests that the estimates are robust to the partial reconstruction of the trade matrix.

1.3.5 Analysis: Estimation and Inference

Finally, the last step in the implementation of the blocking estimator is the regression adjustment within the blocks, described in Equation 1.4 in Section 3.1. For each block the procedure requires running a linear regression with the same set of covariates used for predicting the probability of PTA formation, since those are the factors that can also directly influence trade. The regression controls for year-into-force fixed effects.[18] Within each block and for each time period $T = \{A, S, M, L\}$ the regression takes the form:

$$s_{ij} = \alpha + \tau \text{PTA}_{ij} + \delta \mathbf{Z}_{ij} + \varepsilon_{ij} \tag{1.6}$$

This leads to nine estimates of $\hat{\tau}$ for each T, one for each block (standard errors are clustered at country-pair level). To estimate the effect of PTAs in the entire sample, the block estimates are averaged, using as weights the number of treated units in each block, as shown in Equation 1.5 in Section 3.1.

There is an important aspect of the estimation that relates to the fact that PTAs are being concluded in different points in time. The outcome variable is the average normalized market share at different horizons before and after agreement's entry into force. For treated units, the year of entry into force is well defined and known, so the average shares are easily constructed around that year. Each of the nine blocks contains agreements with different years of entry into force. For example, in the first block, the short run outcome for Israel-USA pair is calculated as the average normalized market share from 1985 (the year of entry into force) to 1989; the short run outcome for Canada-Israel is the average share from 1997 (the year of entry into force) to 2001.

For control units, however, by definition there is no year of entry into force. Since within each block there are treated pairs with different years of entry into force, normalized market shares for the control group are calculated for the same control pairs

[18] Note that these are still defined for control units, corresponding to the treatment years within the block.

around those different treatment years. Continuing with the example above, in the first block, USA-Denmark is a control pair (i.e. never had a PTA), so its average short run normalized market share is calculated in both 1985-1989 and 1997-2001. Thus, this data structure as a form of re-sampling of the outcomes from the control distribution for different treatment years within the block.

The question is whether the re-sampled structure of the data makes a difference for how to interpret the outcomes. For the point estimates there will be no difference, since the coefficient will still show the difference between the average outcomes for treated and control pairs within each block. More precisely, the type of variation used in such estimation is still cross-sectional, where same pairs in different years are treated as different pairs.

Such data structure, however, would require a special inference procedure. Appendix E details two different methods to derive the distribution of standard errors and point estimates. The first method is a standard bootstrap procedure applied within each block: it samples observations with replacement, performs the regression as in Equation 1.6, calculates the mean and the standard error, and repeats these steps one thousand times. The second method performs the same regression analysis with the same number of iterations, but the re-sampling method is tailored specifically for the structure of the data: it samples observations only from the control group, while keeping the treated observations intact at each re-sampling step. Both procedures show that the point estimates, i.e. the nine $\hat{\tau}$ coefficients for each block, correspond to the means of the simulated distributions, while the standard errors are systematically lower without re-sampling. In what follows this paper reports the (more conservative) standard errors which correspond to the means of the distributions resulting from the bootstrap procedure.

1.4 Results

This section summarizes the main results of the estimation, including the estimates in the full sample and across blocks. The second part of this sections reveals and discusses the magnitude of the selection bias which arises in case of using alternative research designs.

1.4.1 Average PTA Effects

The first set of estimates is the average effects of PTAs in different time periods across all blocks. These estimates are obtained by taking a weighted average of all individual block estimates for a particular time period. Table 1.6 shows the average treatment effects of PTAs on their members and the bootstrapped standard errors. The estimates represent the percentage increase in the average normalized market shares caused by PTAs. The full effect of a PTA is a 48% increase in the normalized market share after ten years since the entry of agreement into force. One third of that total effect (16%) is already realized in anticipation, i.e. five years prior to agreement's entry into force. The implementation in the short run (five years since entry into force) is responsible for additional 20 percentage points.

Table 1.6: Average PTA effects in different time periods.

	Anticipation [t-5; t=0)	Short Run (t=0; t+5]	Medium Run (t+5; t+10]	Long Run (t+10; t+15]
Coefficient	0.15	0.32	0.39	0.39
Std. Err.	0.054	0.061	0.065	0.069
Percent	16%	37%	48%	48%

Note: 'Coefficient' is the weighted average of the block estimates from estimating Equation 1.6 for each block within a given time period. 'Standard error' is the mean of the standard error distribution from the bootstrap procedure described in Appendix E. The percentage increase of normalized market shares of treated pairs relative to controls is calculated using the standard formula for interpreting dummy variable coefficients: $\exp(\hat{\tau}) - 1$.

These average effects, however, are not the same across blocks. An additional intuition unveils when comparing the dynamics of PTA effects across different types of country pairs. Figure 1.5 shows the point estimates for each of the nine blocks in anticipation, short, medium and long run. The anticipation effect (the upper-left panel of Figure 1.5) is driven entirely by country pairs in lower-index blocks, while there are no effects for natural trading partners. In the long run these differences in effects across blocks largely disappear, and the same effects are observed within every block (the lower-right panel of Figure 1.5).

The anticipation effects of PTAs have been highlighted in the previous literature (Egger et al. (2020)), and the suggested mechanisms include the actual reduction in trade costs prior to agreement's official entry into force; and firm behavior. The first explanation has been described by policy makers (see, for example US Trade Repre-

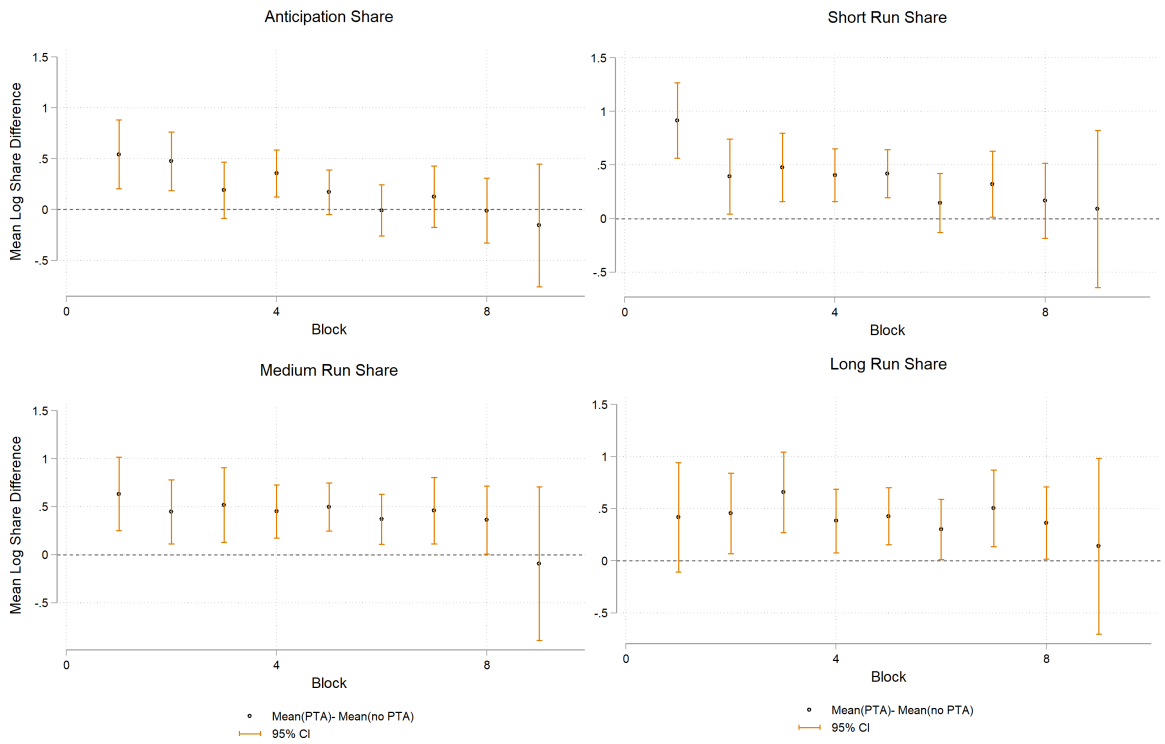

Figure 1.5: Average treatment effects within blocks in different time periods.

Note: The figure plots the point estimates of $\hat{\tau}$'s from Equation 1.6 for each of the nine blocks and each time period. The 95% confidence interval is calculated using the standard errors obtained from the bootstrap procedure described in Appendix E

sentative description of the steps involved into PTA implementation), and relates to a technical procedure which ensures that countries comply with their PTA obligations on the day of entry into force. PTA implementation is a complex process involving the cooperation of many government bodies (ministries, agencies, customs authorities), and the gradual preparation for the actual day of entry into force is necessary. The second explanation relates to the idea that firms adjust their behavior in expectation. Higher future profits would encourage firms to invest more into the new markets and increase trade before the agreement's entry into force.

In light of the second finding – that the dynamic effects are different across types of country pairs – both explanations are reasonable. First, the reduction in trade barriers for less natural trading partners is likely to be disproportionately large relative to country pairs which are close and trade a lot with each other. This explanation emphasizes the potential heterogeneity in the size of trade cost shock, rather than the varying responses of country pairs. Second, the trade behavior of firms may also

differ across different destinations. In distant markets characterized by weak trade connections, firms might want to establish market presence in anticipation of the reduction in trade barriers. In markets of natural trading partners firms might be willing to wait until the barriers are actually reduced.

1.4.2 Estimates and Selection

This subsection discusses the magnitudes of the resulting PTA effects and compares them to the estimates obtained by applying alternative research designs. Recall that the main purpose of the empirical strategy in this paper is to reduce the size of the bias associated with countries' selection into PTAs. In cases which correspond to the standard empirical gravity setup (linear model with fixed effects), the expected sign of the bias is positive, and thus the coefficient overestimates the effects of PTAs, since both the probability and past trade are associated positively with the treatment and the outcome variable.[19]

Besides accounting for the probability of selecting into PTAs and past trade, additional bias reduction comes from improving the balance in covariate distributions between treated and control pairs. Each step of the design – trimming and blocking – is aimed at reducing a part of the bias by ensuring that pairs are comparable. Trimming helps to get rid of pairs which do not have counterparts in the other treatment group in terms of their probabilities to get a PTA. Blocking further groups observations such that the propensity scores are similar and the covariate distributions within each block are balanced, to make comparisons closer to the randomized experiment setup. Finally, regression adjustment takes care of the remaining differences in covariate distributions without heavy reliance on extrapolation or functional form.

Another source of bias which arises in the standard empirical form of the empirical gravity equation is the incorrect form of controlling for economic size. As mentioned earlier, since size is affected by PTAs, a simple conditioning on size may introduce a bias in the treatment estimates. The sign of this bias is likely positive due to the structure of correlation between trade outcomes, size and treatment. This form of bias can persist even when including exporter-time and importer-time fixed

[19] A negative selection bias would only arise if there were omitted variables which would be either correlated positively with trade outcomes and negatively with PTAs, or correlated negatively with trade outcomes and positively with PTAs. Since, the regression adjustment controls for all the confounders which previous research used, the potential remaining omitted variable biases would remain the same.

effects, since those encompass all the factors that are varying across origin or destination countries and time.

Figure 1.6 presents the estimates from alternative designs: without blocking, without trimming and blocking, a linear model with three-way fixed effects (standard in the literature), and two non-parametric techniques. The latter are propensity score matching as in Baier and Bergstrand (2009) and entropy balancing as in Egger and Tarlea (2021). Each dot represents the percentage increase implied by the point estimate in a given period using a given estimator.

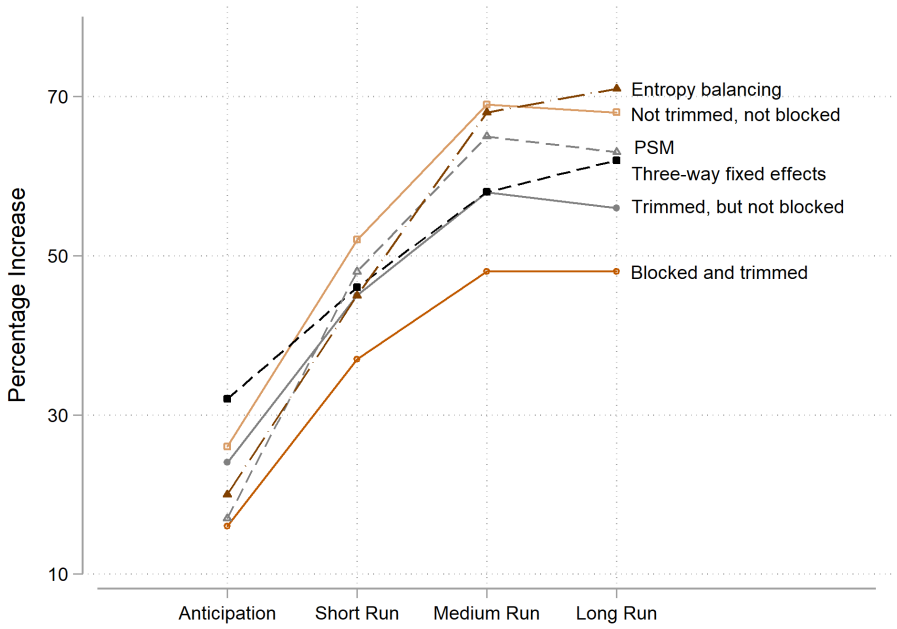

Figure 1.6: Estimates of PTA effects using alternative research designs.

Note: The percentage increase in the outcome variable of treated pairs relative to controls is calculated using the standard formula for interpreting dummy variable coefficients: $\exp(\hat{\tau}) - 1$.

First, it is clear that the estimates presented in Table 1.6 and Figure 1.5 are substantially lower than all the other alternatives estimates in the long run. Second, each part of the research design – trimming and blocking – is responsible for reducing a fraction of the positive selection bias. For example, not blocking the dataset would increase the estimates in each time period by 8-10 percentage points. Trimming has an important bias-reduction property for the long run coefficients: PTA effects are reduced by 20 percentage points when applying the preferred design as opposed to running a regression in the full sample. Third, the comparison to the gravity esti-

mates demonstrates the importance of dealing with economic size. The model with three-way fixed effects (origin-time, destination-time, and country-pair fixed effects) doubles the effects in anticipation - from 16 to 32%, and overestimates the long run effects by 14 percentage points. Finally, an interesting comparison emerges when looking at propensity score matching and entropy balancing estimators: they give very similar estimates in anticipation, but overestimate the effects in the long run by 15% and 23% respectively.[20]

The three-way fixed effects specification may not take into account the positive selection bias of country pair choosing to form a PTA. This may imply that the log-linearized version of the gravity equation may not be a good empirical model to measure the effects of PTAs, or that the real data-generating process is not coming from the structural gravity equation. Setting aside the second possibility, Appendix G provides a numerical example which demonstrates that the empirical version of the gravity equation may indeed not take into account the self-selection bias. In this example, the data generating process is based on the structural gravity equation (following a simple model presented in the next section), and thus it rules out the questions related to the structure of trade.

Given the data-generating process, the simulation estimates the effects of PTAs in two cases: random and non-random assignment of PTAs, for 500 panel datasets (see the details in Appendix G). As Figure 1.G.1 in Appendix G shows, this stylized numerical example clearly demonstrates that the fixed effects estimator may substantially overestimate the effects of non-random PTAs. It is also clear from the figure that the blocking estimator cannot eliminate the bias in the point estimate entirely for the majority of iterations. When combined with the estimates for the standard errors at each iteration, however, the blocking estimator is the only one that includes the true value (see Figure 1.G.2 in Appendix G).

[20]For the propensity score matching estimator, the bias can remain due to the ultimate lack of balance in the covariate distributions, as shown by King and Nielsen (2019). Entropy balancing is essentially a weighting procedure: it calibrates unit weights so that the reweighted treatment and control group satisfy a set of pre-specified balance conditions that incorporate information about known sample moments. The calibration, however, minimizes the divergence from a set of baseline weights chosen by researchers, and thus the method could be inconsistent unless the correct functions are specified.

1.5 From Partial to General Equilibrium: an Application

The estimates presented in Table 1.6 and Figure 1.5 should be interpreted as the average partial equilibrium effects. For example, picking two small countries and endowing them with a PTA will result in their bilateral normalized market shares increasing on average by 48% in the long run, while there will be virtually no effects on other countries. This assumption is plausible if countries are small: diversion of trade from other trading partners would likely be economically small and statistically insignificant. However, it does not mean that the estimates are not suitable for studying bigger PTA formations.

This section uses a standard trade model of Costinot and Rodríguez-Clare (2014) to make predictions about the changes in welfare and trade patterns following the entry into force of the Regional Comprehensive Economic Partnership Agreement (RCEP). RCEP is a free trade area formed between China, Japan, South Korea, Australia, New Zealand and ten Southeast Asian economies (Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar, the Philippines, Singapore, Thailand, Vietnam). The 15 member countries of RCEP account for about 30% of the world's population (around 2.2 billion people) and about 30% of global GDP (29.7 trillion USD), making it the largest trade bloc in history. The agreement was signed in November 2020, and entered into force in the beginning of 2022. The scale and the timing of RCEP make it an interesting and a policy-relevant PTA to study.

1.5.1 The 'Off-the-Shelf' Model

To study the general equilibrium effects and to conduct counterfactual exercises this paper uses the simplest quantitative trade model: the Armington model.[21] The setup and notations closely follow Costinot and Rodríguez-Clare (2014), and are briefly repeated here.

The world economy is composed if $i = 1, \ldots, N$ countries, each endowed with Q_i units of distinct good $i = 1, \ldots, N$. A representative agent in each country has prefer-

[21] The gravity equation, which is a centerpiece of the this model, can be derived from a variety of micro-theoretical foundations and economic environments. The reason to use the simplest model is that it has relatively low data requirements, yet it still captures the main components of the counterfactual exercise. The welfare and trade predictions generated by this model can be interpreted as the lower bound for gains from trade, as shown in Tables 1 and 2 of Costinot and Rodríguez-Clare (2014).

ences characterized by the Constant Elasticity of Substitution (CES) utility function:

$$C_j = \left(\sum_{i=1}^{N} \psi_{ij}^{(1-\sigma)/\sigma} C_{ij}^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)}$$

where C_{ij} is the demand for good i in country j; ψ_{ij} is an exogenous preference parameter, and $\sigma > 1$ is the elasticity of substitution of goods between different countries. The price of good i in country j is P_{ij}, and the consumer price index in country j is given by:

$$P_j = \left(\sum_{i=1}^{N} \psi_{ij}^{(1-\sigma)} P_{ij}^{1-\sigma} \right)^{1/1-\sigma}$$

Trade costs are assumed to be of the iceberg form: $\tau_{ij} > 1$, with $\tau_{ii} = 1$. The price of good i in country j is equal to $P_{ij} = \tau_{ij} P_{ii}$. The domestic price $P_{ii} = Y_i / Q_i$, where Y_i denotes country i's total income. Thus, we can express the price of good i in country j as $P_{ij} = Y_i \tau_{ij} / Q_i$.

Let X_{ij} denote the total value of country j's imports from i, and $E_j = \sum_{i=1}^{N} X_{ij}$ denote country j's total expenditure. Bilateral trade flows satisfy:

$$X_{ij} = \left(\frac{\psi_{ij} P_{ij}}{P_j} \right)^{1-\sigma} E_j$$

Combining the expression for bilateral trade flows, the price index, and the price of good i in country j, the gravity equation is obtained:

$$X_{ij} = \frac{(Y_i \tau_{ij})^{-\varepsilon} \chi_{ij}}{\sum_{l=1}^{N} (Y_l \tau_{lj})^{-\varepsilon} \chi_{lj}} E_j \tag{1.7}$$

where $\chi_{ij} = (Q_i / \psi_{ij})^{\sigma-1}$, and $\varepsilon = \sigma - 1$ is the trade elasticity.

In competitive equilibrium the budget constraint and the goods market clearing imply $E_i = Y_i$, and $Y_i = \sum_{j=1}^{N} X_{ij}$ for all countries. Equation 1.7 together with these two conditions yields the system describing the world income distribution:

$$Y_i = \sum_{j=1}^{N} \frac{(Y_i \tau_{ij})^{-\varepsilon} \chi_{ij}}{\sum_{l=1}^{N} (Y_l \tau_{lj})^{-\varepsilon} \chi_{lj}} Y_j \tag{1.8}$$

In principle, with a simplification that preference parameters do not vary across destinations $\psi_{ij}^{(1-\sigma)/\sigma} = \theta_i$, a numeraire rule for the distribution of incomes ($\sum_i Y_i =$

1), and the data on X_{ij} and Y_i, the model could be calibrated to find the trade costs and demand parameters, by jointly solving Equation 1.7 and Equation 1.8. This, however, is not necessary if the goal is to conduct counterfactual exercises using the model. Instead, this paper uses the approach which became known as the "exact" version of Jones's hat algebra (see, for example, Dekle et al. (2008)).

Consider a shock to trade costs from $\boldsymbol{\tau} = \{\tau_{ij}\}$ to $\boldsymbol{\tau'} = \{\tau'_{ij}\}$ (for example, PTA entry into force). Denote all changes in variables with a 'hat', where $\hat{\nu} = \nu'/\nu$ is the proportional change in any variable ν between the initial and the counterfactual equilibria. Let $\lambda_{ij} = X_{ij}/\sum_l X_{lj}$ be the share of expenditure of country j on goods coming from country i. Since the gravity equation holds in both initial and counterfactual equilibria, the change in the expenditure shares can be expressed using changes in income distributon, changes in trade costs, and the initial expenditure shares:

$$\hat{\lambda}_{ij} = \frac{(\hat{Y}_i \hat{\tau}_{ij})^{-\varepsilon}}{\sum_{l=1}^{N} \lambda_{lj} (\hat{Y}_l \hat{\tau}_{lj})^{-\varepsilon}} \tag{1.9}$$

To then compute changes in the income distribution, use the observation that in the counterfactual equilibrium Equation 1.8 implies:

$$Y'_i = \sum_{j=1}^{N} \lambda'_{ij} Y'_j$$

Combining the two previous expressions we obtain the system of equations defining the changes in the income distribution as follows:

$$\hat{Y}_i Y_i = \sum_{i=1}^{N} \frac{\lambda_{ij} (\hat{Y}_i \hat{\tau}_{ij})^{-\varepsilon} \hat{Y}_j Y_j}{\sum_{l=1}^{N} \lambda_{lj} (\hat{Y}_l \hat{\tau}_{lj})^{-\varepsilon}} \tag{1.10}$$

Equation 1.10 shows that the counterfactual changes in income can be computed without the need to estimate trade costs, endowments or preference shifters. After determining the changes in the income distribution, we the changes in expenditure shares are computed using Equation 1.9. Finally, the changes in real consumption (welfare) are computed[22] using changes in domestic expenditure shares on domestic goods:[23]

[22] In the context of the Armington model the terms 'real consumption changes' and 'welfare changes' are used interchangeably, meaning the percentage change in income that the representative agent would be willing to accept in the lieu of the trade shock.

[23] For the details on the derivation of this result see Costinot and Rodríguez-Clare (2014).

$$\hat{C}_j = \hat{\lambda}_{jj}^{-1/\varepsilon} \tag{1.11}$$

An important thing to note here is that this version of the model studies the static counterfactual equilibrium which would result from the changes in *iceberg trade costs*. In particular, the changes in welfare defined in Equation 1.11 do not take into account the changes in tariff revenue.

There are at least two reasons why this model structure is suitable to study the implications of trade cost reductions such as PTAs. First, in order for the tariff revenue to make a difference for the predictions of the model, the changes in tariffs have to be substantial.[24] For example, Costinot and Rodríguez-Clare (2014) estimate the welfare changes as a function of tariff size, and show that the optimal tariff of around 20% is associated with modest gains from trade (ranging from 0.3% for the US to 1.3% for Ireland). At the same time, the world applied weighted average tariffs since 1988 have not exceeded 10%, and have steadily declined since 1994, reaching 2.7% in 2017 (see Figure 1.D.5 in Appendix D). The data on applied tariffs before 1988 is scarce, but as Bown and Irwin (2015) show, even by the beginning of the Kennedy Round of multilateral trade negotiations in 1964, the average tariffs for the major players in the GATT were about 15%. The average tariffs were reduced to below 10% for the GATT members by the end of the round, and pushed further down by the subsequent multilateral negotiations and the admission of the new members into the GATT.

Second, PTAs include multiple provisions regulating trade in goods which go beyond plain tariff reduction (see, for example, Limão (2016)). Especially since the 1990s, when the majority of PTAs in the studied sample enter into force, trade agreements aim at reducing non-tariff barriers to trade, harmonizing rules, enhancing the efficiency of customs, and covering trade-related rules (such as intellectual property provisions or labor regulations). Therefore, if PTAs were modeled as purely tariff reductions, their trade and welfare effects would likely be substantially underestimated.

Thus, the view about the counterfactual trade cost reductions in this paper is such that PTAs have effects beyond tariffs, and the losses in tariff revenue due to a PTA are not large enough to offset the gains from trade. In fact, the agreement this pa-

[24] The change in welfare in that model would be defined as $\hat{C}_j = \left(\frac{1-\pi_j}{1-\pi_j'}\right)\hat{\lambda}_{jj}^{-1/\varepsilon}$, where π_j and π_j' are the share of tariff revenues in the initial and counterfactual equilibria (see section 4.1 of Costinot and Rodríguez-Clare (2014)).

per studies – RCEP – represents a good case in point. RCEP covers multiple areas relating to trade in goods, trade in services, investment, economic and technical cooperation, and creates new rules for electronic commerce, intellectual property, government procurement, competition, and small and medium sized enterprises.

At the same time, the applied weighted average tariffs in RCEP countries in the year of signature were at 2.12%, comparable with the average world applied tariffs (see Figure 1.D.5 in Appendix D). Table 1.D.14 in Appendix D additionally demonstrates that the highest tariffs among RCEP countries in 2020 were applied by Cambodia (6.21%) and Korea (5.48%), but all the other members have tariffs well below 5%. In fact, the bilateral tariffs of RCEP countries are even lower than the average applied tariffs, since many country pairs had a pre-existing free trade agreement. In particular, the PTA among the ten Southeast Asian nations (ASEAN) was signed in 1992, and completely eliminated tariffs in mutual trade between five countries (Malaysia, Brunei Darussalam, Indonesia, the Philippines, Singapore and Thailand) by 2010, while substantially reducing tariffs among the remaining members. ASEAN as a bloc signed a trade deal with Japan in 2008, with Australia and New Zealand in 2009, with China in 2010, and with Korea in 2010. Thus, since tariff revenue losses are not large for the RCEP countries after the formation of the free trade area, the model outlined in the previous section is suitable to study the effects of this trade agreement.

1.5.2 Application: Regional Comprehensive Economic Partnership

This section shows how the empirical estimates from the first part of the paper can be combined with the model described above, to study policy-relevant questions. The counterfactual exercises use the data for the year 2015, with 88 countries (see Table 1.D.15 in Appendix D) forming 7744 country pairs.[25] The model uses trade flows for that year and computes the income distribution as a share of each country in the total world income.

This section presents two types of counterfactual exercises. The first one endows RCEP countries with a trade agreement using the long run average estimates ob-

[25]Most of the domestic trade flows, in particular, for the RCEP members, which are necessary to conduct the general equilibrium exercises, are available for year 2015. At the time of writing, UNIDO manufacturing production data for 2019 and 2018 is unavailable for such countries as China, Japan, and Korea, as well as many others.

tained in the empirical part of the paper.[26] Setting the trade elasticity $\varepsilon = 5$,[27] and given the long run estimate of 48% increase in normalized market shares, the shock corresponds to a 9.6% reduction in iceberg trade costs for the RCEP members. The second type of the counterfactual exercise exploits the full heterogeneity of the empirical estimates. Although the model does not feature any dynamics, it can still be used it in a series of static exercises to study changes in real consumption and trade reallocation in the anticipation period and long run.

Counterfactuals: Long Run. Given the 9.6% reduction in iceberg trade costs for all RCEP members, the model predicts the simple average reduction in welfare of 0.05% (see Figure 1.D.6 in Appendix D for the distribution of changes in welfare, income, expenditure shares and normalized market shares across all countries). Weighted by the initial share in the total world income, however, the average change in welfare is predicted to be positive, although negligible (0.0005%). Figure 1.7 maps the percentage changes in real consumption following trade shock.

Naturally, RCEP members are the winners in terms of welfare after the PTA formation. The biggest gain is recorded for Myanmar, with a 18.3% increase in real consumption. The effect comes from both the increase in size by 9%, and the reduction in the price index by 7.9% (see Table 1.D.15 in Appendix D for the decomposition of the changes in real welfare into size and price effects). Myanmar experiences by far the largest effect, followed by Cambodia with 1.98%. The simple average gain for RCEP economies equals 1.75% (0.24% without Myanmar). However, since large gains are recorded for smaller countries, like Myanmar and Cambodia, while China, Japan and Korea gain less than one percent each, the weighted average gains are quite modest (0.0018%).

For the rest of the world changes in real consumption are negligible, constituting less than half a percent change on average. The biggest gains outside of the block

[26] Since the estimates in the empirical part do not differentiate between different types of trade agreements, this simplification implies that the exercise treats RCEP as an 'average' trade agreement. It is a plausible approximation, since RCEP includes many of the elements of modern trade agreements, while tariff levels among its members are at the world average.

[27] Appendix H provides the sensitivity checks using alternative values of elasticity. The role of elasticity is two-fold in the model: on one hand, it amplifies the trade effects of trade cost changes, but on the other hand it decreases the magnitude of reductions in iceberg trade costs. Appendix H demonstrates that the values of elasticity influence primarily the distribution of the growth rates of normalized market shares for the RCEP economies, while having relatively little differences in the welfare growth rate distribution.

are recorded for Congo, with the increase of 1.1%. The main losers from the formation of RCEP in the long run are small countries outside of the block, such as French Polynesia (8.3% reduction in real consumption), Lebanon (6.2% reduction), and El Salvador (3.1% reduction). For all of these countries, even though the price index is decreasing, the reduction in size dominates (again, for the decomposition see Table 1.D.15 in Appendix D).

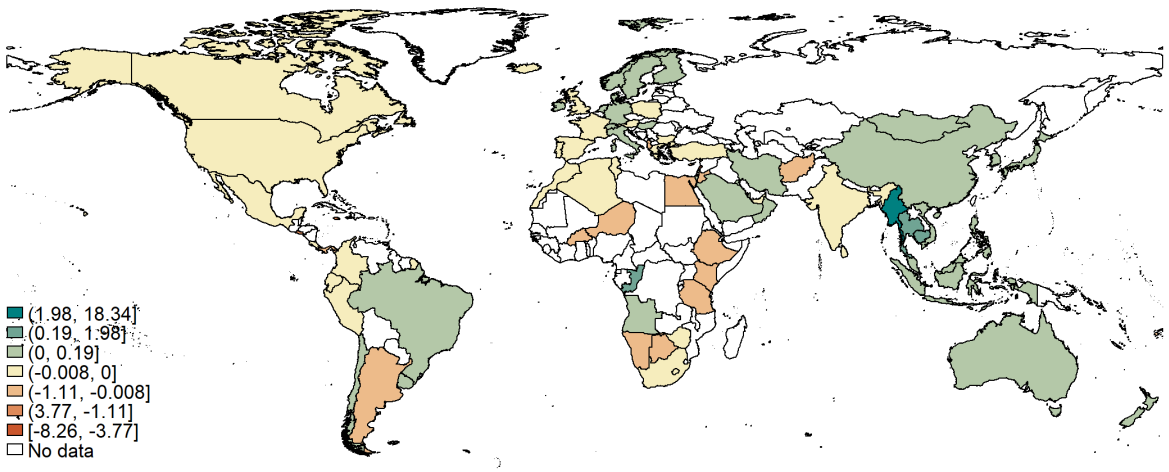

Figure 1.7: Percentage changes in welfare following the RCEP entry into force in the long run.

Note: The shock corresponds to a 9.6% reduction in iceberg trade costs for the RCEP members (using the estimated PTA effects and the trade elasticity of $\varepsilon = 5$).

Next, the model can be used to analyze the changes in the trade patterns following the shock. Figure 1.D.7 in Appendix D plots the distributions of the gross growth rates of normalized market shares of different groups of countries. In the new equilibrium, almost all RCEP members redirect trade towards each other (on average their normalized market shares increase by 56.24%), while reducing exports to the outside world (on average normalized market shares with the outsiders fall by 23.65%). Similarly, the countries outside of RCEP start trading more within themselves (on average, outsiders' normalized market shares in mutual exports increase by 21.37%). As an example of trade pattern change, Figure 1.8 maps the changes of China's normalized market shares with other countries. China increases its normalized market shares primarily with the RCEP countries, such as Malaysia (76.72%), South Korea (67.95%), and Indonesia (58.08%). Among the countries that China trades less with in the new equilibrium are small economies, which are highly dependent on China's

trade, but are not a part of RCEP, such as Macao (-73.26%) and Hong Kong (-64.58%). Notably, China decreases the share with its largest market – the domestic one – by a considerable 2.3%.

As shown in Section 4, alternative research designs overestimate the effects of preferential trade agreements. Appendix I repeats the counterfactual exercise using the estimates from the three-way fixed effects gravity regression. The model uses the partial equilibrium estimate of 68% and the same value of elasticity, $\varepsilon = 5$, which translates into 13.6% reduction in iceberg trade costs. The average changes in real consumption are very similar for the two cases, and the simple t-test cannot reject the null hypothesis that the difference is equal to zero for the two welfare vectors. Since the changes in real consumption are negligible for most of the countries in both versions of the model, this result is not surprising. There are, however, substantial differences in growth rates of normalized market shares for the RCEP countries. While the mean increase in the baseline model is 56.2%, it is almost double when using the gravity estimates (90.9%). Appendix I demonstrates that correctly identifying partial equilibrium estimates matters for the predictions of normalized market shares growth rates of the directly affected countries.

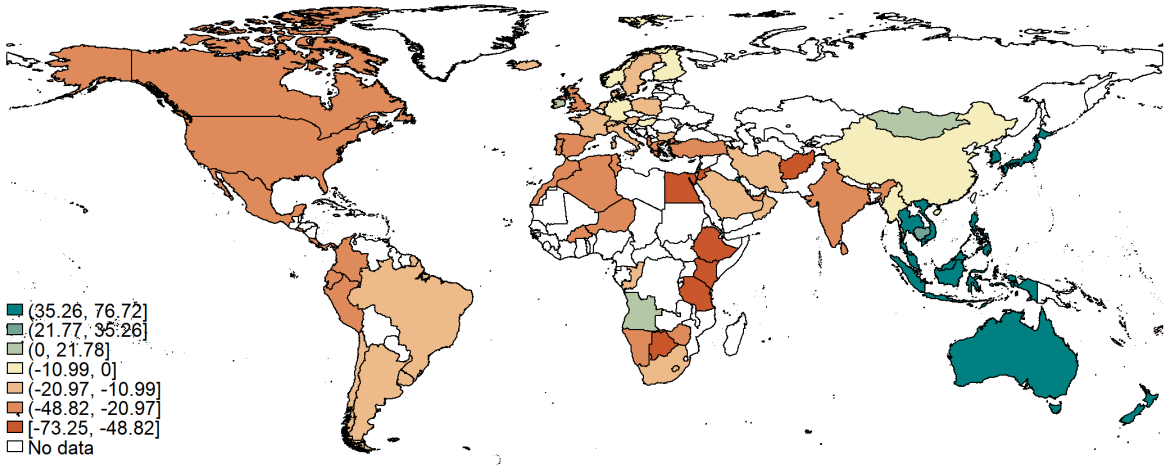

Figure 1.8: Percentage changes in China's normalized market shares with other countries following the RCEP entry into force in the long run.

Note: The shock corresponds to a 9.6% reduction in iceberg trade costs for the RCEP members (using the estimated PTA effects and trade elasticity of $\varepsilon = 5$).

Counterfactuals: Transition to the Long Run. To construct the reductions in iceberg trade costs for different time periods, this exercises uses point estimates from the

empirical part corresponding to different blocks and the value of the trade elasticity $\varepsilon = 5$.[28] In anticipation there are substantial differences for point estimates, while in the long run they are similar across blocks (with the exception of block nine). Table 1.D.16 in Appendix D gives more details in the coefficients and the corresponding reductions in iceberg trade costs used in the counterfactual general equilibrium exercise. Among the RCEP economies, there is only one country pair which belongs in the first block (i.e. the lowest probability of forming a trade agreement), which is Myanmar and New Zealand. Other examples of pairs in lower-index blocks include Myanmar and Korea or Australia and Cambodia. Blocks nine and eight have the most number of pairs (32 and 33 pairs respectively), indicating that the majority of RCEP members are natural trading partners. Those blocks include pairs such as Vietnam and Thailand, or China and Korea. The trade costs reductions are applied sequentially, i.e. the counterfactual equilibrium resulting from the shocks in anticipation period is used as a baseline equilibrium for the long run shocks.[29]

Figure 1.9 maps the percentage changes in real consumption in anticipation and long run using the heterogeneous block estimates. In anticipation the only country which experiences a decline in welfare (although negligible) is Japan.[30] With the exception of Myanmar, which increases its real consumption by 4.03%, changes in welfare for RCEP countries in anticipation are negligible (simple average of 0.06%, and weighted average of 0.0005%). In the long run, again, Myanmar's gain of 9.72% by far exceed those of other RCEP countries (simple average gain of 0.11% and weighted average gain of 0.0007%).

Similarly to the previous counterfactual exercise, the model can also be used to analyze the changes in trade patterns after the shock in anticipation and in the long run. Table 1.D.17 in Appendix D provides the model-implied average changes in normalized market shares of RCEP members by block. Normalized market shares of RCEP members in trade with each other increase on average by 15.76% in anticipation, and by 25.84% in the long run. In anticipation less natural trading partners within RCEP (blocks 1-4) experience growth in mutual normalized market shares, by

[28] Again, Appendix H provides the sensitivity checks using alternative values of elasticity. It demonstrates that trade flows (normalized market shares) can be sensitive to elasticity values, while it is not true for the welfare growth rates.

[29] I.e. the reduction in the iceberg trade costs from the anticipation to long run period is defined as the difference between these two periods.

[30] This happens because all country pairs including Japan as an exporter or importer are sorted into higher-index blocks, which have no reductions in trade costs in anticipation.

36.37% on average, while natural trading partners do not experience any substantial changes in bilateral trade. In the long run, on the contrary, natural trading partners are the ones experiencing most growth (41.16%), while pairs distributed to lower blocks experience mild changes in trade patterns.[31]

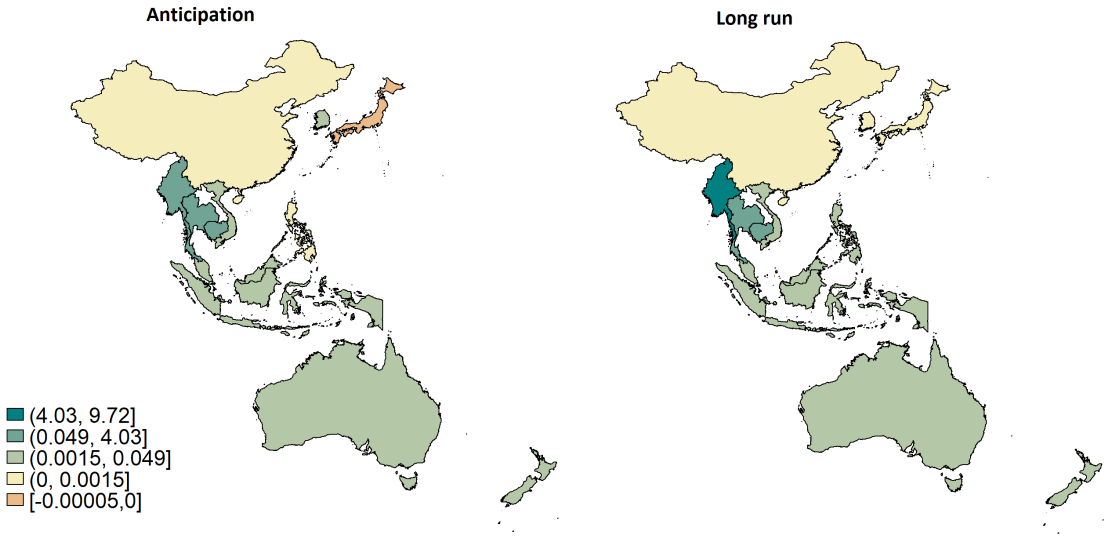

Figure 1.9: Percentage changes in welfare in anticipation of RCEP formation, and in the long run.

Note: The shock corresponds to reductions in iceberg trade costs specified in Table 1.D.16 in Appendix D for different blocks (using the estimated PTA effects and the trade elasticity of $\varepsilon = 5$).

Figure 1.10 maps percentage changes of China's normalized market shares with its RCEP partners in anticipation and long run. In anticipation China decreases trade with a few more natural trading partners, such as Japan in block nine (-16.04%), Philippines in block nine (-15.28%), and Vietnam in block five (-13.81%); while redirecting trade towards Indonesia in block four (30.89%), and New Zealand in block two (25.53%). In the long run, China increases its trade with all RCEP members (except Myanmar), with normalized market shares for Vietnam (+68.67%), Philippines (+51.63%) and Malaysia (+50.25%) rising the most. China also reduces domestic trade in anticipation of RCEP (by 2.54%), while there is almost no change in it in the long run (0.08%).[32]

[31]With the exception of block one, which contains only one country pair (Myanmar and New Zealand).

[32]The two counterfactual exercises provide different perspectives on the formation of PTAs. The first exercise assumes larger changes in trade costs over the period over 20 years years (long run). The

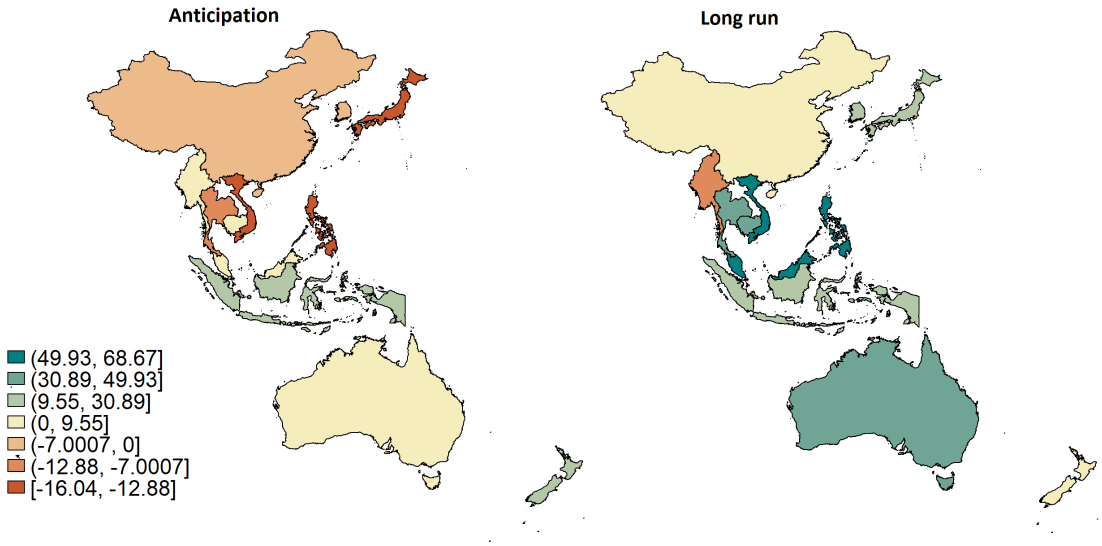

Figure 1.10: Percentage changes in China's normalized market shares with other RCEP members in anticipation of RCEP formation, and in the long run.

Note: The shock corresponds to reductions in iceberg trade costs specified in Table 1.D.16 in Appendix D for different blocks (using the estimated PTA effects and the trade elasticity of $\varepsilon = 5$).

To sum up the results of the two exercises, the trade shocks from RCEP formation in the model have small effects on the real consumption of most countries. However, the model predicts large trade creation effects, i.e. the increase in trade (normalized market shares) of RCEP economies within the PTA.

1.6 Conclusion and Future Research

This paper estimates the effects of preferential trade agreements on trade between members. Using the causal inference framework to address the self-selection bias, it shows that PTAs gradually increase bilateral trade, starting prior to the agreement's entry into force. Fifteen years later member countries trade 48% more relative to similar pairs without an agreement. Although the long-run effects are similar across country pairs, there is substantial heterogeneity in the dynamic responses. In partic-

estimate used in this exercise is a weighted average of the individual block estimates. The second one considers smaller changes in the years preceding PTA formation and in the first five years after the PTA enters into force (anticipation and short run), followed by additional reductions some ten years after that. The estimates used in this exercise are different across pairs. Thus, the cumulative gains from the anticipation and long run of the second exercise are not supposed to add up to the gains from the first exercise.

ular, only non-natural trading partners react in anticipation.

The findings in this paper open up a few important questions for future research. A natural first question is what is driving the increase in trade between non-natural trading partners prior to the actual reduction in trade costs. One potential explanation requires looking closer at the firm behavior. To understand this phenomenon both an appropriate theoretical model and a thorough empirical investigation are necessary. On the theory side, one could think about alternative models of firm behavior, where, for example, firms would want to invest in a future prospective market, acquire or expand a customer base, and reap the benefits of the first-mover advantage in remote markets.

To empirically check these theoretical alternatives, in a separate project I use Colombian monthly firm shipment-level data across different destinations. The collected data includes extensive information on Colombia's PTA partners from 2006 to 2020, as well as detailed information about the negotiation process (the announcement of the negotiations, the timing of the negotiation rounds, the treaty signature, etc.). Colombia has signed a number of trade agreements both with natural (for example, Peru), and non-natural trading partners (for example, Israel). Exploiting the variation across destinations for Colombian firms exports, the project analyzes trade patterns over the life cycles of trade agreements. The universe of transactions is instrumental in disentangling the contributions of extensive and intensive margins of trade.

The results of the empirical investigation of Colombia's firm-level trade flows across destinations open the next set of questions: why do we observe these differences in responses across destinations? Are different types of firms selecting into exporting to a particular type of the destination? Or are the same firms behaving differently in different destinations? These are the questions that the project is seeking to answer next.

There are alternative explanations to the observed anticipatory effects that also deserve attention in future research. One possibility of why firms might increase their trade volumes in anticipation of a PTA is that the change in regulation itself might involve an additional cost. An example of such costs are the complex rules of origins that some PTAs impose. Combining product-level data for Colombia with specific provisions for rules of origins or standards within PTAs could shed light on whether firms react to the associated regulatory changes by increasing their trade volumes prior to the actual change.

Another possible reason for the anticipatory increase in trade may be that there

are actual trade cost reductions prior to a PTA's entry into force. Since trade treaties represent a complex set of rules, the regulators might want to start implementing the changes beforehand, to avoid a customs blockage on the day of entry into force. Checking this hypothesis would require high-quality and high-frequency data on applied tariffs or other regulations. Collecting such data from the bilateral schedules laid out in detail in the attachments to each trade agreement would contribute to the efforts of the trade economists to understand the impacts of well-defined trade policies.

The second prospective avenue for future research concerns a puzzle uncovered in this paper: the type of country pair does not seem to define the type of the trade agreement this pair will sign. In principle, the hypothesis would be that natural trading partners tend to sign more comprehensive and binding trade agreements. This paper, however, estimated the effects of an 'average' trade agreement precisely because the characteristics of country pairs cannot predict the various features of PTAs (for example, being a customs union or a free trade area) within blocks. In addition, there is limited variation in agreement types across blocks. Since the empirical design used in this paper cannot be used to study each type of trade agreement separately due to power issues, the future research plans aim at tackling this issue in a more hands-on manner.

The first step is to classify PTAs according to their structure, content, coverage, depth, length, and legal enforceability. Given the high dimensionality of PTA heterogeneity, the first challenge is to understand the features of agreements that matter most for trade flows. In an ongoing project I utilize the digitized legal texts of concluded preferential trade agreements to construct a novel dataset with individual characteristics of PTAs. The next step is to use this new comprehensive dataset to study the effects of the relevant regulatory features of PTAs on trade flows.

Finally, this project draws attention to the importance of selection bias in estimating the effects of trade policy. While there are tools to estimate the consequences and welfare implications of implemented PTAs, there is limited knowledge about the mechanisms that lead to PTA formation in the first place.

This paper used a prediction model to empirically estimate the probability of self-selecting into a PTA given characteristics of country pairs. While some factors – such as economic size, geographical proximity and past trade – have been highlighted in the literature, they still do not explain all the patterns of PTA formation. For example, given these factors, we should observe a strict hierarchy in PTA formation (akin

to that to firms exports across destinations). Yet, there is a large number of trade agreements concluded between non-natural trading partners. Additionally, comparative advantage or factor endowments can act as possible economic drivers behind the formation of PTAs. Besides economic forces, there is a number of political economy factors (military coalitions or past war conflicts) that matter for the formation of PTAs. While the model of selection is ultimately a theory question, it should be informed by stylized empirical facts. Given that the results of this paper indicate that self-selection produces a large bias that may ultimately misinform policy decisions, the understanding of the assignment mechanism (i.e. how countries decide to form PTAs) should be high on the research agenda.

Appendix

1.A Data Construction

Bilateral Trade

To construct trade flows from origin i to destination j, I unite the following databases: International Trade and Production Database for Estimation (ITPD-E); WTO Structural Gravity Database; IMF Direction of Trade Statistics Database (data retrieved in 2018); World Trade Flows (WTF) bilateral cross-sectional data; NBER-United Nations Trade Data; and CEPII Gravity Dataset.

Table 1.A.1 shows the parameters of each raw dataset: the number of unique countries and country pairs, the time span of the data, the number of observations, the number of missing values; and whether the dataset is a balanced panel. Since ITPD-E, WTO, IMF and WTF datasets only report positive trade flows, they do not contain missing values. However, these datasets, if transformed into balanced panels, will contain a lot of gaps in both cross-sectional and time dimensions. The CEPII Dataset itself collects trade data from several sources, including UN Comtrade, CEPII BACI Database, and IMF Direction of Trade Statistics. The number of missing values varies across different sources.

Table 1.A.1: Metadata for raw bilateral trade datasets.

Name	Countries	Pair	Years	Observations	Balance	Missing
ITPDE-E	237	43,623	2000-2016	714,951	No	0
WTO	229	48,711	1980-2016	972,692	No	0
IMF	218	47,030	1948-2017	2,710,148	No	0
WTF	263	50,456	1984-2015	750,556	No	0
NBER	201	23,750	1962-2000	926,250	Yes	499,365
CEPII	248	61,034	1948-2019	3,661,898	No	UN exporter: 2,843,970 UN importer: 2,731,663 BACI: 3,056,279 IMF exporter: 2,770,880 IMF importer: 2,687,346

Note: The number of observations for the CEPII Gravity Dataset is reported after deleting non-existing countries and domestic trade flows.

Since the datasets use different country identifiers, I use concordances to use ISO-3 codes as identifiers throughout. I also make sure that the values are reported in USD across all data sources. I proceed to unite the datasets in the following order:

1. Merge ITPD-E and WTO datasets, gaining 193,597 trade flow observations.

2. Merge the resulting dataset with IMF data, gaining additionally 561,915 observations.

3. Merge the WTF and NBER datasets, and then merge the resulting dataset with the one created at the previous step, resulting in 242,534 additional observed trade flows.

4. Finally, I unite the dataset resulting from step 3, with the CEPII dataset, and construct the final trade volume variable in the following order:

    - Start with IMF data reported by the exporter;
    - Substituting the missing values with UN Comtrade data reported by the exporter (gaining 188,441 observations);
    - Substituting the missing values with UN Comtrade data reported by the importer (gaining 118,152 observations);
    - Substituting the missing values with IMF data reported by the importer (gaining 30,860 observations);
    - Substituting the missing values with BACI data reported by the exporter (gaining 1,228 observations);
    - Substituting the missing values with data constructed in steps 1-3 (gaining 611,237 observations);

I then delete countries that did not exist throughout the whole period of time from 1960 to 2019. The resulting dataset contains 210 unique customs territories, forming 43,890 pairs over the period of 1960-2019. The total number of observations is 2,633,400 in a balanced panel. The number of missing observations is 1,613,684. I then use this dataset for imputation (see Appendix B).

Domestic Trade

In order to construct domestic trade flows from i to i, I complement the data from ITPD-E and WTO with data from TradeProd Database and UNIDO's INDSTAT Rev. 4 Database. Table 1.A.2 shows the characteristics for the datasets with domestic flows (for ITPD-E and WTO datasets) and production (for TradeProd and INDSTAT databases): the number of unique countries, year coverage, and the number of observations.

Table 1.A.2: Metadata for raw domestic trade datasets.

Name	Countries	Years	Observations
ITPD-E	115	2000-2014	1,356
WTO	160	1980-2016	3,645
TradeProd	180	1980-2006	4,514
INDSTAT	137	1980-2016	3,349

ITPD-E and WTO datasets contain ready-made information on domestic trade flows for some countries and years. In particular, after merging them I have information on 3,084 domestic flows out of the total 7,104 observations (for 192 unique exporters over the period from 1980 to 2016). I then add observations from CEPII TradeProd database, additionally gaining 2,286 observations. I then add observations from INDSTAT Database, gaining 256 observations. Note that since CEPII TradeProd and INDSTAT report production data, I calculate the domestic trade flows as the difference between production and total exports of a country in a given year. I then use this dataset to show that normalized market shares calculated with and without domestic trade flows do not have substantial differences (see Appendix C).

PTAs

To construct the PTA indicator and extract the information about the agreements, I use Design of Trade Agreements Database (DESTA version 2.0, Dür et al. (2014)). The dataset contains all trade agreements ever concluded, both notified and not notified to the WTO, as well as:

- Superseding agreements: for example, Andean Group was formed through a series of agreements – Cartagena Agreement 1969, Quito Protocol 1988, Trujillo Protocol 1997, Sucre Protocol 2003;
- Overlapping agreements: for example, Colombia and Peru are both in Andean Group (Bolivia, Colombia, Ecuador, and Peru) and in Pacific Alliance (Chile, Colombia, Mexico and Peru);
- Accessions: for example, Venezuela joined Andean Community in 1973;
- Withdrawals: for example, Venezuela withdrew from Andean Community in 2006.

To take into account agreements' dynamic, I use the following cleaning protocol:

1. Start with the list of all baseline treaties (without accessions or withdrawals);
2. Filter only Free Trade Areas (FTAs) and Customs Unions (CUs), i.e. delete all Partial Scoope Agreements (PSAs), Framework Agreements, Services Agreements, and Sectoral Agreements;
3. Clean from superseding agreements, amendment protocols, revisions, leaving only the earliest agreements;
4. Represent the dataset in dyadic form;
5. Clean from overlapping agreements[33];
6. Separately recode accessions and withdrawals to dyadic form. For accessions,the entry into force is coded as the year of accession (there are 852 of such country pairs over the whole period). For withdrawals, I code only 'real' withdrawals, i.e. only the cases when countries stop having any type of formal preferential trade arrangement:
    - Brazil-Venezuela from 2006 to 2012: Venezuela exited Andean Community to join MERCOSUR, but was not a member until 2012;
    - Eritrea with Angola, Lesotho, Mozambique, Namibia, Tanzania when the latter exited COMESA;
    - Georgia with Belarus, Kyrgyzstan, Tajikistan when Georgia exited CIS;
    - The rest of the 486 country pairs which formally withdrew from PTAs had another PTA in place. For these pairs, the withdrawal is related to restructuring, for example, joining the EC and thus withdrawing former agreements, while joining those that the EC has.
7. Create a symmetric matrix.

[33]If two overlapping agreements were in the same year, leave the 'strongest' in terms of agreement characteristics (has a national treatment clause, is a Customs Union, is a bilateral agreement, has the metadata available); if two overlapping agreements were in different years, leave the earliest agreement

The resulting dataset contains a total of 9,168 symmetric dyads in 398 unique PTAs (410 PTAs counting accessions). I also collect the metadata for the agreements available in DESTA: the type of agreement (FTA or CU), regional composition, the year of signature, entry into force, the implementation period, the composition (bilateral, plurilateral, region-region), notification to the WTO, the presence of national treatment and third-party MFN provisions. Table 1.D.1 in Appendix D presents the descriptive statistics for the final PTA dataset, after it is merged with trade flows and other variables.

Other Variables

Geographical and cultural characteristics come from CEPII Gravity Dataset. In particular, I use bilateral distances, information on common language, colonial past, legal system, and information on GATT and EU membership. I construct a measure of remoteness as the sum of bilateral distances from a given country to every other country in the sample. To get a country-pair remoteness, I average the remoteness of two countries. I complement these variable with the information from NASA's Earth Observing System Data and Information System (EOSDIS), where I take information on insularity (small island developing economy), and the indicator for being landlocked.

1.B Data Imputation

As shown in Appendix A, even after combining all available data sources containing trade flows, many missing values remain: out of 2,633,400 observations 1,613,684 (or 61%) are missing. Figure 1.B.1 shows the percentage of missing observations for selected variables: trade volume, GDP and distance. Almost 90% of trade data and 70% of GDP data is missing for the period before 1960. Therefore, in everything that follows, I will focus only on the period after 1960.

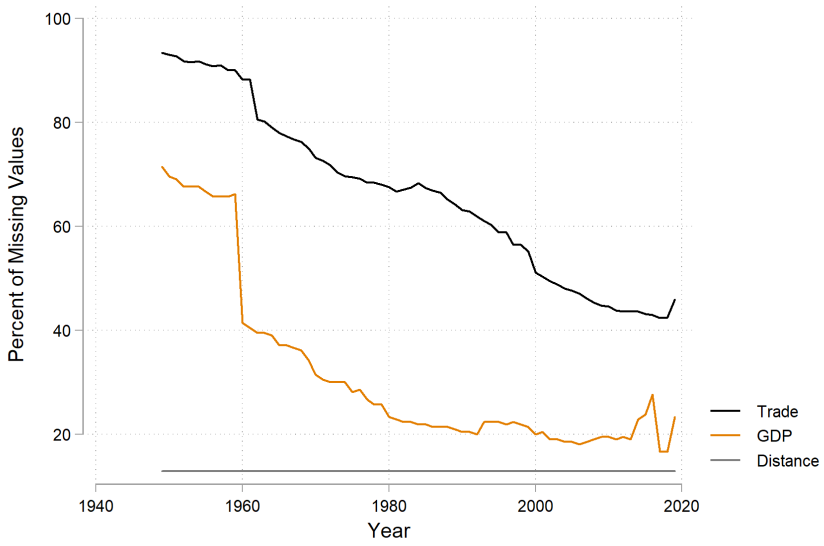

Figure 1.B.1: Percentage of missing observations in the final dataset, 1950-2019.

One way to treat missing observations is to declare them as zeros, assuming that countries do not trade in a given year. The main problem is that it is virtually impossible to distinguish true zero trade flows and non-reported trade volumes. Appendix A demonstrated that adding up various data sources may substantially reduce the number of missing observations, suggesting that some of those flows are not zeros after all. Additionally, there are 35,411 missing trade flow observations for active PTAs (21.09% of all country-pairs with active PTAs). It is unlikely that countries would spend resources to negotiate an agreement if they do not trade. Moreover, there are some data patterns that suggest that some flows might indeed be missing, namely:

- 45,742 observations not missing at t and $t+2$, but missing at $t+1$;
- 21,259 observations not missing at t and $t+3$, but missing at $t+1$ and $t+2$;

- 11,621 observations not missing at t and $t+4$, but missing at $t+1$, $t+2$, and $t+3$;
- 4,664 missing observations for neighbouring countries.

In order to predict the values of missing trade flows, I use the fact that the empirical gravity relationship – even though not suitable for causal interpretation – has very high predictive power. I use a flexible form of log-linearized gravity equation, where I interact bilateral distance with the year indicators, to take account of the change in trade costs over the past 60 years. Using all available data, I estimate the 266 parameters of the following equation:

$$\begin{aligned}\log(X_{ijt}) = \beta_0 + \beta_1 \log(GDP_{it}) + \beta_2 \log(GDP_{jt}) + \sum_{q=2}^{4} \gamma_{qt} Dist_{ij} \times \delta_t + \beta_3 Colony_{ij} + \\ + \beta_4 Comcol_{ij} + \beta_5 Language_{ij} + \beta_6 Contiguity_{ij} + \beta_7 Legal_{ij} + \beta_8 GATT_{it} + \beta_9 GATT_{jt} + \\ + \beta_{10} EU_{it} + \beta_{11} EU_{jt} + \beta_{12} PTA_{ijt} + \beta_{13} NumPTA_{it} + \beta_{14} NumPTA_{jt} + \beta_{15} Landlock_{ij} + \\ + \beta_{16} SIDS_{ij} + \beta_{17} SameReg_{ij} + \beta_{18} \log(Pop_{it}) + \beta_{19} \log(Pop_{jt}) + \varepsilon_{ijt} \quad (1.12)\end{aligned}$$

Since the regression is estimated without domestic trade flows (recall that domestic trade data is only available after 1980), the distance puzzle persists in the estimation (Yotov (2012)). The problem is less pronounced, however, since I am using the flexible specification with distance quartiles: the interaction coefficients for the 75th percentile in Figure 1.B.2 almost do not change, while the ones for the 25th percentile fall only from -0.05 to -0.1, relative to the baseline.

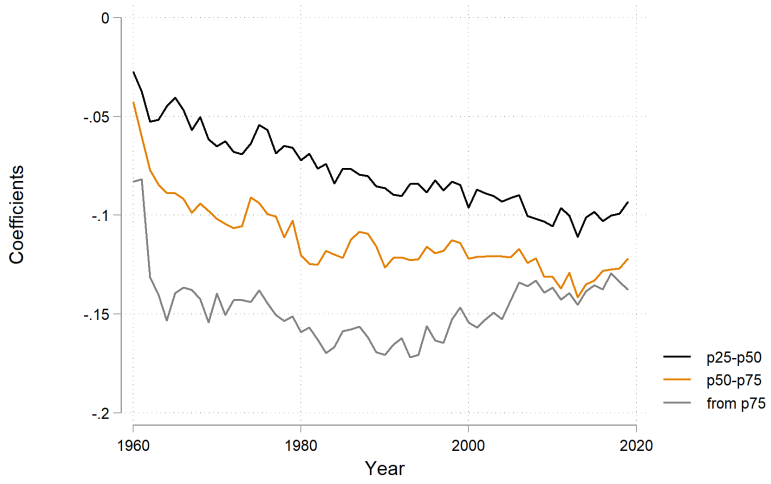

Figure 1.B.2: Distance-Year Interaction Coefficients for Various Distance Percentiles.

After estimating the parameters, I use them to predict the missing trade flows, for country pairs for which I have all the necessary data available. This procedure leads to imputing additional 428,267 missing observations (see Table 1.B.1).

Table 1.B.1: The number of missing observations before and after imputation.

	Missing	Total	Percent Missing
Trade	1,613,663	2,633,400	61.28
Predicted Trade	1,185,396	2,633,400	45.01

The parameters of the model fit are as follows. The adjusted R-squared is 0.62. The 10-fold cross-validation root mean squared error is 2.5 (compared to the mean of 6.64 in the full sample). Figure 1.B.3 plots the actual values of trade against the predicted ones, showing that a large number of observations lie along the 45-degree line.

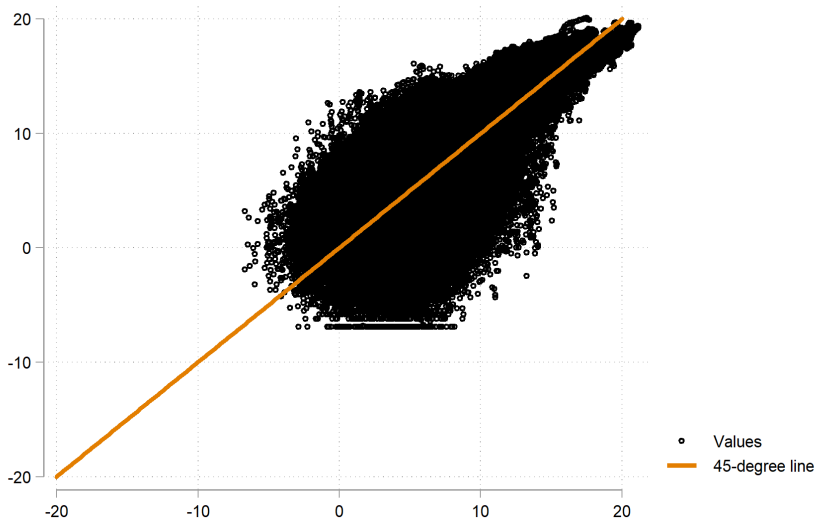

Figure 1.B.3: Actual vs. predicted values of (log) trade.

Importantly, the imputed volumes of trade are never directly used for the blocking procedure or estimation. Instead, I use the values to construct the normalized market shares, which depend not only on trade volumes between two countries, but on the whole matrix of bilateral trade. In this sense, imputation helps me to recover the distribution of normalized market shares. Appendix F implements the whole procedure without imputation, and demonstrates that the conceptual results are unchanged, while the standard errors are larger due to the reduced power.

1.C Domestic Trade

To calculate the normalized market shares in the way consistent with the theoretical framework, I need to take into account the domestic trade. As Santamaría et al. (2020) show, the (log) normalized market shares are (log) deviations between the data and the predictions of the naïve gravity model:

$$\ln s_{ij} = \ln\left(\frac{V_{ij}}{E}\right) - \ln\left(\frac{Y_i}{E}\frac{E_j}{E}\right)$$

where V_{ij} are the sales from origin i to destination j; $E_j = \sum_i V_{ij}$ is the total expenditure of j on all goods, including those coming from j itself; $Y_i = \sum_j V_{ij}$ is the total income of i, including from selling goods to i itself; and $E = \sum_j E_j$ is the total expenditure on all goods, including those sold within the country.

However, the data on production or domestic trade (which is calculated as production minus exports across all destinations) exists only for a very limited number of countries before 1980. To overcome this issue, I collect all available data on domestic trade after 1980 (see Appendix A), construct normalized market shares with and without domestic trade, and compare the two.

Figure 1.C.1 plots the distributions of the normalized market shares with and without the domestic trade. Clearly, the differences in the two measures are very small. Similarly, Figure 1.C.2 shows that the two variables plotted against each other are concentrated along the 45-degree line.

Finally, I run two regressions (with and without covariates) of normalized market shares calculated with domestic trade, s_{ijt}, on normalized market shares calculated without domestic trade, $\tilde{s}_{ijt}$. The results are presented in Table 1.C.1. The coefficient of the univariate regression is 0.99 with the intercept of -0.006, indicating that there is high level of correlation between the two measures. The coefficient of the regression with covariates is slightly smaller, 0.97, but leads to the same conclusion: the normalized market shares calculated with and without domestic trade are highly correlated. I thus proceed to calculate normalized market shares using only international trade data for all years before 1980.

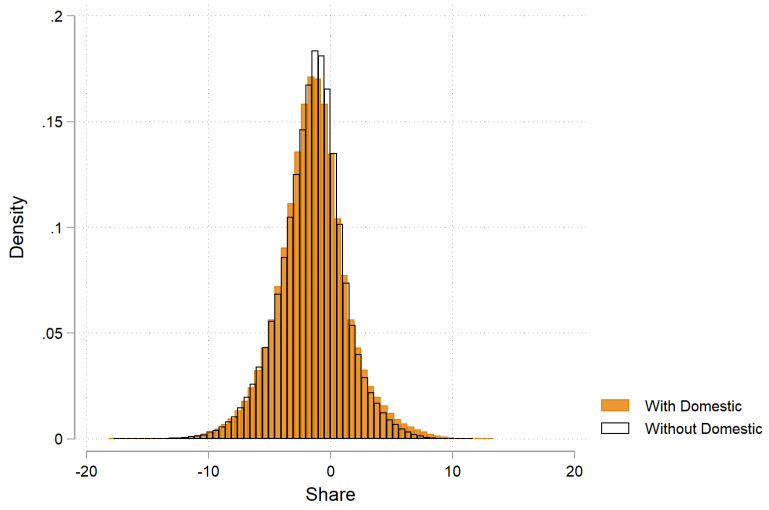

Figure 1.C.1: The distributions of normalized market shares calculated with and without domestic trade after 1980.

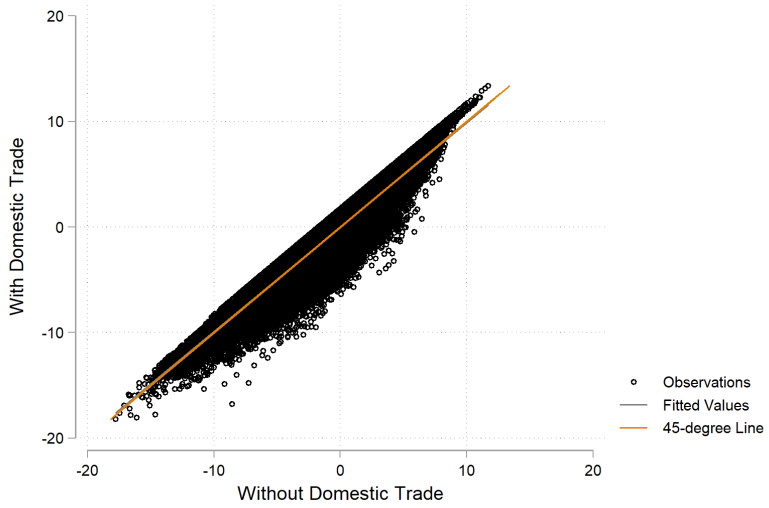

Figure 1.C.2: Normalized market shares without domestic trade against the normalized market shares with domestic trade after 1980.

Table 1.C.1: The coefficients of regressions of normalized market shares with domestic trade on normalized market shares without domestic trade.

	Univariate	Multivariate
$\tilde{s}_{ij}$	0.99***	0.97***
PTA		-0.01*
ln(GDP origin)		-0.02***
ln(GDP destination)		-0.08***
ln(Pop origin)		-0.11***
ln(Pop destination)		-0.6***
ln(Dist)		-0.06***
ln(Area origin)		-0.04***
ln(Area destnation)		-0.02***
Landlock origin		0.25***
Landlock destination		0.16***
Same country		0.08***
Colony		0.04***
Common language		-0.01**
Contiguity		0.05***
Intercept	-0.006***	3.93***
Number of obs.	636,957	549,031
Adj. R-squared	0.95	0.82

Note: Levels of statistical significance correspond to: * p<0.05, ** p<0.01, *** p<0.001.

1.D Various Tables and Figures

Table 1.D.1: Characteristics of PTAs in the final dataset

Indicator		Number of observations	Percentage
Type	FTA	4,065	57.08
	CU	3,057	342.92
Participation	Base Treaty	6,291	88.58
	Accession	811	11.42
Notification	Notified	3,427	48.42
	Not Notified	3,651	51.58
National Treatment	Yes	4,820	67.75
	No	2,294	32.25
Composition	Bilateral	262	3.68
	Plurilateral	3,220	45.21
	Plurilateral and 3rd country	1,192	16.74
	Region-Region	1,637	22.99
	Accession to a PTA	566	7.95
	Inheritance accession	245	3.44
Region	Africa	2,740	38.47
	Americas	382	5.36
	Asia	250	3.51
	Europe	778	10.92
	Oceania	114	1.60
	Intercontinental	2,858	40.13

Table 1.D.2: The standardized differences and t-test for covariate distributions before and after trimming.

	Before Trimming					After Trimming				
	Mean PTA=0	**Mean PTA=1**	**Diff.**	**(Std.Err.)**	**Std. Diff.**	**Mean PTA=0**	**Mean PTA=1**	**Diff.**	**(Std.Err.)**	**Std. Diff.**
Pre-treatment Share	-1.55	-0.40	-1.15***	(0.03)	-0.72	-0.99	-0.53	-0.46***	(0.04)	-0.31
Distance	9.04	7.91	1.13***	(0.01)	1.62	8.42	7.98	0.43***	(0.013)	0.83
Remoteness	9.08	8.96	0.12***	(0.002)	1.05	8.97	8.94	0.03***	(0.002)	0.27
Small Island	0.43	0.19	0.24***	(0.008)	0.53	0.12	0.10	0.02***	(0.008)	0.07
Common Language	0.19	0.3	-0.11***	(0.006)	-0.26	0.21	0.29	-0.08***	(0.01)	-0.19
EU Membership	0.04	0.13	-0.09***	(0.003)	-0.35	0.12	0.16	-0.04***	(0.008)	-0.11
Landlocked	0.25	0.39	-0.14***	(0.006)	-0.29	0.37	0.39	-0.02	(0.012)	-0.04
Common Colonizer	0.14	0.18	-0.04***	(0.006)	-0.11	0.11	0.16	-0.05***	(0.008)	-0.14
Colonial Relationship	0.007	0.014	-0.007***	(0.001)	-0.07	0.02	0.02	0.002	(.003)	0.01
GATT Membership	0.48	0.667	-0.19***	(0.008)	-0.39	0.78	0.77	.0012	(0.01)	0.03
Legal System	0.28	0.47	-0.19***	(0.007)	-0.39	0.38	0.45	-0.07***	(0.01)	-0.15
Pre-treatment PTAs	0.50	1.05	-0.55***	(0.014)	-0.53	1.06	1.28	-0.22***	(0.03)	-0.19
N treated			3,200					2,612		
N control			13,392					4,673		
N Total			16,592					7,285		

Note: Standard errors in parenthesis. Levels of statistical significance correspond to: * p<0.05, ** p<0.01, *** p<0.001. The difference is calculated as the mean(no PTA) - mean(PTA). The standardised differences are calculated using the method of Yang and Dalton (2012). An absolute standardized difference of 0.10 or more indicates that covariates are imbalanced between groups (Austin (2009)).

Table 1.D.3: Balancing t-test of covariates by block

	B1	B2	B3	B4	B5	B6	B7	B8	B9
Pre-treatment Share	-0.25*	-0.08	0.07	0.15*	-0.03	-0.01	-0.03	0.22	-0.80**
	(0.14)	(0.11)	(0.12)	(0.08)	(0.09)	(0.11)	(0.13)	(0.17)	(0.33)
Distance	-0.0003	0.09***	0.04	0.09***	0.02	0.009	-0.23***	-0.11**	-0.53***
	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)	(0.04)	(0.04)	(0.08)
Remoteness	0.001	-0.012*	-0.02*	-0.005	0.0009	0.0009	0.03**	0.02*	0.05***
	(0.007)	(0.006)	(0.007)	(0.005)	(0.006)	(0.006)	(0.009)	(0.009)	(0.01)
Small Island	-0.03	-0.03	0.01	-0.03*	0.005	-0.0008	0.11***	-0.03	0.21***
	(0.03)	(0.03)	(0.02)	(0.18)	(0.02)	(0.02)	(0.03)	(0.02)	(0.04)
Common Language	-0.17***	0.03	0.06	0.11***	0.13***	0.16***	0.17***	-0.17**	-0.21**
	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)	(0.04)	(0.04)	(0.06)	(0.11)
EU Membership	-0.02	0.01	-0.07**	0.03*	0.05*	-0.02	-0.04	-0.01	-0.06
	(0.03)	(0.02)	(0.02)	(0.01)	(0.03)	(0.03)	(0.03)	(0.04)	(0.09)
Landlocked	0.09*	0.009	-0.009	-0.13***	0.06*	-0.04	-0.19***	-0.17***	-0.37***
	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.05)	(0.06)	(0.10)
Common Colonizer	0.02	0.04	0.03	0.02	0.01	-0.01	-0.12***	-0.03	-0.3***
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)	(0.04)	(0.09)
Colonial Relationship	0.02	0.004	-0.004	0.01	-0.03***	0.001	0.002	0.02	0.07***
	(0.01)	(0.009)	(0.01)	(0.007)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
GATT Membership	0.07*	0.02	0.06*	0.07**	0.02	-0.04	-0.14***	-0.27***	-0.22**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.04)	(0.05)	(0.08)
Legal System	-0.03	0.07*	0.07*	-0.005	-0.04	0.01	-0.15***	0.06	-0.05
	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)	(0.05)	(0.06)	(0.11)
Pre-treatment PTAs	0.08**	0.1***	0.07**	0.03	0.02	-0.09***	-0.13***	-0.16***	-0.35***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.05)	(0.06)	(0.1)

Note: Standard errors in parenthesis. Levels of statistical significance correspond to: * p<0.05, ** p<0.01, *** p<0.001. The difference is calculated as the mean(no PTA) - mean(PTA).

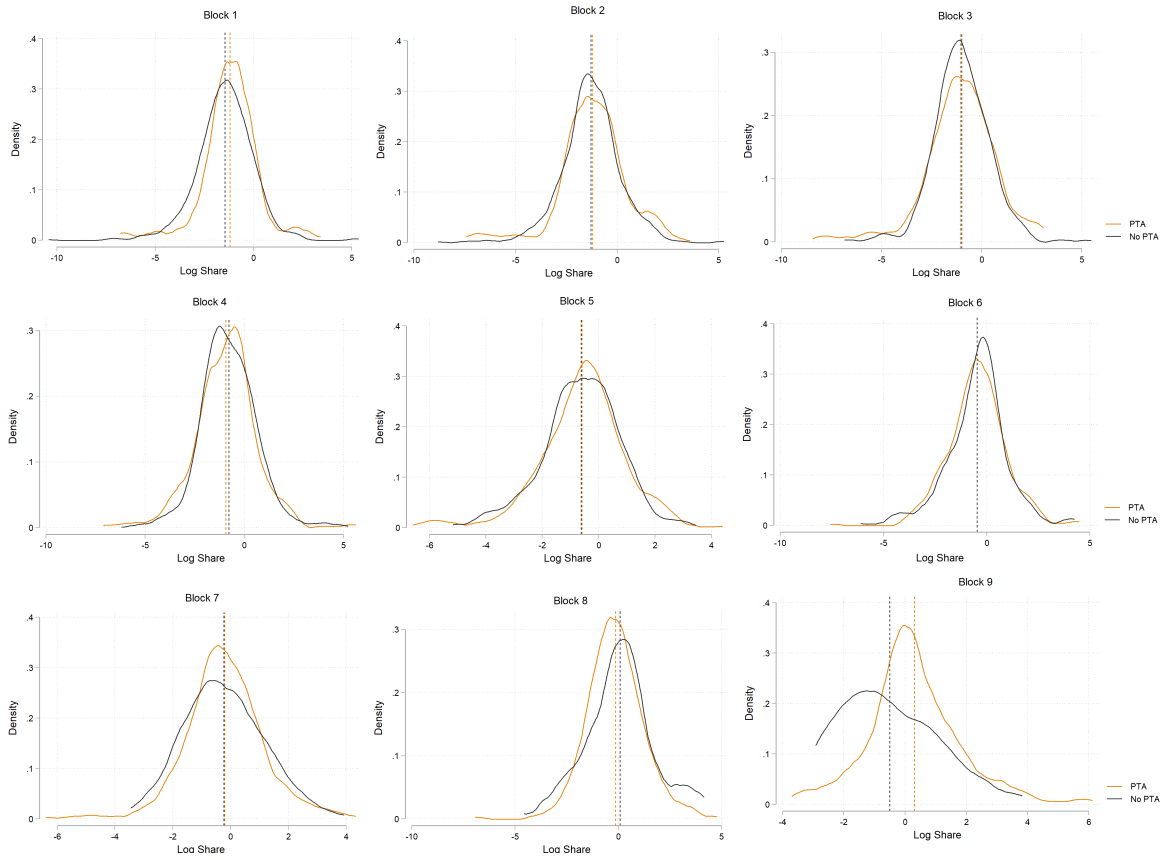

Figure 1.D.1: Distribution of the pre-PTA normalized market shares by treatment group and by block.

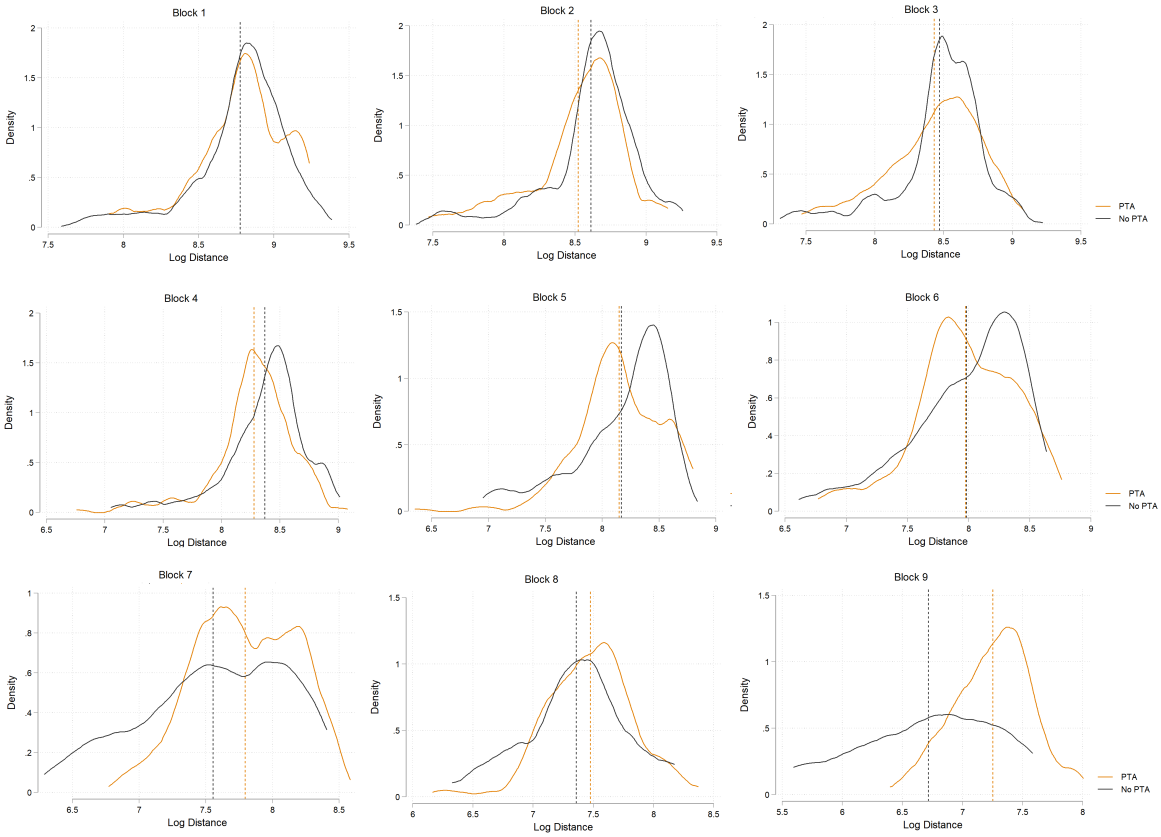

Figure 1.D.2: Distribution of log distance by treatment group and by block.

Table 1.D.4: The probability of having a customs union by block.

	B1	B2	B3	B4	B5	B6	B7	B8	B9
Pre-treatment Share	0.0557	-0.0160	-0.0126	-0.0527	-0.0913	-0.0800	0.0603	-0.167	0.0831
	(0.58)	(-0.23)	(-0.17)	(-1.05)	(-1.39)	(-1.05)	(0.65)	(-1.87)	(0.67)
Distance	-1.626	-1.499	-3.258**	-2.308***	-0.941	-0.0774	-0.668	-0.480	-0.887
	(-1.19)	(-1.46)	(-2.67)	(-4.31)	(-0.81)	(-0.06)	(-0.50)	(-0.49)	(-0.49)
Remoteness	-12.28**	-7.008**	-10.40**	-6.506***	-0.966	-0.229	-3.188	4.668	-2.059
	(-3.23)	(-2.58)	(-3.19)	(-4.07)	(-0.31)	(-0.07)	(-0.89)	(1.62)	(-0.40)
Small Island	-0.239	0.0307	-1.760**	-1.059**	-0.743	0.463	-0.700	-0.404	-2.524*
	(-0.34)	(0.06)	(-2.66)	(-2.97)	(-1.20)	(0.68)	(-0.93)	(-0.67)	(-2.04)
Common Language	1.066*	0.471	0.341	0.200	-0.0120	-0.434	-0.190	0.669	1.412
	(2.04)	(1.09)	(0.70)	(0.83)	(-0.03)	(-1.04)	(-0.36)	(1.60)	(1.88)
EU Membership					-1.887**	-3.147***	-1.533	-2.614***	-0.793
					(-2.92)	(-4.18)	(-1.93)	(-3.57)	(-0.84)
Landlocked	0.589	0.758*	1.455***	0.667**	1.371***	0.798*	1.426***	1.277***	1.371*
	(1.32)	(2.31)	(3.86)	(3.26)	(3.92)	(2.10)	(3.31)	(3.40)	(2.30)
Common Colonizer	0.228	0.104	1.320*	0.762*	0.566	1.131*	1.367**	0.472	2.335**
	(0.35)	(0.21)	(2.37)	(2.48)	(1.27)	(2.26)	(2.64)	(0.94)	(2.85)
Colonial Relationship		0.349					0.160		
		(0.38)					(0.14)		
GATT Membership	-0.173	0.132	0.382	0.268	0.570**	0.974***	1.100***	1.775***	2.541***
	(-0.57)	(0.58)	(1.40)	(1.52)	(2.61)	(3.68)	(3.79)	(5.44)	(5.07)
Legal System	0.0274	-0.299	-0.0394	0.165	0.0619	-0.242	0.817**	1.196***	-0.440
	(0.09)	(-1.29)	(-0.16)	(1.04)	(0.32)	(-1.15)	(3.02)	(3.94)	(-1.11)
Pre-treatment PTAs	-2.839***	-1.199***	-1.155**	-0.493*	-0.648**	0.371	-0.948**	-0.207	-0.981*
	(-4.16)	(-3.41)	(-3.26)	(-2.47)	(-2.83)	(1.48)	(-3.05)	(-0.63)	(-1.98)
Constant	122.2**	73.61*	118.3**	75.65***	14.42	0.656	31.69	-40.93	22.14
	(2.77)	(2.29)	(3.07)	(4.20)	(0.40)	(0.02)	(0.77)	(-1.29)	(0.39)
N Obs	1123	1214	837	1260	929	692	505	441	271
Pseudo R-squared	0.120	0.051	0.075	0.038	0.144	0.169	0.231	0.311	0.401

Note: t-statistics in parenthesis. Levels of statistical significance correspond to: * p<0.05, ** p<0.01, *** p<0.001.

Table 1.D.5: The probability of having a national treatment provision by block.

	B1	B2	B3	B4	B5	B6	B7	B8	B9
Pre-treatment Share	0.0266	-0.00181	-0.0678	-0.0441	-0.0308	-0.0614	0.0881	-0.151	0.0186
	(0.33)	(-0.03)	(-0.98)	(-0.94)	(-0.54)	(-0.87)	(1.04)	(-1.79)	(0.17)
Distance	2.018*	-0.585	0.439	-1.306**	-0.956	0.751	-3.226**	0.882	3.159*
	(2.32)	(-0.65)	(0.51)	(-2.88)	(-0.97)	(0.67)	(-2.64)	(0.99)	(2.03)
Remoteness	2.188	-2.405	0.451	-3.593**	-2.726	2.146	-9.226**	5.192	5.967
	(0.87)	(-1.01)	(0.19)	(-2.60)	(-1.04)	(0.71)	(-2.87)	(1.89)	(1.33)
Small Island	1.089*	0.353	-0.0673	-0.395	-0.938	0.158	-2.524***	0.139	0.579
	(2.16)	(0.72)	(-0.13)	(-1.25)	(-1.73)	(0.25)	(-3.58)	(0.25)	(0.56)
Common Language	0.147	0.0802	-0.490	0.00621	0.0916	-0.644	0.882	-0.573	-1.597*
	(0.40)	(0.20)	(-1.25)	(0.03)	(0.27)	(-1.66)	(1.84)	(-1.49)	(-2.53)
EU Membership	1.428	1.428	1.428	1.428	-1.580**	-2.871***	0.847	-0.220	-2.318**
	(1.53)	(1.53)	(1.53)	(1.53)	(-3.20)	(-4.46)	(1.25)	(-0.41)	(-2.80)
Landlocked	-0.768*	0.370	0.233	0.159	0.893**	-0.0261	1.303***	0.592	-0.194
	(-2.34)	(1.24)	(0.82)	(0.86)	(2.97)	(-0.07)	(3.36)	(1.74)	(-0.39)
Common Colonizer	-1.475**	-0.368	-0.245	0.139	0.359	0.623	2.198***	1.202*	0.613
	(-2.74)	(-0.80)	(-0.55)	(0.48)	(0.91)	(1.34)	(4.50)	(2.57)	(0.93)
Colonial Relationship		0.391					-2.326*		
		(0.44)					(-2.27)		
GATT Membership	-0.652*	0.0889	0.102	0.297	0.264	0.603**	0.884***	1.169***	0.693
	(-2.39)	(0.41)	(0.40)	(1.75)	(1.33)	(2.59)	(3.34)	(4.20)	(1.73)
Legal System	0.0246	-0.249	-0.128	0.0284	-0.0590	-0.641**	0.217	-0.985***	-0.437
	(0.10)	(-1.15)	(-0.57)	(0.19)	(-0.35)	(-3.25)	(0.89)	(-3.74)	(-1.31)
Pre-treatment PTAs	-1.815***	-1.193***	-0.988**	-0.353	-0.0327	0.294	0.00153	0.886**	0.387
	(-3.86)	(-3.53)	(-3.04)	(-1.93)	(-0.17)	(1.23)	(0.01)	(2.90)	(0.94)
Constant	-38.75	24.81	-9.251	41.72**	-25.75	0.656	105.5**	-53.70	-74.93
	(-1.35)	(0.88)	(-0.34)	(2.70)	(-0.73)	(0.02)	(2.84)	(-1.79)	(-1.53)
N Obs	1123	1214	837	1260	929	692	505	441	271
Pseudo R-squared	0.060	0.035	0.026	0.020	0.133	0.169	0.107	0.185	0.164

Note: t-statistics in parenthesis. Levels of statistical significance correspond to: * p<0.05, ** p<0.01, *** p<0.001.

Table 1.D.6: The probability of having a third-party MFN provision by block.

	B1	B2	B3	B4	B5	B6	B7	B8	B9
Pre-treatment Share	0.0543	-0.116	-0.0864	0.0137	-0.0214	0.00621	-0.0282	-0.118	0.0188
	(0.54)	(-1.65)	(-0.96)	(0.25)	(-0.35)	(0.09)	(-0.30)	(-1.37)	(0.16)
Distance	-0.844	1.216	-2.095	-0.767	-0.551	-1.704	-1.026	2.035*	3.855*
	(-0.53)	(0.96)	(-1.09)	(-0.94)	(-0.51)	(-1.41)	(-0.77)	(2.19)	(2.21)
Remoteness	-11.51**	0.404	-6.685	-1.866	-1.597	-4.100	-4.676	7.930**	8.078
	(-2.67)	(0.12)	(-1.33)	(-0.86)	(-0.55)	(-1.27)	(-1.31)	(2.79)	(1.65)
Small Island	0.379	1.131	-2.153*	-0.772	-0.931	-1.724*	-2.461**	-0.879	0.884
	(0.47)	(1.74)	(-2.16)	(-1.70)	(-1.59)	(-2.45)	(-3.03)	(-1.38)	(0.78)
Common Language	0.848	-0.463	0.147	-0.225	-0.0487	0.424	0.0420	-0.262	-0.612
	(1.45)	(-0.94)	(0.21)	(-0.74)	(-0.13)	(1.07)	(0.08)	(-0.65)	(-0.89)
EU Membership	0.597	-0.847	1.799*	-0.514	-0.0392	1.327*	0.505	-1.543**	-3.451***
	(0.53)	(-1.06)	(2.03)	(-1.12)	(-0.08)	(2.18)	(0.68)	(-2.76)	(-3.64)
Landlocked	0.384	-0.127	1.126*	-0.0585	0.836**	1.066**	1.790***	0.749*	0.424
	(0.80)	(-0.33)	(2.12)	(-0.23)	(2.64)	(2.98)	(4.27)	(2.17)	(0.79)
Common Colonizer	-0.105	-0.848	1.066	0.114	0.269	1.461**	1.904***	-0.325	-0.905
	(-0.15)	(-1.51)	(1.46)	(0.32)	(0.64)	(3.01)	(3.56)	(-0.66)	(-1.24)
Colonial Relationship			0.398	-0.539	2.405***	-0.374	0.917		
			(0.42)	(-0.81)	(3.55)	(-0.49)	(0.89)		
GATT Membership	-0.215	-0.0438	0.363	0.0555	0.254	0.928***	0.845**	1.880***	0.819*
	(-0.69)	(-0.20)	(1.24)	(0.32)	(1.24)	(3.68)	(2.92)	(6.03)	(2.00)
Legal System	-0.237	-0.472*	-0.0309	0.0812	0.213	0.0430	0.346	0.315	0.639
	(-0.80)	(-2.13)	(-0.12)	(0.52)	(1.27)	(0.22)	(1.29)	(1.17)	(1.74)
Pre-treatment PTAs	-2.753***	-1.368***	-1.038**	0.0491	-0.431*	0.345	-0.524	-0.133	-0.0000712
	(-4.05)	(-4.06)	(-2.99)	(0.27)	(-2.13)	(1.51)	(-1.73)	(-0.43)	(-0.00)
Constant	108.6*	-15.63	75.29	22.06	17.58	47.89	48.09	-87.45**	-99.56
	(2.14)	(-0.40)	(1.25)	(0.86)	(0.51)	(1.27)	(1.18)	(-2.82)	(-1.85)
N Obs	1123	1214	837	1260	929	692	505	441	271
Pseudo R-squared	0.119	0.044	0.064	0.018	0.056	0.084	0.229	0.225	0.298

Note: t-statistics in parenthesis. Levels of statistical significance correspond to: * p<0.05, ** p<0.01, *** p<0.001.

Table 1.D.7: The probability of notification to the WTO by block.

	B1	B2	B3	B4	B5	B6	B7	B8	B9
Pre-treatment Share	0.194	0.417**	-0.126	0.245**	0.0927	0.0555	-0.222	0.0169	0.182
	(1.41)	(2.67)	(-0.79)	(2.58)	(1.06)	(0.62)	(-1.88)	(0.18)	(1.09)
Distance	3.003**	-0.640	-3.031	-4.937***	-1.293	-4.260**	0.802	-0.949	3.104*
	(2.84)	(-0.30)	(-1.16)	(-3.89)	(-0.86)	(-2.90)	(0.50)	(-0.94)	(1.96)
Remoteness	10.59**	6.950	-3.867	-16.97***	-4.212	-15.43***	-0.239	-8.999**	-38.66***
	(3.07)	(1.23)	(-0.54)	(-4.38)	(-1.07)	(-3.84)	(-0.06)	(-2.60)	(-4.20)
Small Island		-0.963	-1.363	-0.547	-0.831	-1.732*	0.127	-0.862	
		(-0.81)	(-1.11)	(-0.93)	(-0.99)	(-2.08)	(0.13)	(-1.32)	
Common Language	-0.500	-0.816	3.006**	2.453***	0.205	1.445**	-0.142	1.175**	-0.930
	(-0.80)	(-0.76)	(2.88)	(4.69)	(0.38)	(2.93)	(-0.22)	(2.71)	(-1.28)
EU Membership	0.320	2.859*	10.84***	3.330***	0.741	3.133***	1.199	2.214***	0.291
	(0.44)	(2.54)	(4.93)	(4.75)	(1.05)	(4.24)	(1.38)	(3.66)	(0.29)
Landlocked	-2.604**	-1.420	-0.485	-2.521**	-1.423**	0.00120	-1.462**	-0.381	-2.134**
	(-3.26)	(-1.68)	(-0.58)	(-3.17)	(-3.01)	(0.00)	(-2.83)	(-1.02)	(-3.06)
Common Colonizer					-0.994	-0.237	-1.589*	-1.778**	-2.952**
					(-1.48)	(-0.34)	(-2.15)	(-2.69)	(-2.99)
Colonial Relationship		0.338	2.592	-2.298*	1.195	-1.267	-0.0745	-0.239	
		(0.27)	(1.92)	(-2.44)	(1.47)	(-1.45)	(-0.06)	(-0.20)	
GATT Membership	-0.860	-0.473	-1.757**	-2.011***	-0.333	-0.258	-0.332	0.403	-2.275***
	(-1.36)	(-0.84)	(-2.95)	(-5.50)	(-1.07)	(-0.86)	(-0.90)	(1.34)	(-3.83)
Legal System	0.0124	0.415	0.0559	-0.167	0.527*	0.278	-0.666	-0.646*	-0.690
	(0.03)	(0.90)	(0.10)	(-0.56)	(2.36)	(1.12)	(-1.92)	(-2.26)	(-1.32)
Pre-treatment PTAs	1.066	0.306	-8.059***	1.703***	0.862***	1.108***	2.006***	0.602	1.686*
	(1.86)	(0.40)	(-3.70)	(4.45)	(3.34)	(3.93)	(5.50)	(1.90)	(2.32)
Constant	-124.6***	-60.38	56.40	190.2***	46.43	169.1***	-5.034	86.01*	324.3***
	(-3.30)	(-0.89)	(0.66)	(4.30)	(1.00)	(3.65)	(-0.10)	(2.35)	(3.87)
N Obs	1123	1214	837	1260	929	692	505	441	271
Pseudo R-squared	0.221	0.258	0.382	0.289	0.176	0.223	0.328	0.261	0.619

Note: t-statistics in parenthesis. Levels of statistical significance correspond to: * p<0.05, ** p<0.01, *** p<0.001.

Table 1.D.8: The probability of having a late agreement (after 1993) by block.

	B1	B2	B3	B4	B5	B6	B7	B8	B9
Pre-treatment Share	0.103	-0.0874	-0.0505	-0.0410	0.0929	0.0454	0.0861	0.0173	0.201
	(1.26)	(-1.26)	(-0.74)	(-0.82)	(1.58)	(0.70)	(0.99)	(0.22)	(1.75)
Distance	1.207	2.385	-3.221**	-2.995***	-6.487***	-5.091***	-1.052	2.621**	2.298
	(0.95)	(1.94)	(-2.93)	(-5.32)	(-8.46)	(-6.98)	(-0.84)	(3.02)	(1.36)
Remoteness	-0.202	6.454*	-5.869*	-7.179***	-16.82***	-10.44***	-6.419	2.766	-1.317
	(-0.06)	(2.04)	(-2.05)	(-4.47)	(-7.84)	(-4.93)	(-1.94)	(1.06)	(-0.27)
Small Island	0.799	1.509*	-1.875**	-1.184***	-3.373***	-2.860***	-0.885	2.616***	1.517
	(1.21)	(2.38)	(-3.10)	(-3.39)	(-7.59)	(-6.08)	(-1.25)	(4.18)	(1.35)
Common Language	0.203	-0.891	0.820	0.772**	2.070***	1.447***	0.845	-1.196**	-0.243
	(0.43)	(-1.86)	(1.96)	(3.26)	(7.13)	(5.09)	(1.72)	(-3.26)	(-0.38)
EU Membership	0.459	-1.126					-1.185	-3.582***	-1.456
	(0.64)	(-1.55)					(-1.65)	(-6.14)	(-1.67)
Landlocked	-0.645	-0.493	1.045**	0.418*	2.068***	1.303***	1.241**	-0.316	0.0544
	(-1.63)	(-1.33)	(3.03)	(2.09)	(8.18)	(4.98)	(3.10)	(-0.97)	(0.10)
Common Colonizer	-1.116	-1.309*	0.836	0.468	2.102***	2.670***	1.328**	-0.145	1.093
	(-1.77)	(-2.39)	(1.64)	(1.56)	(6.47)	(7.81)	(2.68)	(-0.33)	(1.57)
Colonial Relationship				-1.950*	-2.043*		-1.760		1.134
				(-2.43)	(-2.42)		(-1.69)		(0.75)
GATT Membership	-0.652*	-0.0830	0.0832	-0.0959	0.221	0.320	-0.429	-0.254	-0.516
	(-2.36)	(-0.38)	(0.35)	(-0.59)	(1.14)	(1.46)	(-1.58)	(-0.97)	(-1.21)
Legal System	-0.158	-0.594**	-0.0515	0.123	0.259	-0.227	0.622*	0.102	-0.0743
	(-0.65)	(-2.73)	(-0.23)	(0.82)	(1.49)	(-1.18)	(2.44)	(0.39)	(-0.22)
Pre-treatment PTAs	-1.108**	-1.186***	-1.045**	-0.305	0.0504	0.546*	0.221	0.512	-0.0646
	(-2.75)	(-3.72)	(-3.11)	(-1.61)	(0.25)	(2.36)	(0.77)	(1.76)	(-0.15)
Constant	-10.02	-79.63*	77.94*	87.93***	200.8***	131.6***	64.55	-43.38	-3.944
	(-0.25)	(-2.08)	(2.28)	(4.77)	(8.11)	(5.50)	(1.69)	(-1.51)	(-0.07)
N Obs	1123	1214	837	1260	929	692	505	441	271
Pseudo R-squared	0.041	0.035	0.067	0.054	0.142	0.133	0.167	0.156	0.202

Note: t-statistics in parenthesis. Levels of statistical significance correspond to: * p<0.05, ** p<0.01, *** p<0.001.

Table 1.D.9: The probability of having a plurilateral agreement by block.

	B1	B2	B3	B4	B5	B6	B7	B8	B9
Pre-treatment Share	0.0996	-0.0542	-0.0462	-0.0116	0.0280	-0.130	0.0350	-0.0464	-0.134
	(1.09)	(-0.69)	(-0.62)	(-0.22)	(0.44)	(-1.88)	(0.41)	(-0.54)	(-0.95)
Distance	2.140	2.486	-1.993	-3.085***	-2.553*	-0.425	-2.790*	-1.836	5.526*
	(1.47)	(1.78)	(-1.62)	(-5.15)	(-2.24)	(-0.39)	(-2.24)	(-1.94)	(2.15)
Remoteness	-0.402	4.272	-4.803	-9.390***	-5.651	-2.239	-5.730	-3.000	1.002
	(-0.10)	(1.19)	(-1.50)	(-5.45)	(-1.87)	(-0.76)	(-1.77)	(-1.05)	(0.14)
Small Island	1.067	1.511*	-1.794**	-1.609***	-2.291***	-0.792	-3.074***	-2.848***	-1.755
	(1.41)	(2.11)	(-2.58)	(-4.24)	(-3.72)	(-1.25)	(-4.21)	(-4.44)	(-1.13)
Common Language	-0.492	-1.049	0.596	0.828***	0.705	-0.0186	0.845	1.655***	0.740
	(-0.89)	(-1.91)	(1.27)	(3.33)	(1.78)	(-0.05)	(1.73)	(4.03)	(0.70)
EU Membership	0.226	-0.758			-2.079**	-2.869***	-0.0530	0.580	-3.041*
	(0.29)	(-0.96)			(-3.17)	(-4.33)	(-0.08)	(1.06)	(-2.31)
Landlocked	-0.771	-0.431	0.986**	0.531*	0.857*	0.0216	0.687	1.009**	0.595
	(-1.77)	(-1.04)	(2.61)	(2.54)	(2.57)	(0.06)	(1.76)	(2.89)	(0.71)
Common Colonizer	-0.818	-0.999	0.652	0.698*	0.849*	0.948*	1.785***	1.078*	1.691
	(-1.21)	(-1.69)	(1.20)	(2.25)	(1.97)	(2.06)	(3.57)	(2.19)	(1.29)
Colonial Relationship		1.638		-1.772*	-0.550		-2.466*	-1.393	4.914
		(1.70)		(-2.19)	(-0.58)		(-2.39)	(-1.17)	(1.24)
GATT Membership	-0.969**	-0.477*	-0.0825	-0.173	0.297	0.275	0.999***	1.492***	1.032
	(-3.23)	(-2.09)	(-0.33)	(-1.04)	(1.48)	(1.23)	(3.76)	(5.15)	(1.70)
Legal System	-0.135	-0.417	0.0488	0.252	0.0499	-0.485*	0.381	-0.0199	-0.787
	(-0.51)	(-1.78)	(0.20)	(1.63)	(0.28)	(-2.51)	(1.54)	(-0.07)	(-1.70)
Pre-treatment PTAs	-0.877*	-1.336***	-1.781***	-0.367	0.156	0.435	0.508	1.023**	0.932
	(-2.09)	(-3.62)	(-4.30)	(-1.87)	(0.78)	(1.83)	(1.79)	(3.19)	(1.46)
Constant	-16.31	-60.86	58.09	108.3***	70.01*	22.77	71.29	38.41	-47.18
	(-0.35)	(-1.40)	(1.52)	(5.49)	(1.97)	(0.66)	(1.90)	(1.22)	(-0.61)
N Obs	1123	1214	837	1260	929	692	505	441	271
Pseudo R-squared	0.046	0.034	0.085	0.055	0.132	0.145	0.124	0.181	0.403

Note: t-statistics in parenthesis. Levels of statistical significance correspond to: * p<0.05, ** p<0.01, *** p<0.001.

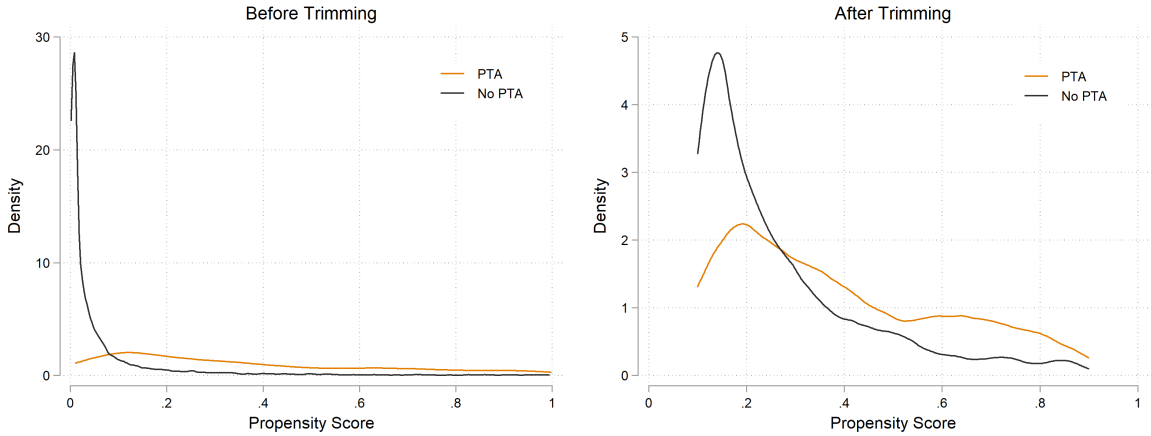

Figure 1.D.3: The estimated probability of concluding a PTA in 1970-2005 for pairs that do not trade in 1960-1965.

The probability is estimated using a logit model, where all covariates are the same as in the baseline estimation, except the logarithm of the pre-treatment normalized market share is substituted by the value of the pre-treatment normalized market share (including zeros). The trimming cutoff is the same as in the baseline exercise.

Table 1.D.10: The number of missing observations in anticipation imputed in the pre-treatment period, by block and treatment status.

		Total obs.	Unique obs.	Total zeros	Unique zeros	Unique imputed
Block 1	PTA=1	115	115	1	1	0
	PTA=0	13,104	1008	334	82	32
Block 2	PTA=1	186	186	0	0	0
	PTA=0	13,364	1028	348	99	38
Block 3	PTA=1	180	180	1	1	0
	PTA=0	8,541	657	242	60	16
Block 4	PTA=1	387	387	3	3	0
	PTA=0	16,587	873	373	56	11
Block 5	PTA=1	405	405	2	2	0
	PTA=0	10,480	524	109	23	7
Block 6	PTA=1	380	380	0	0	0
	PTA=0	7,488	312	107	16	1
Block 7	PTA=1	352	352	0	0	0
	PTA=0	3,366	153	41	8	4
Block 8	PTA=1	360	360	2	2	0
	PTA=0	2,106	81	62	12	0
Block 9	PTA=1	247	247	3	3	0
	PTA=0	480	24	9	3	0

Note: in every block control units are re-sampled for every year of the treatment. Thus, the number of unique control units is smaller than the total number of units. The unique number of zeros counts country pairs which had a zero average normalized market share in any period for which they were re-sampled. The number of unique imputed values represents the number of country pairs which had a zero in anticipation, and had a missing value imputed in the pre-treatment period.

Table 1.D.11: The number of missing observations in the short run imputed in the pre-treatment period, by block and treatment status .

		Total obs.	Unique obs.	Total zeros	Unique zeros	Unique imputed
Block 1	PTA=1	115	115	8	8	7
	PTA=0	13,104	1008	386	77	30
Block 2	PTA=1	186	186	5	5	3
	PTA=0	13,364	1028	440	86	32
Block 3	PTA=1	180	180	5	5	4
	PTA=0	8,541	657	352	56	14
Block 4	PTA=1	387	387	12	12	7
	PTA=0	16,587	873	475	51	9
Block 5	PTA=1	405	405	7	7	5
	PTA=0	10,480	524	153	23	7
Block 6	PTA=1	380	380	5	5	5
	PTA=0	7,488	312	133	14	1
Block 7	PTA=1	352	352	2	2	2
	PTA=0	3,366	153	37	8	4
Block 8	PTA=1	360	360	1	1	1
	PTA=0	2,106	81	87	12	0
Block 9	PTA=1	247	247	0	0	0
	PTA=0	480	24	17	3	0

Note: in every block control units are re-sampled for every year of the treatment. Thus, the number of unique control units is smaller than the total number of units. The unique number of zeros counts country pairs which had a zero average normalized market share in any period for which they were re-sampled. The number of unique imputed values represents the number of country pairs which had a zero in the short run, and had a missing value imputed in the pre-treatment period.

Table 1.D.12: The number of missing observations in the medium run imputed in the pre-treatment period, by block and treatment status .

		Total obs.	Unique obs.	Total zeros	Unique zeros	Unique imputed
Block 1	PTA=1	115	115	6	6	6
	PTA=0	13,104	1008	376	66	27
Block 2	PTA=1	186	186	6	6	4
	PTA=0	13,364	1028	451	83	36
Block 3	PTA=1	180	180	6	6	6
	PTA=0	8,541	657	366	50	13
Block 4	PTA=1	387	387	10	10	6
	PTA=0	16,587	873	517	55	11
Block 5	PTA=1	405	405	6	6	5
	PTA=0	10,480	524	169	25	8
Block 6	PTA=1	380	380	6	6	6
	PTA=0	7,488	312	135	16	2
Block 7	PTA=1	352	352	4	4	4
	PTA=0	3,366	153	13	6	3
Block 8	PTA=1	360	360	1	1	1
	PTA=0	2,106	81	126	13	2
Block 9	PTA=1	247	247	2	2	0
	PTA=0	480	24	22	2	0

Note: in every block control units are re-sampled for every year of the treatment. Thus, the number of unique control units is smaller than the total number of units. The unique number of zeros counts country pairs which had a zero average normalized market share in any period for which they were re-sampled. The number of unique imputed values represents the number of country pairs which had a zero in the medium run, and had a missing value imputed in the pre-treatment period.

Table 1.D.13: The number of missing observations in the long run imputed in the pre-treatment period, by block and treatment status .

		Total obs.	Unique obs.	Total zeros	Unique zeros	Unique imputed
Block 1	PTA=1	115	115	4	4	4
	PTA=0	13,104	1008	427	73	28
Block 2	PTA=1	186	186	6	6	4
	PTA=0	13,364	1028	506	90	34
Block 3	PTA=1	180	180	7	7	7
	PTA=0	8,541	657	402	52	13
Block 4	PTA=1	387	387	7	7	6
	PTA=0	16,587	873	602	58	14
Block 5	PTA=1	405	405	6	6	6
	PTA=0	10,480	524	214	31	10
Block 6	PTA=1	380	380	4	4	4
	PTA=0	7,488	312	158	15	2
Block 7	PTA=1	352	352	2	2	2
	PTA=0	3,366	153	18	8	6
Block 8	PTA=1	360	360	1	1	1
	PTA=0	2,106	81	151	15	4
Block 9	PTA=1	247	247	0	0	0
	PTA=0	480	24	25	3	0

Note: in every block control units are re-sampled for every year of the treatment. Thus, the number of unique control units is smaller than the total number of units. The unique number of zeros counts country pairs which had a zero average normalized market share in any period for which they were re-sampled. The number of unique imputed values represents the number of country pairs which had a zero in the long run, and had a missing value imputed in the pre-treatment period.

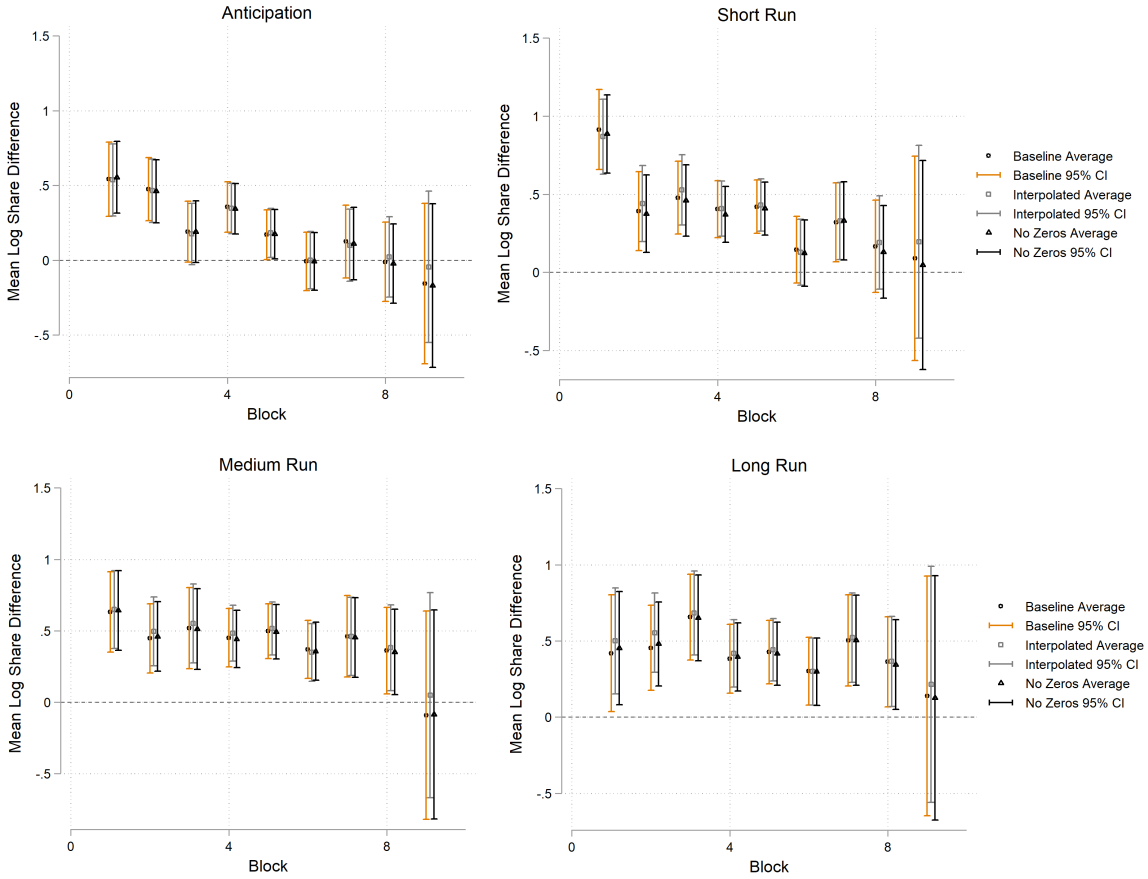

Figure 1.D.4: Comparison of the estimates obtained using (1) the baseline procedure; (2) Interpolated sample; and (3) Sample excluding zeros in average calculations.

Note: baseline procedure conditions the analysis on positive trade flows the the pre-treatment period in the imputed sample. The average normalized market share is calculated assuming zeros for missing values for those pairs that used to trade in the pre-treatment period, but have missing values in later periods. The interpolation estimate implements the same procedure as the baseline estimate, but does so for the imputed and interpolated sample. The estimate obtained without considering zeros is still using data conditional on positive trade flows in the pre-treatment period, but does not assume missing trade flows as zeros in later years. The average normalized market shares are calculated using only the available data, discarding the missing values.

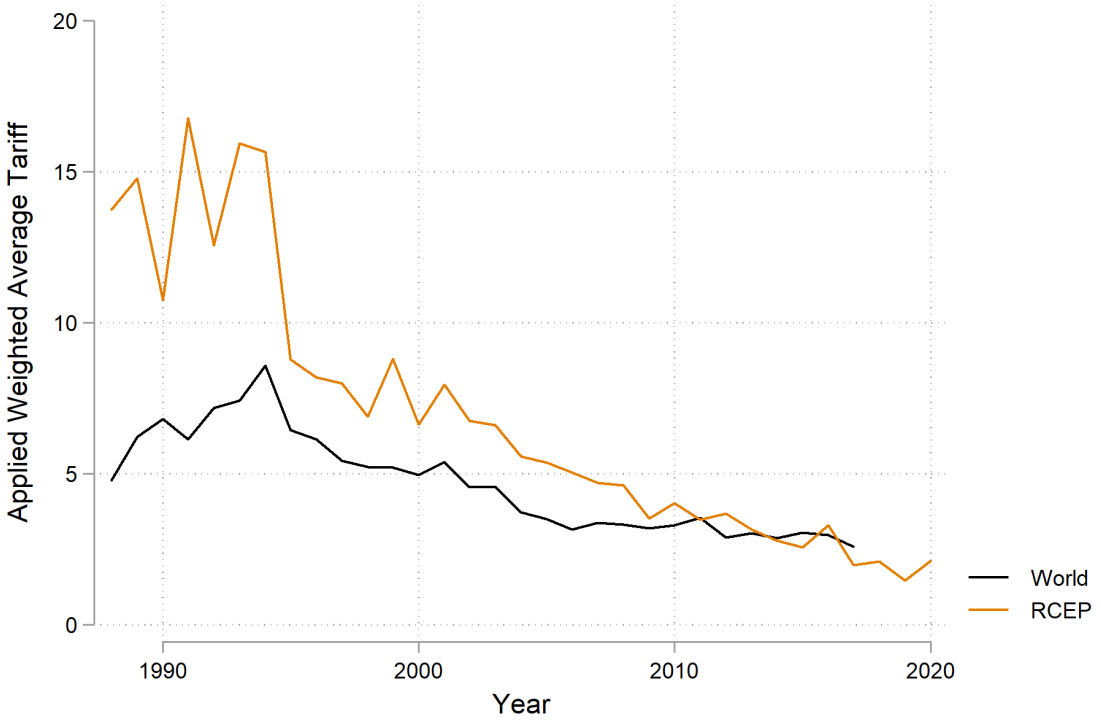

Figure 1.D.5: Applied weighted average tariffs in the world and in RCEP countries, 1988-2020.

Source: World Bank

Table 1.D.14: Applied weighted average tariffs by RCEP country in 1988 and 2020.

Country / Region	1988	2020
Australia	18.56*	0.71
Brunei Darussalam	4.43*	0.02
China	32.17*	2.47
Indonesia	14.54*	2.04
Japan	4.12	2.22
Cambodia	16.43*	6.21
Korea, Rep.	13.95	5.48
Lao PDR	14.06*	0.97
Myanmar	4.13*	1.81*
Malaysia	14.4	3.6
New Zealand	11.24*	0.85
Philippines	22.5	1.67
Singapore	3.26*	0.05
Thailand	33.65*	3.52*
Vietnam	15.19*	1.34
RCEP	14.84	2.20
World	4.79	2.59*

Note: values indicated with stars are not available for the corresponding year, and are presented for the nearest available year. In particular, Myanmar in 2020 is in Myanmar 2019; Thailand in 2020 is Thailand in 2015; World in 2020 is World in 2017; Australia in 1988 is Australia in 1991; Brunei in 1988 is Brunei 1992; China in 1988 is China in 1992; Indonesia in 1988 is Indonesia in 1989; Cambodia in 1988 is Cambodia in 2001; Laos in 1988 is Laos in 2000; Myanmar in 1988 is Myanmar in 1996; New Zealand in 1988 is New Zealand in 1992; Singapore in 1988 is Singapore in 1989; Thailand in 1988 is Thailand in 1989; Vietnam in 1988 is Vietnam in 1994.
Source: World Bank

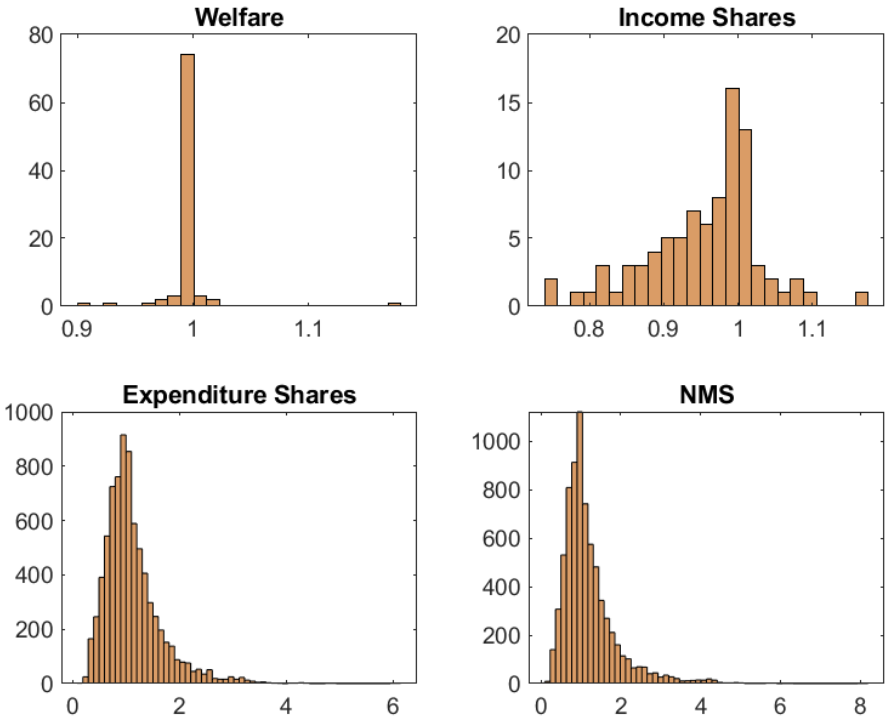

Figure 1.D.6: Distributions of gross growth rates of welfare (real consumption), income shares, expenditure shares, and normalized market shares (NMS) after RCEP formation, in the long run, for all countries.

Note: Top left panel plots the distribution of the gross growth rates of real consumption / welfare ($\hat{C}_j$). Top right panel plots the distribution of the gross growth rates of income shares ($\hat{E}_j$). Bottom left panel plots the distribution of the gross growth rates of expenditure shares ($\hat{\lambda}_{ij}$). The bottom right panel plots the distribution of the gross growth rates of normalized market shares ($\hat{s}_{ij}$).

Table 1.D.15: Decomposition of welfare changes into size and price effects: ross growth rates of size ($\hat{E}$), price ($\hat{P}$) and welfare ($\hat{C}$).

Country name	Size	Price	Welfare	Country name	Size	Price	Welfare
Afghanistan	0.7919	0.7991	0.9910	Jamaica	0.9360	0.9589	0.9762
Angola	1.0548	1.0546	1.0002	Jordan	0.8901	0.8902	0.9999
Albania	0.9237	0.9240	0.9997	Japan	1.0014	1.0014	1.0000
Andorra	0.9697	0.9830	0.9864	Kenya	0.8560	0.8561	0.9999
United Arab Emirates	0.9737	0.9737	1.0000	Cambodia	0.9984	0.9790	1.0198
Argentina	0.9910	0.9917	0.9993	South Korea	1.0268	1.0267	1.0001
Australia	1.0107	1.0106	1.0001	Kuwait	1.1414	1.1412	1.0001
Austria	0.9973	0.9973	1.0000	Lebanon	0.8531	0.9097	0.9378
Burkina Faso	0.9714	0.9791	0.9921	Sri Lanka	0.9243	0.9243	1.0000
Bulgaria	0.9818	0.9818	1.0000	Lesotho	0.9631	0.9841	0.9787
Bahrain	0.9270	0.9270	0.9999	Luxembourg	0.9898	0.9904	0.9995
Bermuda	0.8831	0.8851	0.9977	Macao	0.7885	0.7900	0.9981
Brazil	1.0013	1.0013	1.0000	Morocco	0.9420	0.9420	0.9999
Botswana	0.8816	0.8818	0.9998	Maldives	0.8232	0.8232	1.0000
Canada	0.9535	0.9535	1.0000	Mexico	0.9587	0.9588	1.0000
Switzerland	0.9866	0.9866	1.0000	Myanmar	1.0901	0.9212	1.1834
Chile	0.9990	0.9989	1.0000	Mongolia	1.0636	1.0634	1.0001
China	1.0236	1.0236	1.0000	Mauritius	0.9100	0.9112	0.9988
Congo	1.0009	0.9903	1.0107	Malaysia	1.0374	1.0372	1.0002
Colombia	0.9472	0.9472	1.0000	Namibia	0.9114	0.9117	0.9997
Costa Rica	0.9380	0.9380	1.0000	Niger	0.9232	0.9235	0.9996
Cyprus	0.9194	0.9196	0.9998	Netherlands	0.9936	0.9935	1.0001
Germany	1.0138	1.0137	1.0000	Norway	1.0272	1.0271	1.0001
Denmark	1.0047	1.0047	1.0000	New Zealand	1.0000	0.9999	1.0001
Algeria	0.9606	0.9607	1.0000	Oman	1.0007	1.0007	1.0001
Ecuador	0.9580	0.9580	1.0000	Panama	0.9145	0.9147	0.9998
Egypt	0.8983	0.8984	0.9999	Peru	0.9752	0.9752	1.0000
Spain	0.9779	0.9779	1.0000	Philippines	1.0045	1.0044	1.0001
Ethiopia	0.8450	0.8452	0.9997	Poland	0.9970	0.9970	1.0000
Finland	1.0092	1.0092	1.0000	Portugal	0.9739	0.9739	1.0000
Fiji	0.9003	0.9006	0.9998	Fr. Polynesia	0.8648	0.9427	0.9173
France	0.9871	0.9871	1.0000	Qatar	1.0652	1.0650	1.0002
United Kingdom	0.9581	0.9581	1.0000	Saudi Arabia	1.0002	1.0002	1.0000
Greece	0.9400	0.9401	0.9999	El Salvador	0.9049	0.9339	0.9690
Greenland	1.0079	1.0043	1.0035	Sweden	1.0034	1.0034	1.0000
Hong Kong	0.8279	0.8356	0.9908	Thailand	1.0115	1.0053	1.0062
Hungary	1.0061	1.0061	1.0000	Tunisia	0.9625	0.9626	1.0000
Indonesia	1.0144	1.0144	1.0001	Turkey	0.9571	0.9571	1.0000
India	0.9473	0.9473	1.0000	Tanzania	0.8794	0.8797	0.9997
Ireland	1.0315	1.0314	1.0001	Uruguay	0.9986	0.9970	1.0016
Iran	1.0006	1.0006	1.0000	United States of America	0.9426	0.9426	1.0000
Iceland	0.9863	0.9863	1.0000	Vietnam	0.9994	0.9992	1.0001
Israel	0.9834	0.9834	1.0000	South Africa	0.9905	0.9905	1.0000
Italy	1.0034	1.0034	1.0000	Zimbabwe	0.9521	0.9521	0.9999

Note: Welfare is defined as the change in real consumption, $C_j = E_j/P_j$, where E_j is the total expenditure, and P_j is the price index. This table decomposes the changes in welfare into changes in size ($\hat{E}_j$) and changes in the price index ($\hat{P}_j$).

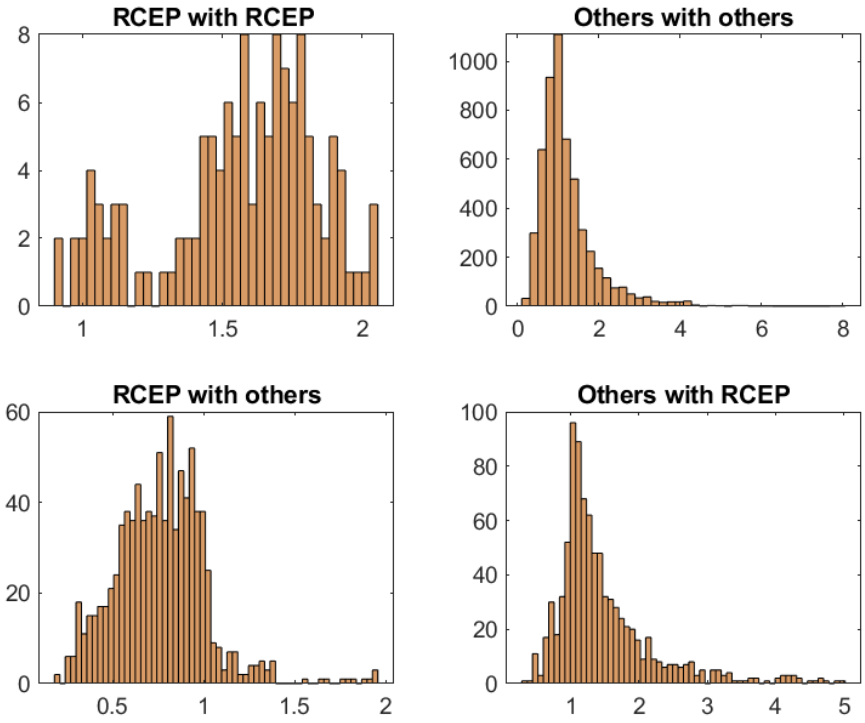

Figure 1.D.7: Distributions of gross growth rates of normalized market shares of different country groups, in the long run.

Note: Top left panel plots the distribution of the gross growth rates of normalized markets shares of RCEP countries with other members of RCEP (excluding domestic trade). Top right panel plots the distribution of the gross growth rates of normalized markets shares of countries outside of RCEP among each other (including domestic trade). Bottom left panel plots the distribution of the gross growth rates of normalized markets shares of RCEP members (as exporters) with countries outside of RCEP (as importers). Bottom right panel plots the distribution of the gross growth rates of normalized markets shares of countries outside of RCEP (as exporters) with RCEP members (as importers).

Table 1.D.16: Block coefficients and corresponding percentage iceberg trade cost reductions use in the counterfactual general equilibrium exercise.

Block	Number of RCEP pairs	Anticipation coefficient	Anticipation iceberg trade cost reduction	Long run coefficient	Long run iceberg trade cost reduction
1	1	0.54	10.83	0.63	12.63
2	16	0.39	7.98	0.46	9.11
3	15	0.19	3.81	0.52	10.41
4	14	0.36	7.15	0.44	8.97
5	15	0	0	0.50	10.00
6	3	0	0	0.37	7.43
7	6	0	0	0.50	10.08
8	30	0	0	0.37	7.37
9	32	0	0	0.15	3.05

Note: The coefficients correspond to regression adjustment coefficients for each block, resulting from a blocking procedure applied to year 2015, following the methodology outlined in the empirical section of the paper. Zero coefficients correspond to block point estimates that were not statistically significant. The corresponding iceberg trade cost reductions were calculated using the trade elasticity of $\varepsilon = 5$.

Table 1.D.17: Average percentage change of normalized market shares in RCEP members' mutual trade, in anticipation and long run.

Block	Anticipation	Long run
1	38.06	-24.55
2	47.85	0.57
3	20.59	36.39
4	38.99	14.29
5	0.43	53.16
6	1.92	51.17
7	-1.04	47.78
8	-2.79	40.01
9	-2.15	13.69

Note: The counterfactual exercise is carried out using block coefficients and corresponding iceberg trade cost shocks presented in Table 1.D.16.

1.E Standard Errors

As described in the body of the paper, at the analysis stage the structure of the data is such that the same control country pairs appear multiple times for different time periods within each block. This appendix deals with the consequences of such setup for the estimation of the means and the sampling variances within each block. In order to relax the assumption that the control units are independent observations, I run two simulation exercises: bootstrap and re-sampling from control distribution. Both methods demonstrate that the point estimates of the $\hat{\tau}$ from Equation 1.6 are very close to the mean of the simulated distribution; while standard errors are systematically higher in the simulations.

Bootstrap

The first method is a standard bootstrap procedure. For each $T = \{A, S, M, L\}$ and for each block, I re-sample observations with replacement, run the regression using Equation 1.6, calculate the mean and the standard error at each iteration; perform this procedure one thousand times. This will give me a whole distribution of block means and standard errors. Since I do this for each time period (pre-treatment, anticipation, short, medium, and long run) and each of the nine blocks, there are a total of 45 distributions. In the interest of space, I will report the means of the simulated point estimates and standard errors distributions along with the their counterparts without re-sampling; and provide a visualisation of the typical distribution.

Table 1.E.1 reports the results for $\hat{\tau}$'s and the means for their simulated distributions obtained using bootstrap. With the exception of the pre-treatment period, all the point estimates of the mean are almost exactly the same as the means of the simulated distributions. The slightly higher differences between the two estimates for the pre-treatment period, however, do not change the conceptual results, as the point estimates are still not statistically significant, given the standard errors. Figure 1.E.1 shows simulated distribution and the point estimate for the anticipation period for the nine blocks, visually re-enforcing the reported results in the table. This is a typical picture for all other periods as well.

Similarly, Table 1.E.2 reports the means of the simulated distributions for the standard errors, as well as standard errors obtained using the data without re-sampling. The main conclusion is that the bootstrapped standard errors are systematically higher

than their counterparts in the full sample. Figure 1.E.2 confirms this conclusion visually.

Table 1.E.1: The point estimates and the means of the simulated distributions using bootstrap, by block and time period

	Pre-Treatment		Anticipation		Short Run		Medium Run		Long Run	
	Point Estimate	Distribution Mean	Point Estimate	Distribution Mean	Point Estimate	Distribution Mean	Point Estimate	Distribution Mean	Point Estimate	Distribution Mean
B1	0.296	0.258	0.543	0.542	0.914	0.914	0.633	0.632	0.419	0.421
B2	-0.002	-0.065	0.475	0.473	0.393	0.399	0.448	0.455	0.455	0.444
B3	-0.144	-0.127	0.191	0.190	0.478	0.481	0.519	0.520	0.657	0.648
B4	-0.095	-0.173	0.356	0.358	0.405	0.403	0.451	0.449	0.383	0.380
B5	0.025	0.110	0.171	0.173	0.420	0.416	0.498	0.500	0.428	0.433
B6	0.066	0.238	-0.007	-0.007	0.146	0.146	0.370	0.372	0.301	0.304
B7	0.065	0.210	0.125	0.122	0.321	0.327	0.461	0.455	0.504	0.504
B8	-0.042	0.236	-0.011	-0.014	0.167	0.168	0.361	0.360	0.364	0.369
B9	1.080	1.226	-0.156	-0.161	0.089	0.104	-0.091	-0.073	0.140	0.153

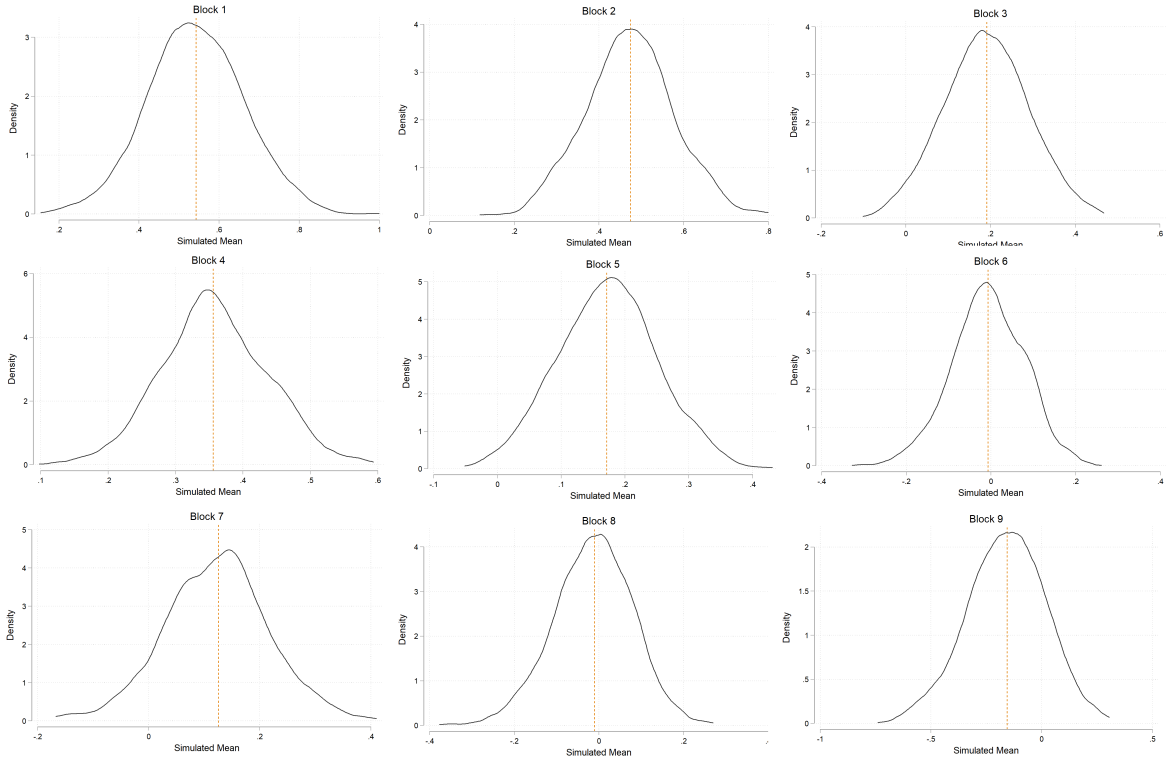

Figure 1.E.1: The bootstrap-simulated distributions and the point estimates of $\hat{\tau}$ for anticipation period, by block.

Table 1.E.2: The standard errors and the means of the simulated distributions of the standard errors using bootstrap, by block and time period

	Pre-Treatment		Anticipation		Short Run		Medium Run		Long Run	
	Full Sample	Distribution Mean	Full Sample	Distribution Mean	Full Sample	Distribution Mean	Full Sample	Distribution Mean	Full Sample	Distribution Mean
B1	0.128	0.201	0.127	0.173	0.131	0.179	0.143	0.195	0.196	0.268
B2	0.124	0.173	0.108	0.148	0.129	0.178	0.124	0.170	0.143	0.197
B3	0.135	0.209	0.104	0.142	0.119	0.162	0.145	0.199	0.144	0.197
B4	0.093	0.128	0.086	0.117	0.093	0.125	0.104	0.141	0.115	0.156
B5	0.090	0.132	0.085	0.112	0.087	0.114	0.098	0.129	0.106	0.139
B6	0.105	0.156	0.100	0.129	0.109	0.141	0.104	0.133	0.113	0.147
B7	0.131	0.155	0.124	0.154	0.129	0.157	0.145	0.176	0.152	0.187
B8	0.189	0.228	0.135	0.163	0.150	0.179	0.154	0.180	0.151	0.176
B9	0.340	0.363	0.273	0.307	0.333	0.373	0.371	0.408	0.401	0.430

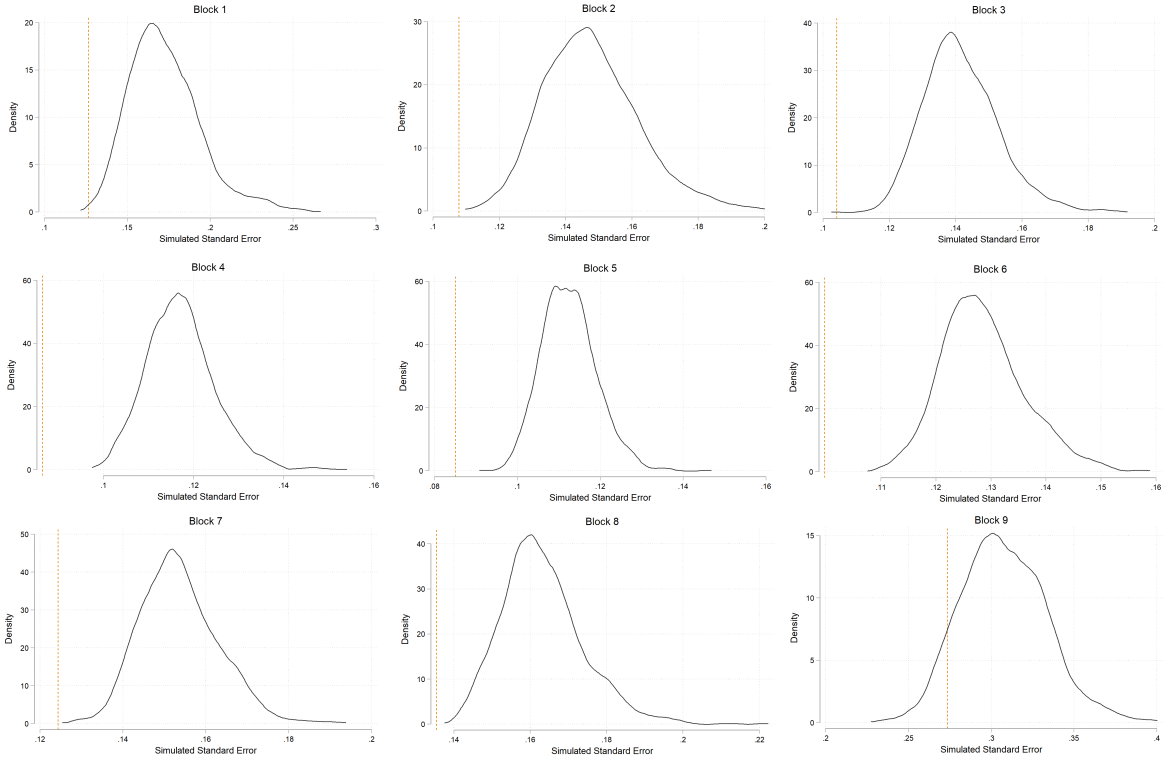

Figure 1.E.2: The bootstrap-simulated distributions of standard errors and the estimates of standard errors using the full sample in anticipation period, by block.

Re-Sampling From Control Distribution

In the bootstrap exercise all observations were randomly re-sampled with replacements. However, given the structure of my data, I know that the non-independent observations are only the ones in the control group. Thus, I perform a simulation which is more tailored to my data structure. In particular, it randomly re-samples observations only from the control group, while leaving the treated observations the same within each sample. The algorithm is similar to the bootstrap: for each $T = \{A, S, M, L\}$ and for each block, I sample the observations, run the regression using Equation 1.6, calculate the mean and the standard error at each iteration; perform this procedure one thousand times. The number of control observations sampled at each iteration is approximately equal to the number of unique control pairs within each block.

Similarly to bootstrap results, Table 1.E.3 shows that there are no big differences between the point estimates of the $\hat{\tau}$'s and the means of the simulated distributions. Differently from the bootstrap, however, the results suggest that the means should be slightly higher for every block and time period. For most blocks, however, the differences are small, as confirmed visually in Figure 1.E.3, which plots the distributions for the nine blocks in the anticipation period as an example.

Similarly, the re-sampling method confirms the results of the bootstrap estimation for the standard errors. Table 1.E.4 compares the standard errors obtained from the full sample estimation and the mean of the simulated distribution of the standard errors. Again, the simulated standard errors are systematically higher than those from the full sample.

Comparison

Finally, Figure 1.E.5 compares the standard errors obtained with three different methods: by estimating the full sample, by performing a bootstrap procedure, and by re-sampling from the control distributions, in different time periods, across all blocks. The conclusion is that the bootstrap standard errors are larger than those obtained by the other two methods. I therefore use these more conservative standard errors in the body of the paper to report the statistical significance of the point estimates.

Table 1.E.3: The point estimates and the means of the simulated distributions using re-sampling from the control group, by block and time period

	Pre-Treatment		Anticipation		Short Run		Medium Run		Long Run	
	Point Estimate	Distribution Mean	Point Estimate	Distribution Mean	Point Estimate	Distribution Mean	Point Estimate	Distribution Mean	Point Estimate	Distribution Mean
B1	0.296	0.318	0.543	0.547	0.914	0.922	0.633	0.681	0.419	0.509
B2	-0.002	0.046	0.475	0.526	0.393	0.454	0.448	0.522	0.455	0.579
B3	-0.144	-0.122	0.191	0.197	0.478	0.508	0.519	0.564	0.657	0.694
B4	-0.095	-0.029	0.356	0.374	0.405	0.435	0.451	0.583	0.383	0.551
B5	0.025	0.132	0.171	0.202	0.420	0.457	0.498	0.553	0.428	0.518
B6	0.066	0.137	-0.007	-0.119	0.146	0.058	0.370	0.311	0.301	0.224
B7	0.065	0.082	0.125	0.163	0.321	0.353	0.461	0.498	0.504	0.566
B8	-0.042	-0.038	-0.011	-0.013	0.167	0.162	0.361	0.361	0.364	0.362
B9	1.080	1.128	-0.156	0.084	0.089	0.270	-0.091	0.004	0.140	0.327

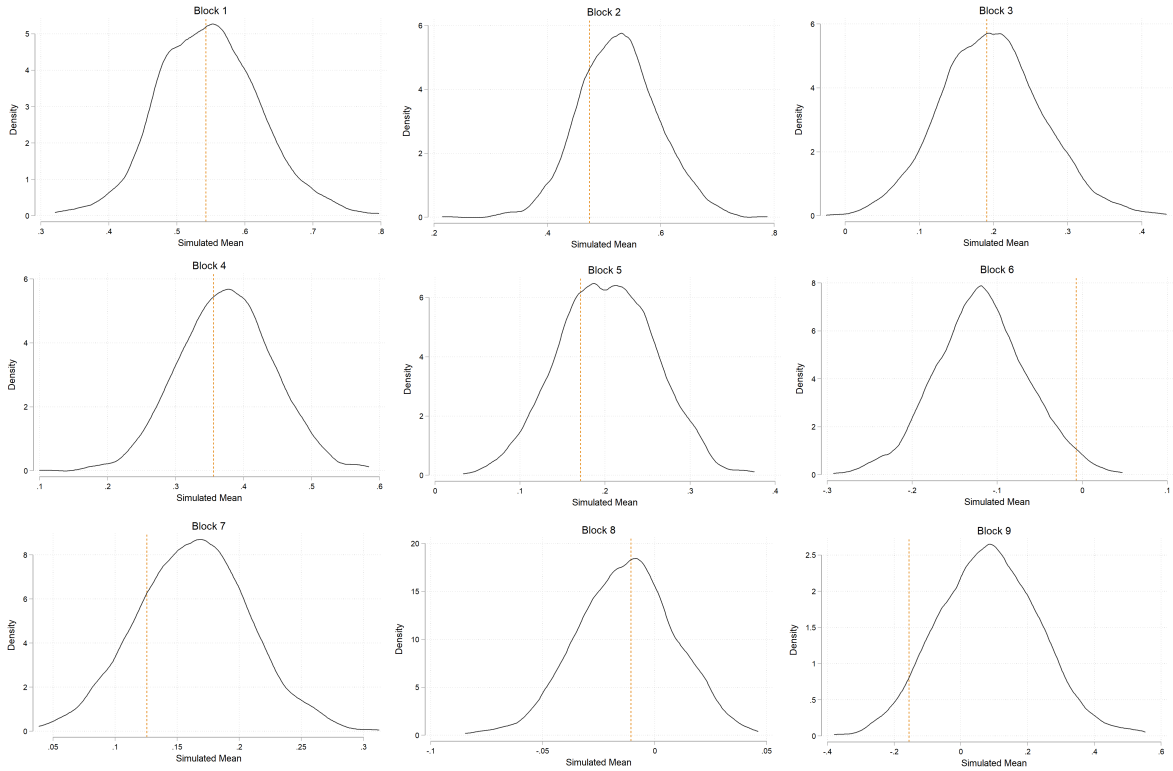

Figure 1.E.3: The simulated distributions using re-sampling from the control group, and the point estimates of $\hat{\tau}$ for anticipation period, by block.

Table 1.E.4: The standard errors and the means of the simulated distributions of the standard errors using re-sampling from the control distribution, by block and time period

	Pre-Treatment		Anticipation		Short Run		Medium Run		Long Run	
	Full Sample	Distribution Mean	Full Sample	Distribution Mean	Full Sample	Distribution Mean	Full Sample	Distribution Mean	Full Sample	Distribution Mean
B1	0.128	0.144	0.127	0.147	0.131	0.152	0.143	0.162	0.196	0.203
B2	0.124	0.143	0.108	0.132	0.129	0.153	0.124	0.149	0.143	0.160
B3	0.135	0.150	0.104	0.125	0.119	0.140	0.145	0.161	0.144	0.168
B4	0.093	0.122	0.086	0.119	0.093	0.128	0.104	0.135	0.115	0.148
B5	0.090	0.106	0.085	0.106	0.087	0.108	0.098	0.119	0.106	0.126
B6	0.105	0.121	0.100	0.117	0.109	0.126	0.104	0.122	0.113	0.129
B7	0.131	0.135	0.124	0.131	0.129	0.135	0.145	0.150	0.152	0.154
B8	0.189	0.190	0.135	0.138	0.150	0.152	0.154	0.156	0.151	0.152
B9	0.340	0.352	0.273	0.309	0.333	0.363	0.371	0.393	0.401	0.403

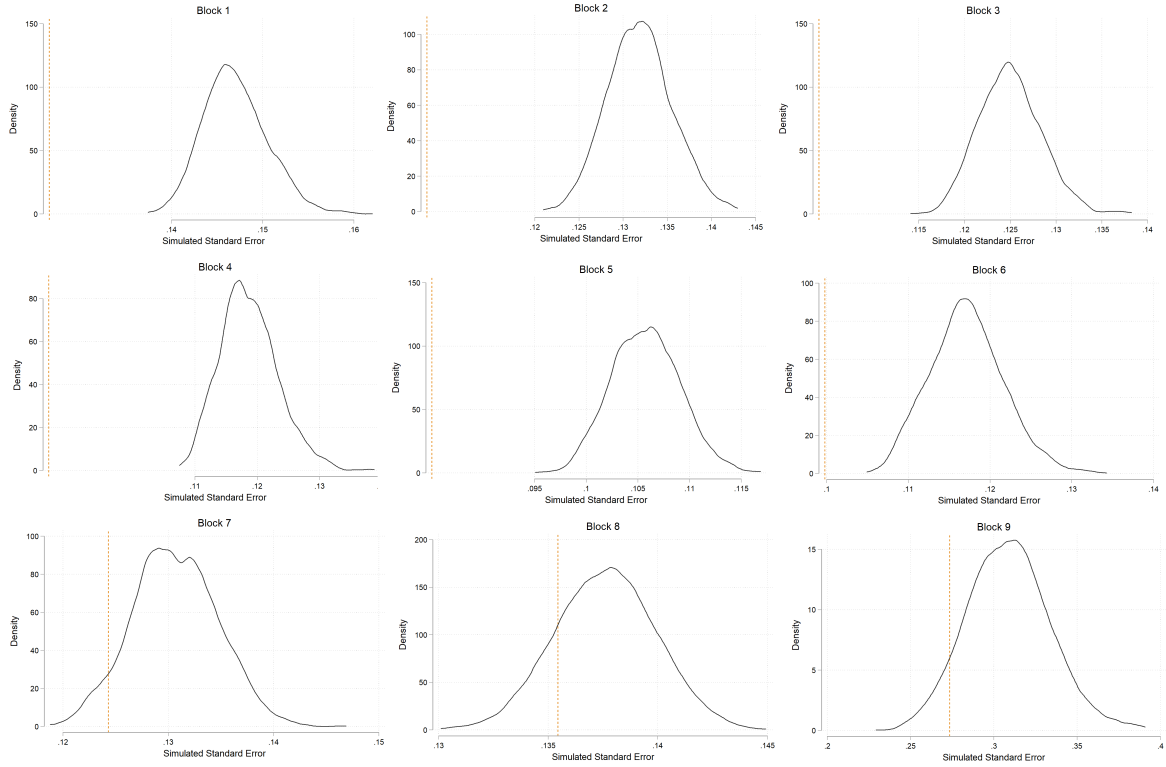

Figure 1.E.4: The simulated distributions of standard errors using re-sampling from the control group, and the estimates of standard errors in the full sample, for anticipation period, by block.

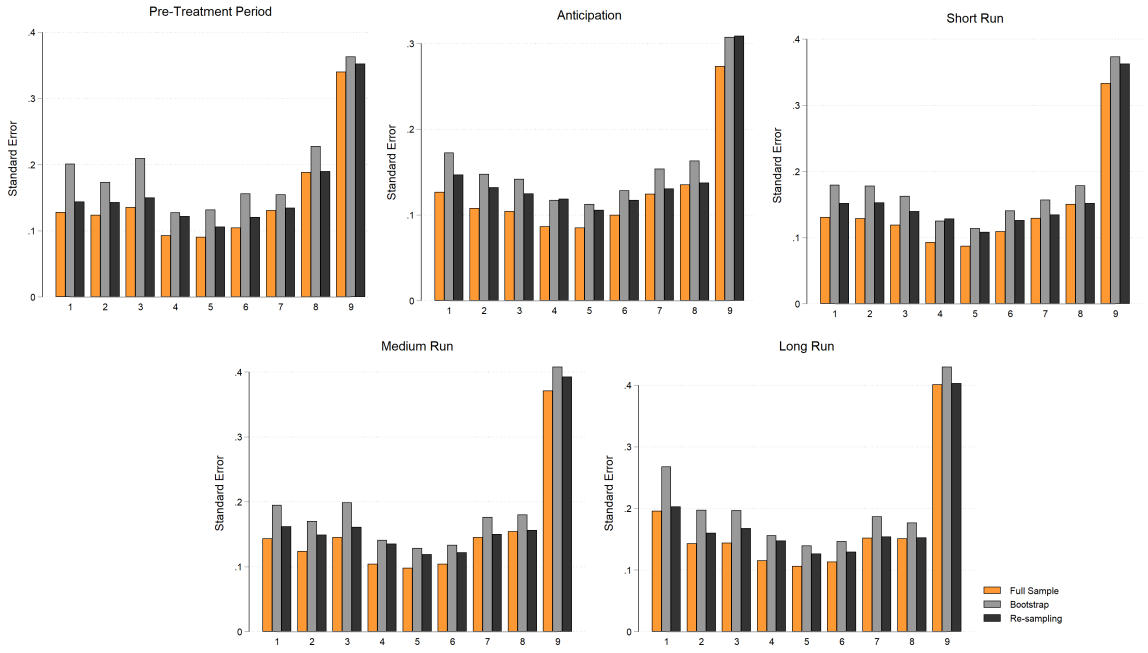

Figure 1.E.5: The comparison of the standard errors obtained by estimating the full sample, using the bootstrap, and the re-sampling from the control distribution.

1.F Results without Imputation

This appendix implements the causal inference framework on the data without imputed values. The main conclusion is that the conceptual results remain intact: PTAs gradually increase trade; and in anticipation only non-natural trading partners react to the PTA shock. However, the estimates are noisier, and the standard errors are higher due to the reduced power. Moreover, the magnitude of the averages is slightly reduced.

To understand why, let me first present the comparison between the normalized makers shares calculated using raw data and the data with imputed values. The correlation between the two shares is 0.98, and 0.99 between their logs. Table 1.F.1 shows the summary statistics for the raw (not imputed) shares and shares obtained after imputing the trade volumes. First, the number of observations is substantially higher for the (log) shares calculated with imputed data. The differences in means across the entire sample suggest that imputation leads to lower average shares for both pairs with and without PTAs. The standard deviation for the raw shares is slightly higher for all types of pairs.

Table 1.F.1: Summary statistics of normalized market shares calculated with and without imputation

		N Obs	Mean	Std. Dev	Min	Max
PTA=0	Raw	2,465,521	2.60	161.33	0	78,081
	Imputed	2,465,521	2.55	159.21	0	78,249
	log(Raw)	887,269	-1.94	2.71	-19.72	11.26
	log(Imputed)	1,455,399	-2.12	2.29	-19.70	11.27
PTA=1	Raw	167,897	18.24	165.09	0	22,826
	Imputed	167,897	17.69	143.95	0	12,672
	log(Raw)	132,468	-0.38	2.92	-17.81	10.03
	log(Imputed)	157,681	-0.48	2.75	-17.81	9.45

Note: The normalized market shares are substituted with zeros whenever they are missing.

Figure 1.F.1 plots the average normalized market shares by year for countries with and without PTAs. For both series the shares using imputed trade track closely the shares calculated in the raw data. Figure 1.F.2 reveals the main differences between the two shares: the distribution for the shares with imputed data is slightly skewed to the right (left panel), and particularly so for the control units (right panel). Such

situation occurs because because many missing values (i.e. values that are imputed) occur for smaller and poorer countries which tend to under-report their trade.

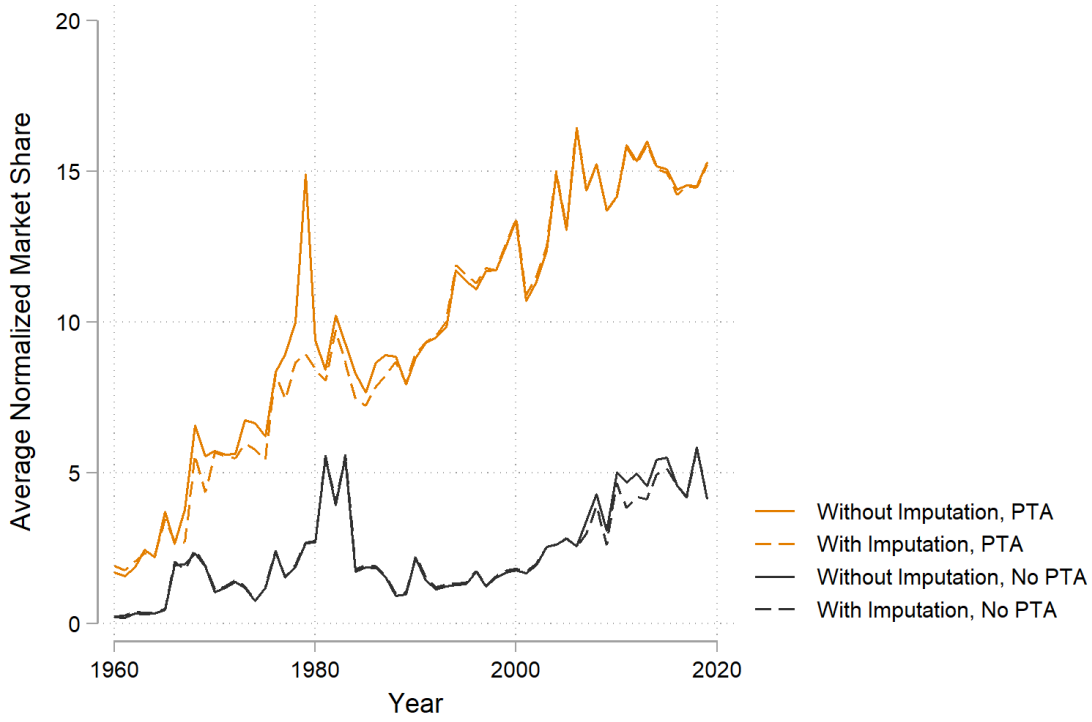

Figure 1.F.1: Average normalized market shares calculated using raw data and data with imputed trade volumes for pairs with and without PTAs, 1960-2019.

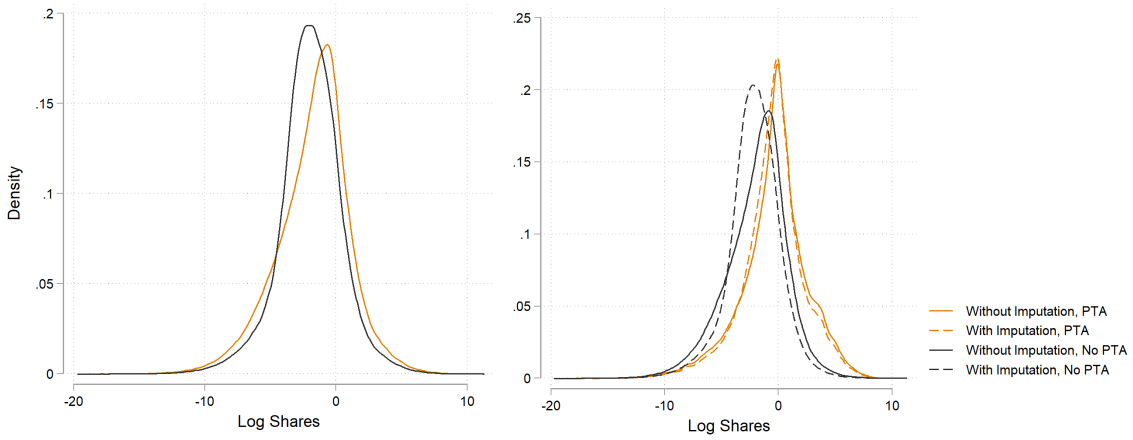

Figure 1.F.2: The distribution of normalized market shares calculated using raw data and data with imputed trade volumes in the full sample (left panel) and by treatment group (right panel)

Now let me report the results of the entire study, using the dataset with normalized market shares where the trade volumes were not imputed. I use exactly the same procedure as in the body of the paper. The blocking procedure groups pairs into ten subsamples. Table 1.F.2 shows the percentage increases in normalized market shares of the country pairs with PTAs relative to control pairs for different time periods. Comparing the results with Table 1.6, we can conclude the magnitudes of the point estimates are lower. Moreover, the estimates for the anticipation and short run period are not statistically significant. This happens due to both the decreased average estimates, and the increased standard errors (recall that the standard deviation of the measures is higher in the case of raw data). Figure 1.F.3 plots the means of each block, and the weighted average across blocks, along with 95% confidence intervals. Overall, it visually confirms the result of PTA effects kicking in gradually over time.

Table 1.F.2: Average PTA effects in different time periods.

	Anticipation [t-5; t=0)	Short Run (t=0; t+5]	Medium Run (t+5; t+10]	Long Run (t+10; t+15]
Coefficient	0.03	0.14	0.29	0.28
Std. Err.	0.07	0.07	0.07	0.07
Percent	3%	15%	34%	32%

Note: 'Coefficient' is the weighted average of the block estimates from estimating Equation 1.6 for each block within a given time period. 'Standard error' is the mean of the standard error distribution from the bootstrap procedure described in Appendix E. The percentage increase of normalized market shares of treated pairs relative to controls is calculated using the standard formula for interpreting dummy variable coefficients: $\exp(\hat{\tau}) - 1$.

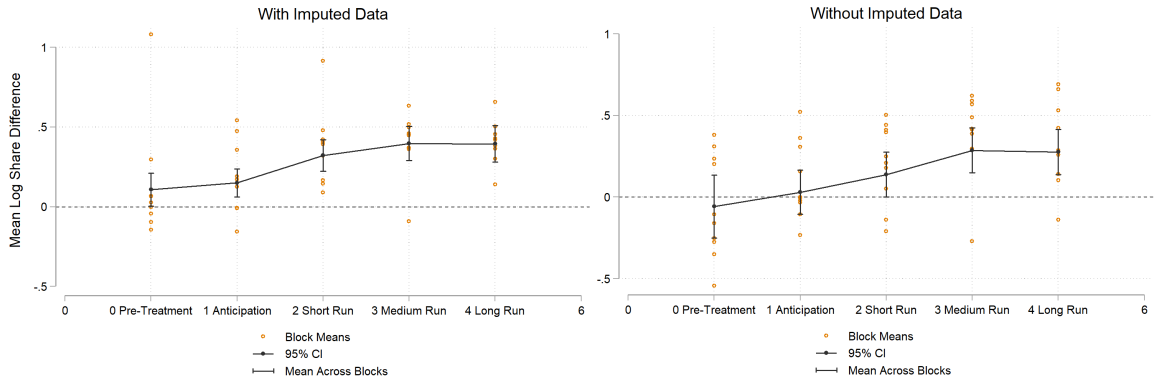

Figure 1.F.3: Block means and average PTA effects for different time periods for normalized market shares calculated using imputed trade volumes (left panel) and using raw data (right panel)

Finally, Figure 1.F.4 shows the point estimates for each block and each time period. In general, the results appear to be much noisier, in particular for the anticipation and the long run. However, we can still observe that for some of the lower-index blocks – corresponding to non-natural trading partners – the anticipation effects are present and are statistically significant; and are on average higher than for natural trading partners. In the short and medium run we observe a gradual increase in point estimates for all types of country pairs. These results are carried on to the long run period, although with increased standard errors for many blocks.

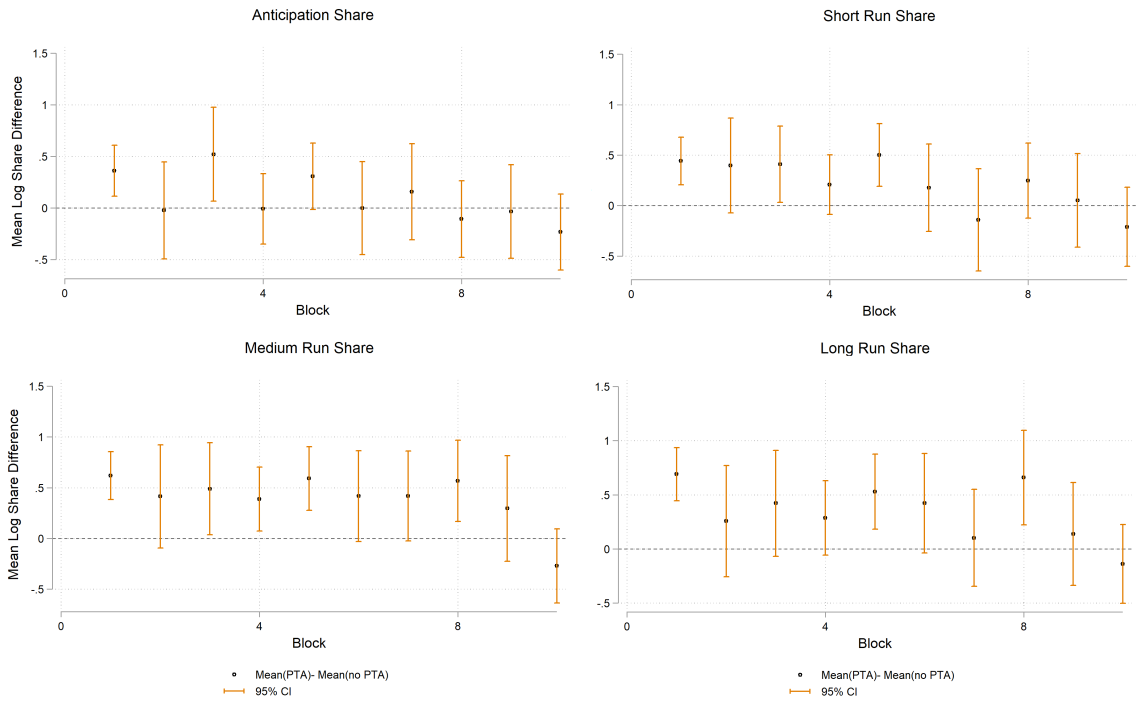

Figure 1.F.4: Average treatment effects within blocks in different time periods.

Note: The figure plots the point estimates of $\hat{\tau}$'s from Equation 1.6 for each of the nine blocks and each time period. The 95% confidence interval is calculated using the standard errors obtained from the bootstrap procedure described in Appendix E.

1.G Numerical Simulations

Figure 1.6 shows the point estimates of the blocking estimator are consistently lower than those obtained by applying other methods. This appendix constructs a numerical simulation which demonstrates that the blocking estimator performs better in a model with non-random PTA assignment.

This stylized numerical example starts off by creating an economy using the gravity model described in Section 6. A small modification concerns the structure of trade costs which are now assumed to have two components: transport costs t_{ij} and trade policy cost β_{ij}. The total trade cost then has the following form:

$$\tau_{ij} = t_{ij}\beta_{ij}$$

where $\beta_{ii} = 1$ and $t_{ii} = 1$.

Then, every period I augment the trade cost, and use the 'exact hat algebra' in a series of static model exercises to get the new equilibrium income distribution and trade flows. In each simulation iteration the parameters of the initial economy are drawn from uniform distributions, and are then augmented by trade shocks, generating panel datasets of trade flows. I simulate 500 such datasets with 50 countries and 10 periods each. In each period the transport costs t_{ij} reduce by 5% for all country pairs where $i \neq j$. The 10% reductions in trade policy costs β_{ij} (reflecting a PTA formation) are designed in two distinct cases:

1. Random PTA assignment: any country pair gets a PTA with a probability of 30%.

2. Non-random PTA assignment: country pairs which are more important to each other than their average trading partner have a higher probability of getting a PTA. In particular, recalling the intuition behind the normalized market shares, if $\bar{s}_{ij} = 1/2(s_{ij} + s_{ji}) > 1$, then a pair gets a PTA with a probability of 60%, while other pairs get a PTA with a probability of 30%.

In each simulated panel dataset I estimate the effects of PTAs (reductions in trade policy costs) using three different estimators: the Ordinary Least Squares (OLS) estimator, the Fixed Effects (FE) estimator, and the blocking estimator.[34]

[34] In this stylized numerical example the application of propensity score matching and entropy balancing does not make much sense, since there are no covariates. Due to the lack of covariates, the blocking estimator blocks on the the distribution of normalized markets shares.

Figure 1.G.1 plots the distribution of the estimates obtained by different estimation methods. The left panel corresponds to the datasets simulated using the random PTA assignment, which the right panel shows the distribution of estimates in case of non-random PTA assignment. The dashed vertical line in both cases indicates the true reduction in trade costs. In case of random PTAs, as expected, all estimators are able to capture correctly the true trade cost reductions. [35] For the non-random reductions, however, the OLS and the FE estimators tent to overestimate the effects of PTAs to a larger extent than the blocking estimator.

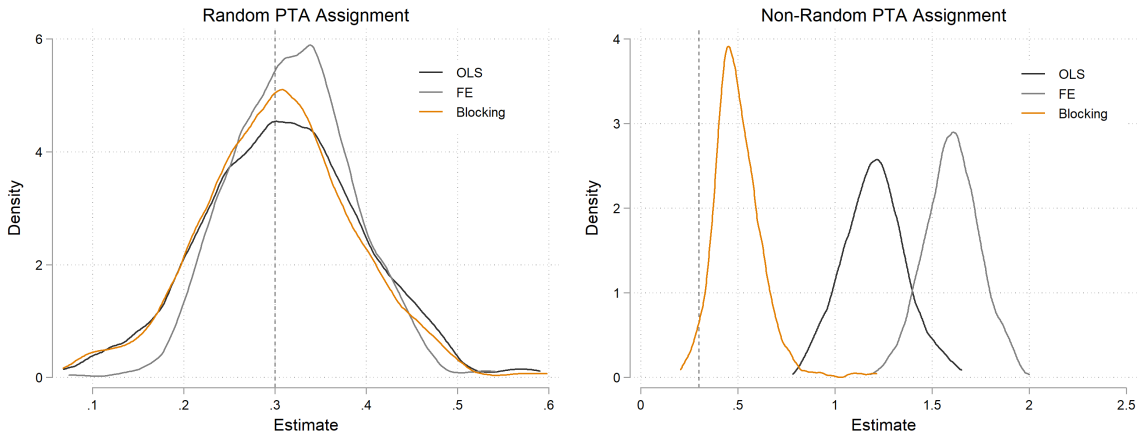

Figure 1.G.1: The distribution of the estimates obtained by applying different types of estimators in the simulated datasets.

Note: For each type of estimator the kernel density is estimated using the 500 point estimates from different estimators. The dotted vertical line indicates the true reductions in trade costs.

This numerical simulation demonstrates that the blocking estimator performs better in case of non-random PTA assignment, even when data is generated by the gravity model. Clearly, as highlighted by the applied empirical literature, the non-parametric methods are not immune to biases, and this stylized simulation confirms this fact for the distribution of the means. However, when combined with the confidence interval estimation, the blocking estimator is the only one that includes the true value, as shown in Figure 1.G.2.

[35] The distributions in Figure 1.G.1 represent only the point estimates, which, combined with the standard errors would contain the true estimate in the confidence intervals.

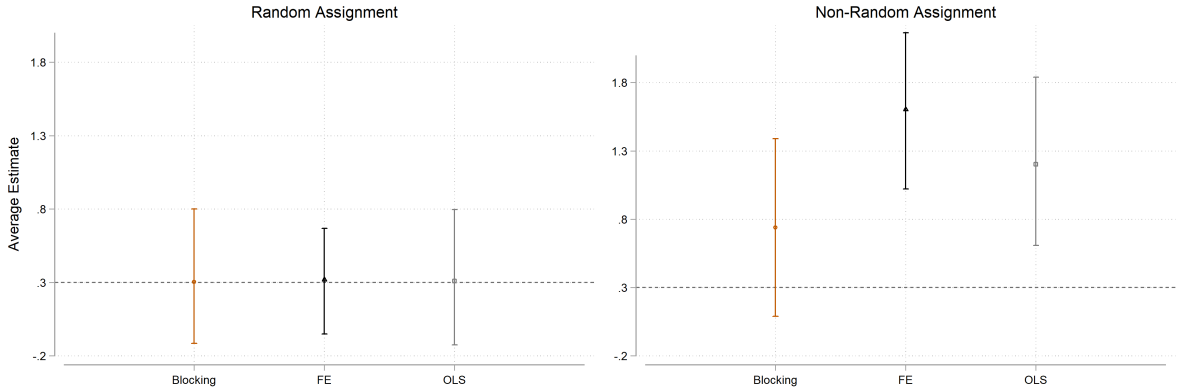

Figure 1.G.2: The mean and the standard errors of the estimates obtained by applying different types of estimators in the simulated datasets.

Note: For each type of estimator the kernel density is estimated using the 500 point estimates from different estimators. The dotted vertical line indicates the true reductions in trade costs.

1.H Sensitivity of General Equilibrium Estimates

In the baseline version of the counterfactual exercises the trade elasticity is set at the conventional value, $\varepsilon = \sigma - 1$ with $\sigma = 6$. This appendix repeats two counterfactual exercises presented in the main body of the paper using different values of trade elasticity.

The first exercise is conducted for the counterfactual equilibrium in the long run, i.e. when there is no heterogeneity across types of country pairs. The reduction in the iceberg trade costs is defined using the estimate from the empirical part of the paper of 48% increase in the long run, and the reductions in trade costs are defined using the different values of elasticity: $\varepsilon = 3$ (reduction in iceberg trade costs is 16%), $\varepsilon = 5$ (baseline reduction of 9.6%) and $\varepsilon = 7$ (reduction in iceberg trade costs of 6.86%). Thus, the role of elasticity is two-fold in the model: on the one hand, it amplifies the trade effects of trade cost changes, but on the other hand it decreases the magnitude of reductions in iceberg trade costs.

Table 1.H.1 below shows the main moments of the distributions of gross growth rates for different variables. The first one is the distribution of changes in welfare (real consumption): in the baseline model specification the average change in real consumption is 0.05%, and is very similar across different specifications. The larger the value of elasticity, the smaller is the standard deviation: the distribution 'shrinks', with minimum values rising (from -17.09% to =6.15%), and maximum values decreasing (from 28.48% to 12.27%). With larger values of trade elasticity the average normalized market shares for all countries also become smaller, with average growth of 26.08% for $\varepsilon = 3$ and 18.21% for $\varepsilon = 7$. Similarly, the dispersion of the distribution reduced with larger elasticity values. Unpacking the changes in the shares into trade between RCEP members and outsiders shows that the countries that are directly affected by the shock increase their shares more with higher value of trade elasticity: the mean increase is 32.34% for $\varepsilon = 3$, while with $\varepsilon = 7$ normalized market shares of RCEP countries more than double. At the same time, the outsiders are redirecting trade relatively less for higher values of elasticity: the increase in normalized market shares for low values of elasticity is 26.08%, while it is 18.21% for higher elasticity value.

The second counterfactual exercise presented in the main text utilizes the heterogeneity in point estimates across blocks. Table 1.H.2 presents the point estimates and the corresponding percentage reductions in iceberg trade costs by block, depending

on the value of elasticity. These reductions are used in the counterfactual exercises to compute the changes in welfare and normalized market shares in anticipation and short run.

Table 1.H.1: Descriptive statistics for the distributions of gross growth rates of real consumption, and normalized market shares, following the trade cost shock in the long run, for different values of elasticity.

Statistic	$\varepsilon = 3$	$\varepsilon = 5$	$\varepsilon = 7$
Welfare (real consumption)			
Mean	0.9977	0.9995	0.9995
Std	0.0401	0.0233	0.0161
Min	0.8291	0.9173	0.9385
Max	1.2848	1.1834	1.1227
NMS of all countries			
Mean	1.2608	1.1958	1.1821
Std	0.7707	0.6866	0.6694
Min	0.1549	0.1474	0.1460
Max	8.6934	8.0558	7.9141
NMS of RCEP with RCEP			
Mean	1.3234	1.5624	1.5274
Std	0.1868	0.2827	0.2662
Min	0.8671	0.9010	0.9066
Max	1.6890	2.0531	1.9943
NMS of others with others			
Mean	1.2608	1.2137	1.1821
Std	0.7707	0.7009	0.6694
Min	0.1549	0.1474	0.1460
Max	8.6934	8.0558	7.9141

Note: The values are calculated using the model presented in Section 6 for different values of the elasticity of substitution. The values correspond to different statistics in the distributions of gross growth rates of different variables. The top panel is the mean, the standard deviation, the minimum and the maximum values of the growth rates for welfare (real consumption) for all country pairs. The second panel presents the statistics for the distribution of the growth rates of normalized market shares for all countries. The third panel presents the statistics for the distribution of the growth rates of normalized market shares of RCEP members trading with each other. Finally, the last panel presents the statistics for the distribution of the growth rates of normalized market shares of pairs outside of RCEP trading with each other.

Table 1.H.2: Block coefficients and corresponding percentage iceberg trade cost reductions use in the counterfactual general equilibrium exercise, for different values of elasticity.

Block	Anticipation coefficient	Anticipation iceberg trade cost reduction			Long run coefficient	Long run iceberg trade cost reduction		
		$\varepsilon = 3$	$\varepsilon = 5$	$\varepsilon = 7$		$\varepsilon = 3$	$\varepsilon = 5$	$\varepsilon = 7$
1	0.54	18.05	10.83	7.74	0.63	21.05	12.63	9.02
2	0.39	13.30	7.98	5.70	0.46	15.18	9.11	6.51
3	0.19	6.34	3.81	2.72	0.52	17.34	10.41	7.43
4	0.36	11.92	7.15	5.11	0.44	14.95	8.97	6.41
5	0	0	0	0	0.50	16.67	10.00	7.14
6	0	0	0	0	0.37	12.39	7.43	5.31
7	0	0	0	0	0.50	16.80	10.08	7.20
8	0	0	0	0	0.37	12.29	7.37	5.27
9	0	0	0	0	0.15	5.09	3.05	2.18

Note: The coefficients correspond to regression adjustment coefficients for each block, resulting from a blocking procedure applied to year 2015, following the methodology outlined in the empirical section of the paper. Zero coefficients correspond to block point estimates that were not statistically significant. The corresponding iceberg trade cost reductions were calculated using different values of trade elasticity.

Table 1.H.3 compares the changes in welfare (real consumption) for different values of trade elasticity for the RCEP members. The values are presented in percentage changes, and it is clear from the table that with the exception of Myanmar and Cambodia, RCEP members experience negligible changes in welfare. The differences in welfare generated by varying the levels of trade elasticity are also small, with larger values of elasticity generating slightly smaller gains in anticipation and long run. This happens due to the fact that larger values of trade elasticity correspond to lower reductions in iceberg trade costs, as shown in Table 1.H.2. For countries that are most affected, varying the levels of elasticity has large effects: for Myanmar, for example gains in anticipation are 7.05% for the value of $\varepsilon = 3$, and 'only' 2.82% for $\varepsilon = 7$. Similarly, Table 1.H.4 presents the percentage changes in average normalized market shares by block in anticipation and long run, for varying levels of elasticity, and demonstrates that, on average, larger elasticity values produce smaller changes in normalized market shares (again, due to reduced size of the shock).

Table 1.H.3: Percentage changes in welfare (real consumption) for RCEP members following the trade cost shock in anticipation and long run, for different values of elasticity.

Country	Period	$\varepsilon = 3$	$\varepsilon = 5$	$\varepsilon = 7$
Australia	Anticipation	0.0045	0.0026	0.0019
	Long run	0.0104	0.0061	0.0043
China	Anticipation	0.0006	0.0004	0.0003
	Long run	0.0004	0.0002	0.0002
Indonesia	Anticipation	0.0045	0.0027	0.0019
	Long run	0.0048	0.0028	0.0020
Japan	Anticipation	-0.0001	-0.0001	-0.00001
	Long run	0.0019	0.0011	0.0008
Cambodia	Anticipation	1.0833	0.6686	0.4844
	Long run	1.6171	0.9255	0.6478
South Korea	Anticipation	0.0025	0.0016	0.0011
	Long run	0.0025	0.0015	0.0010
Myanmar	Anticipation	7.0461	4.0294	2.8212
	Long run	16.1542	9.7187	6.9716
Malaysia	Anticipation	0.0127	0.0080	0.0059
	Long run	0.0141	0.0082	0.0058
New Zealand	Anticipation	0.0059	0.0034	0.0024
	Long run	0.0033	0.0019	0.0014
Philippines	Anticipation	0.0016	0.0010	0.0008
	Long run	0.0078	0.0046	0.0032
Thailand	Anticipation	0.1237	0.0787	0.0578
	Long run	0.5283	0.3061	0.2155
Vietnam	Anticipation	0.0036	0.0022	0.0016
	Long run	0.0162	0.0092	0.0064
Average	Anticipation	0.6907	0.3999	0.2816
	Long run	1.5301	0.9155	0.6550

Note: The values are calculated using the model presented in Section 6 for different values of the elasticity of substitution. The trade elasticity parameter is defined in the model as $\varepsilon = \sigma - 1$. The values correspond to percentage changes in real consumption for RCEP members in anticipation and long run. Trade cost shocks in different periods are defined using the values specified in Table 1.D.16.

Table 1.H.4: Percentage changes in average normalized market shares of RCEP members' trade with each other, by block, following the trade cost shock in anticipation and long run, for different values of elasticity.

Block	Period	$\varepsilon = 3$	$\varepsilon = 5$	$\varepsilon = 7$
1	Anticipation	39.89	38.06	37.32
	Long run	-25.61	-24.55	-24.11
2	Anticipation	49.27	47.85	47.30
	Long run	0.32	0.57	0.67
3	Anticipation	20.30	20.59	20.75
	Long run	37.18	36.39	36.07
4	Anticipation	39.45	38.99	38.86
	Long run	14.45	14.29	14.24
5	Anticipation	0.24	0.43	0.53
	Long run	55.55	53.16	52.22
6	Anticipation	1.66	1.92	2.05
	Long run	52.93	51.17	50.48
7	Anticipation	-1.35	-1.04	-0.89
	Long run	49.59	47.78	47.09
8	Anticipation	-3.28	-2.79	-2.56
	Long run	41.26	40.01	39.51
9	Anticipation	2.64	-2.15	-1.92
	Long run	13.69	13.69	13.69
Average	Anticipation	15.95	15.76	15.72
	Long run	26.59	25.84	25.54

Note: The values are calculated using the model presented in Section 6 for different values of the elasticity of substitution. The trade elasticity parameter is defined in the model as $\varepsilon = \sigma - 1$. The values correspond to percentage changes average normalized market shares of RCEP members' trade with each other, by block, in anticipation and long run. Trade cost shocks in different periods are defined using the values specified in Table 1.D.16.

1.I Comparison of General Equilibrium Estimates

In Figure 1.6 I demonstrated that the point estimates of the empirical gravity model with three-way fixed effects are larger in the long run, compared to the blocking estimator (68% vs. 48% increase in normalized market shares). This appendix aims to answer the question of how much this difference in partial equilibrium estimates translates into the general equilibrium predictions. I use the 68% point estimate in the baseline version of the model with trade elasticity $\varepsilon = 5$, which translates into 13.6% reduction in iceberg trade costs in the long run (compared to 9.6% in the baseline).

Table 1.I.1 presents the main moments of the distributions of gross growth rates for different variables for the case of iceberg trade cost reductions obtained using the blocking estimator (baseline) and empirical gravity model with three-way fixed effects. The average changes in real consumption are very similar (a simple t-test cannot reject the null hypothesis that the difference is equal to zero for the two welfare vectors). This, however, is not surprising, given the magnitudes of the changes in welfare: they are negligible for a vast majority of countries in the sample. Similarly, we cannot reject the null hypothesis of no difference between the means of vectors of normalized market shares for all countries (panel two of Table 1.I.1).

The averages for all countries, however, are hiding some important differences between the two model which matter for individual countries. For example, there is a considerable difference for Myanmar, with 18.3% increase in real consumption predicted by the baseline model, and 27.7% increase predicted by the gravity-based model. Moreover, panel three of Table 1.I.1 shows that the normalized market shares of RCEP countries trading with each other are very different for the two estimates: the mean increase in normalized market shares for the baseline model is 56.24%, while it is almost double of that when using gravity-based trade cost estimate (90.87% increase). Figure 1.I.1 clearly demonstrates the large differences in the two distributions (t-statistics for the difference in means is -23.89). Table 1.I.2 also shows the percentage increases in average normalized market shares for RCEP members for the baseline and gravity based estimates. Gravity-based estimate predicts average shares which are 61.56% larger on average than the baseline model averages. Thus, if we were to use gravity estimates in the general equilibrium exercise we would substantially overestimate the trade reallocation for the RCEP countries.

Table 1.I.1: Descriptive statistics for the distributions of gross growth rates of real consumption, and normalized market shares, following the trade cost shock in the long run, for baseline and gravity-based estimates.

Statistic	Baseline	Gravity-based
Welfare (real consumption)		
Mean	0.9995	1.0007
Std	0.0233	0.0323
Min	0.9173	0.9175
Max	1.1834	1.2772
NMS of all countries		
Mean	1.1958	1.1984
Std	0.6866	0.6950
Min	0.1474	0.1468
Max	8.0558	8.0811
NMS of RCEP with RCEP		
Mean	1.5624	1.9087
Std	0.2827	0.4421
Min	0.9010	0.8717
Max	2.0531	2.6393
NMS of others with others		
Mean	1.2137	1.1984
Std	0.7009	0.6950
Min	0.1474	0.1468
Max	8.0558	8.0811

Note: The values are calculated using the model presented in Section 6 for different values of iceberg trade cost reductions. The values correspond to different statistics in the distributions of gross growth rates of different variables. The top panel is the mean, the standard deviation, the minimum and the maximum values of the growth rates for welfare (real consumption) for all country pairs. The second panel presents the statistics for the distribution of the growth rates of normalized market shares for all countries. The third panel presents the statistics for the distribution of the growth rates of normalized market shares of RCEP members trading with each other. Finally, the last panel presents the statistics for the distribution of the growth rates of normalized market shares of pairs outside of RCEP trading with each other.

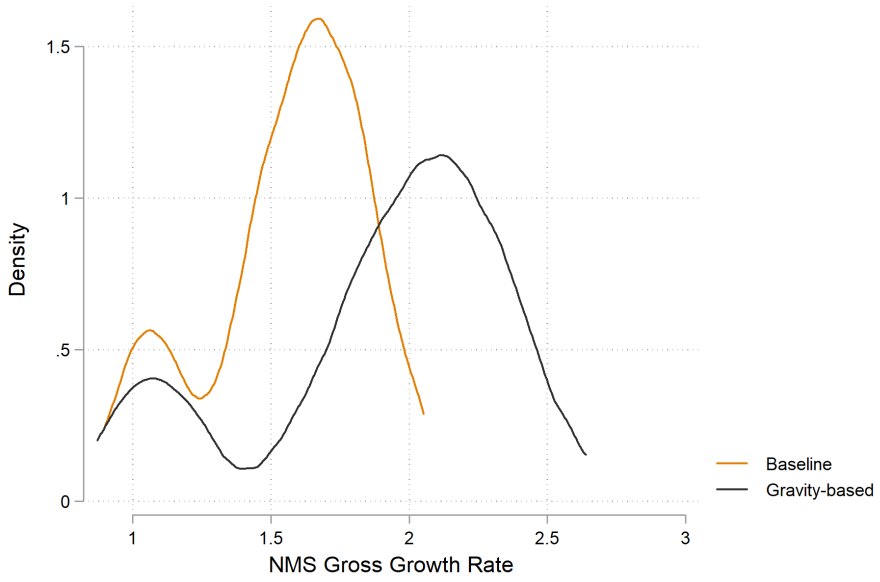

Figure 1.I.1: The distribution of gross growth rates of normalized market shares for RCEP countries' mutual trade, for baseline and gravity-based estimates.

Table 1.I.2: Average percentage increase in normalized market shares for RCEP members in trade with each other, following the trade cost shock in the long run, for baseline and gravity-based estimates.

Country	Baseline	Gravity-based
Australia	61.08	94.08
China	48.33	85.86
Indonesia	57.30	94.64
Japan	70.95	109.93
Cambodia	75.82	122.85
Korea	45.40	76.99
Myanmar	5.91	8.16
Malaysia	35.92	66.78
New Zealand	72.61	109.49
Philippines	67.60	105.17
Thailand	60.70	96.40
Vietnam	73.29	120.04

Chapter 2

Trade Agreements: a Historical Perspective and Modern Trends

"Everything new is well-forgotten old"
—Russian proverb

2.1 Introduction

Being the contemporaries of the 21st century, we are used to thinking about the system of the global commercial relationships as a collection of rules established by the General Agreement on Tariffs and Trade (GATT)—the rule system countries are largely following today. This system, however, was not created from scratch. Centuries of evolution of the commercial relationships led to the formation of the rule system that the countries use today.

What are the important features of trade agreements, and how did they evolve over time? In this chapter I take a long-run perspective and analyze the evolution of trade agreements from the 'first wave of globalization'—the beginning of the 19th century—to modern times. I collect a novel dataset containing information about trade agreements dating back to 1815. I show that many of the features of the contemporary trading system were developed long before the GATT was signed; and that the usage of separate provisions was adjusted and polished along with the political and economic developments of the different historical periods.

Economic history books identify fundamental features of historical events, and summarize the broad trends in commercial relationships over centuries. For exam-

ple, the major work by Findlay and O'Rourke (2007) traces the history of the international economy from its earliest beginnings to the 21st century. Brown (2009) focuses on the development of the multilateral trading system and outlines the main negotiation issues in a given historical period. Published papers, on the other hand, focus more on isolated historical periods and theoretically or quantitatively study the consequences of commercial policies or historical events on various outcomes. A case in point are the studies focusing on quantifying the effects of the treaty network that formed in late 19th century in Europe (see, for example, Accominotti and Flandreau (2008) or Lampe (2009)).

To complement the existing literature this paper presents and measures a particular aspect of commercial relationships, namely, their contractual side, over a long period of time. Analyzing the 'hard' and measurable elements of the commercial treaties in the historical context can help understand better to what extent trade policy was forming the economic and commercial relationships, and to what extent trade policy was merely responding to the political and economic challenges of a given time period. The analysis of trade agreements' textual data can also trace the design features' formation, geographical spread, and the evolution of treaty content, leading up to the contemporary rule system.

This chapter summarizes and presents the information from texts of the treaties in two different ways. First, by reading the treaty texts, I identify and code the key design features that were agreed upon by the signatories. These features are broadly defined and described in Section 3.1 and Appendix B, and also constitute the underlying principles of the modern system: reciprocity, the most-favoured-nation (MFN) principle, the national treatment, the different aspects of the market access rules, and other commercial issues. Additional features (see Section 3.2), such as membership and textual characteristics (length, scope and enforceability) help create a more complete picture of other dimensions across which agreements differ from each other. Second, I analyze the evolution of these principles over time in terms of the changes in the design of the features, prevalence of the features in different time periods, and the geographical distribution of treaties.

The evolution of the treaty texts in the periods identified for the purpose of the analysis largely reflects the political and economic situation of the time. In the early 19th century, following the disruptions of the Napoleonic wars, the treaty formation reflected the relatively low desire of countries to engage into any new endeavours of trade liberalization in the period of relative peace and stability. In terms of the

treaty content in the early 19th century, the most-favoured nation provisions and the national treatment were widely present in commercial texts, while there was little innovation on other aspects. The agreements included a large number of issues, and bundled together the provisions on the movement of natural persons and commercial policy. The unprecedented development of technology and the search for new markets in the second half of the 19th century led to the expansion of the market access provisions, often in the form of a separate tariff schedule included in the agreements. These market access provisions became the major innovation of the network of treaties formed in Europe in the 1960s-1970s, known as the Cobden-Chevalier network.

The following period, describing the interwar years, was full of disruptive historical events. The war itself, the global flu pandemic, and the Great Depression—all amidst tense political rivalries—found their reflection in the types of treaties formed during this period. In fact, the interwar period can be characterized by a series of important policy innovations. Exemptions and specifications of the exact circumstances in which the more general rules and principles can be applied have become an inherent feature of commercial treaties. Flexibilities, such as security regulations, health-related exceptions, or carve-outs for some more privileged trading partners, which appeared in the interwar commercial treaties, will later manifest themselves in the GATT, and become its prominent institutional feature.

The period after World War II in this chapter is referred to as 'the era of multilateralism.' While the foundations and the main rules of this system still constitute the core of the trading system nowadays, this period is characterised by the creation of the unifying commercial framework. From the point of view of policies that the GATT included, it can be seen as a collection of best practices developed in previous decades, crucially, including those flexibilities which were a common place in the interwar period. The revolutionary innovation was the GATT's ability to establish a multilateral system of trade relations, unlike the networks of bilateral treaties existing earlier. Another important innovation was the GATT's dispute settlement provision, which later evolved into a WTO court, maintaining one of the most important functions of the organization. The structure of the commercial relationships under the GATT also allowed regional integration initiatives and agreements to develop alongside the multilateral negotiations. Some of the most important and deep regional trade relations in Europe were forming during this period, and the ideas of regional integration spread to other parts of the world. The majority of the treaties

at the time focused on deepening market access and cutting tariffs, while including a conservative number of issues, in sharp contrast to the next period.

This next period, starting roughly from the beginning of the 2000s, can be characterized by the emergence and a wide spread of the 'new generation' of commercial treaties. With the GATT / WTO success in reducing trade barriers, and the overall demise of the WTO's negotiating function, a dense network of treaties emerged. These newer treaties spread in terms of their geographical coverage, while simultaneously expanding in their scope. Modern commercial treaties include a large number of issues, many outside of the current WTO mandate. At the same time, the extent to which these treaties can credibly enhance cooperation is unclear: many of the issues discussed in these treaties do not represent a legally enforceable provision.

In general, commercial treaties concluded in each distinct time period reflect the political and economic challenges of the time. The world trading system has developed and matured through a series of shocks, leading up to the familiar system of accumulated best practices countries use today in their commercial relationships. As Irwin and O'Rourke (2011) argue, understanding this evolution can help dealing with the challenges that the system is facing in the years ahead.

This chapter is organized as follows. Section 2 describes the data collection for the historical treaties, and the data sources for the contemporary treaty analysis. Section 3 defines the major features of the commercial treaties used to present the stylized facts. Section 4 describes the evolution of the treaty content, dividing the analysis into four time periods. Section 5 concludes and contemplates the possible future developments.

Definitions. Before proceeding further, it is worth spending some time defining the subject of this chapter more precisely.

First, following the conventional usage, I will use the term 'trade policy' in the meaning of 'commercial policy'. From a more formal viewpoint, the latter is broader, and can encompass policies that do not directly relate to trade in goods or even trade in services Brown (2009). For example, in the 19th century the commercial treaties included provisions on treatment of natural persons or merchants, foreign property rights, or rules of conduct for merchant ships and war ships. More modern commercial policies regulate the treatment of issues related to, for instance, intellectual property rights or government procurement. For the purpose of the analysis in this chapter all these policies will be united under the umbrella term 'trade policy'. The

same identity will be relevant when referring to agreements (or treaties): the term 'trade agreements' will be used interchangeably with 'commercial agreements,' referring to the broader set of policies that were included in these documents.

The customary set of terms that the literature on modern trade agreements is using includes 'preferential' or 'regional' agreements, 'bilateral' and 'plurilateral' agreements, and 'multilateral' trading system (or agreement). The latter is usually used to refer to the GATT/WTO agreement. However, as the WTO (2011) states: "all trade agreements – bilateral, regional, multilateral – are preferential in the sense that their benefits and obligations apply to members only, and non-members are excluded." Indeed, these distinctions can be somewhat blurred. In the 19th century, for example, all the agreements were signed bilaterally, but formed a network uniting many countries under a most-favoured nation regime. At the same time, the modern-days WTO system does not include a number of non-member countries.

Often the object of study is defined by a legalistic definition from the GATT/WTO framework. In particular, the trade agreements are understood as those which conform to the Article XXIV of the GATT (the formation of customs unions and free trade areas). The article states that countries can promote integration if (i) trade barriers are eliminated with respect to 'substantially all trade,' and (ii) trade barriers applied to other countries may not be higher than prior to the formation of preferential agreement. The clarity of this definition makes these types of agreements an attractive object to study. There are, however, several limitations to this definition. First, it does not allow us to study trade agreements which were formed prior to the GATT. Second, there are other forms of trade agreements which are important, but are outside of the scope of this definition. These agreements include, for example, preferential trade agreements that were not notified to the WTO.

For the purpose of this chapter I will use the term 'trade agreements' to refer to voluntary reciprocal treaties between two or more sovereign states which include provisions related to commercial policy. Although this definition is quite broad, it excludes colonial trade arrangements and non-reciprocal preference arrangements. Ultimately the more precise features of trade agreements under this definition will also depend on the data available, which is described in Section 2 and Appendix A. The GATT itself falls under this definition, however—acknowledging the importance of this trade agreement—will be analyzed separately.

2.2 Data Sources

The analysis of the design features and the content of trade agreements hinges crucially on the quality of the data available. Naturally, more modern trade agreements (since the creation of the GATT) are easier to analyze, with multiple databases being available in the digital format. Earlier agreements, however, are more difficult to study. The content and the quality of the data in this study will be defined, to some extent, by the data availability.

To study the agreements in the 19th century, I use two data sources. The primary data are the original commercial treaty texts. I collect and hand-code information about each commercial treaty that can be found across the 161 volumes of the "Consolidated Treaty Series" compiled by Clive Parry in 1969-1980 (Parry (1969)). These volumes cover all the treaties concluded by sovereign units from 1815 to 1919[1]. I choose only the treaties directly related to commerce, with some exceptions (see Appendix A for more detail). Since none of these volumes are yet available in any digital form, the contemporary natural language processing techniques cannot be used to analyze these texts. Instead, I devise a coding system which would make the data on these agreements roughly comparable to the data available on modern trade agreements. This system comprises a binary coding of a number of features: the most-favoured nation principle, the national treatment, market access provisions and a few others described in more detail in the next section.

The secondary data source for the 19th century trade agreements is the "Handbook of Commercial Treaties" by Brauer and Kasten (1922). This book presents the digests of commercial treaties, conventions, and other agreements of commercial interest between all nations, covering the years from 1654 to 1922. These treaties are much broader in scope, and include those which are not directly aimed at fixing the commercial interests. For example, some agreements are devoted to treatment of certain territories, or to river navigation. This data also includes additional and supplementary treaties (see Appendix A for more detail). I extract information about the detailed name of the treaty, its signatories and the year of signature. The name allows to categorize agreements into different groups, and gives a more complete picture of what types of commercial issues were on the agenda in the 19th century. The analysis

[1]The full collection (225 volumes) of the "Consolidated Treaty Series" contains all treaties signed since 1648, but I restrict my sample to those treaties signed after the Congress of Vienna in 1815, since this is the year coined as the start of the 'first wave of globalization.'

using this dataset is presented in the Appendix A.

For the agreements signed after World War I, in the interwar period, and during World War II, I use the League of Nations Treaty Series collections. Again, I go through each of the 205 volumes to search for commercial agreements and hand-code information about them using the same protocol as for the primary data source. These volumes are available in a digital format, but are not of an adequate quality to transform the texts into a machine-readable texts.

To analyze the content of the trade agreements signed after the formation of GATT in 1947, I use three databases available online. The databases have different coverage both relating to their substance, and in terms of the agreements covered.

The first source is the Design of Trade Agreements (DESTA) Database (Dür et al. (2014)). The raw dataset includes 781 agreements signed from 1948 to 2021. After deleting the superseding treaties, amendments and additional protocols, as well as partial scope treaties and framework agreements, the sample includes 556 treaties signed over the same time period. The main advantage of DESTA is that it includes all commercial treaties, and not just the ones listed by the WTO. In addition, DESTA codes a number of content elements of trade agreements, including distinct provisions on commercial issues (such as market access, trade in services, etc.).

The second source is the World Bank's Deep Trade Agreements (DTA) Dataset (Hofmann et al. (2017)). It includes 318 agreements signed from 1958 to 2019. The most attractive feature of the dataset is its content coding: for each agreement the researchers map 52 commercial provisions, dividing them into items under the current WTO mandate, and provisions outside of the WTO. Each agreement is then recorded to contain a certain provision. In addition, this dataset provides the extent of each provision's legal enforceability, if it is mentioned in the text. The DTA coding distinguishes between two types: legally enforceable provisions, and provisions which are excluded from the dispute settlement.

Finally, I use Texts of Trade Agreements (ToTA) Dataset (Alschner et al. (2018)). The raw dataset includes the text corpus of 447 agreements notified to the WTO between 1948 and 2015. After deleting non-reciprocal treaties and agreements which are not translated into English, the remaining sample includes 414 treaties that entered into force between 1949 and 2017. The main advantage of ToTA is the possibility to extract semantic information from the texts (such as, for example, the counts if words with varying degree of legal enforceability), as well as the number of chapters and length of the treaties (see the details on the variable construction in Appendix A).

2.3 Features of Trade Agreements

Trade agreements, like other legal documents and textual objects, are highly dimensional objects. There is a great variety of features that can be attributed to a single contract. The challenge is to identify and summarize the most important features in a number of variables that would characterize the agreement well, and at the same time would reduce the dimensionality of the text.

In the context of trade agreements, some of the most important features pertaining to the institutional design of the trading system are well defined—these features are the underlying principles of the GATT, as well as the characteristics of the degree of market access (described in Section 3.1). There are additional characteristics which help create a more detailed understanding of the treaties, or proxy for their content and complexity (see Section 3.2). Depending on the historical period, all of these features can transform and slightly differ in their exact formulations, keeping, however, their fundamental properties. The description of these fundamental properties and differences are the focus of this section. The next two sub-sections briefly describe the definitions, and Appendix B provides the examples of the treaty texts.

2.3.1 Principles and Design Features

Reciprocity. In international trade, broadly defined, the principle of reciprocity means that favours and benefits should be somehow returned to each other by the trading partners, and mutual concessions with respect to market access and other commercial policies have to be made. In the literature, reciprocity is considered to be a major institutional feature of the modern trading system (Bagwell and Staiger (1999)). It was also an important feature of the trade treaties in the past (see Appendix B for the example of the wording). As stated in the definition in the introduction, only reciprocal treaties will the the focus of this chapter.

Most-Favoured Nation. The MFN principle is fundamental to the modern trading system, and it ensures that countries are not discriminating between their trading partners. The GATT/WTO version of this principle requires members to grant the most favourable tariff or regulatory treatment given to the 'like product' or 'like services and service suppliers' coming from any other member. When looking at the historical context, however, the scope of the application and the types of MNF treat-

ment can differ substantially. For example, the 19th century trade agreements included MFN provision in a variety of forms: applied to natural persons (citizens), officials (consuls), merchants, vessels, or products; and covering the treatment at the port, or treatment related to paying duties and taxes.

Early commercial treaties included an extensive set of provisions regarding natural persons, as opposed to purely trade-related subjects such as merchants, commercial vessels and goods. In the dataset that I construct, the MFN provision only counts as present if it relates to commercial activities—if an agreement only includes MFN treatment of natural persons, such an agreement is considered to not include an MFN provision (see Appendix B for examples of these types of provisions).

Similarly, the MFN provisions distinguished between treating the merchant vessels and goods in the port, and extending the MFN provision to the amounts and types of duties that the countries were paying (see Appendix B). These are the two types of MFN I distinguish when constructing the dataset: some of the earlier treaties contained provisions related to port procedures, but continued to discriminate in terms of duties payable on imports or exports.

There is another feature of the MFN treatment which is highly important in the historical context. After the creation of the GATT, the unconditional form of MFN principle has been universally accepted, but this was not the case in the historical treaties. Many of the earlier treaties contained the so-called 'conditional MFN'. Put simply, the unconditional form of the MFN treatment records no conditions or circumstances in which the reciprocal concessions will be made between the contracting parties; while the conditional form stipulates the concessions being made in return for an equivalent (Brauer and Kasten (1922)). I do not code this distinction, yet in the following chapters I will describe the general trends regarding the application of the two forms of the MFN treatment (see an example of the two forms in the text treaties in the Appendix B).

As in today's MFN provisions, the historical texts often contained some exceptions form the MFN principle. These exceptions included, for example, provisions related to public health, national security, or carve-outs for other preferential trading partners (see Appendix B). Although not all coded as part of the dataset, the following sections will provide a general overview of the evolution of these exceptions.

Finally, in the modern commercial treaties, as provided by the DESTA dataset, the MFN treatment requires parties to the agreement to grant to other parties the most favourable tariff and regulatory treatment granted to any third party (third-party

MFN).

National treatment. The second fundamental principle of the modern trading system is the national treatment, which provides that imported and domestically produced goods and services should be treated equally. Similarly to the MFN provisions, national treatment can be granted with respect to natural persons, or internal duties and charges for imported goods. In the dataset, only the latter is used to distinguish between trade agreements with and without the national treatment clause (see Appendix B).

Market Access. Market access is typically used as an umbrella term for a number of commercial regulations that can be used to restrict imports, such as tariffs or non-tariff barriers. As most early treaties were not nearly as sophisticated and developed as the ones we have today, I view market access in a very broad sense. For example, if a given treaty contains provisions related to the rules of the importation of a specific list of products, or explicitly stipulates tariffs and payments for specific goods, such an agreement would be considered to have a market access provision (see examples in Appendix B). If there is an appendix in the form of a tariff schedule, such an agreement will also be considered as having market access provisions. I separately code the presence of a detailed tariff schedule for a given agreement.

Non-tariff barriers were scarce in the 19th century treaties. After the Spanish flu pandemic in 1918, the non-tariff measures mostly included the sanitary provisions, which were often bundled into the same chapter with provisions on national security. The presence of these measures is also coded in the dataset. Similarly, in the advance of the World War II, quotas and quantitative restrictions have become widespread. I will discuss the general trends on these types of provisions in the subsequent sections (for the wording see Appendix B).

Other Commercial Issues. Both historical and contemporary commercial treaties are typically very comprehensive, and cover a number of issues besides MFN provisions, national treatment and market access. Given the abundance of these issues and also the differences in coverage between historical treaties and modern agreements, it would be difficult to create a coding system that would be relevant for historical comparison.

In particular, many historical treaties include issues that are currently covered in agreements outside of commercial domain, such as conditions of residence or travel for natural persons, immigration and emigration (police protection and civil rights); admission of diplomatic and consular officials, and their rights and activities; possession and disposal of, or succession to, real and personal property; exemption from military service, municipal functions, forced loans, and extraordinary levies; freedom of religion and right of burial; treatment of vessels seeking refuge from damage or shipwreck; and other provisions. In addition, there are issues that relate to commerce, but are absent form modern-day treaties, such as, for example: vehicles and instruments of communication and transportation; navigation, quarantine and harbor regulations, and dues relating thereto; treatment of commercial travelers and their samples; coasting trade and port-to-port trade with foreign cargoes. Finally, there are some issues that were discussed in the historical treaties, and are still present in many modern treaties in some form. These include, for example, conditions for importation, exportation, transit, transfer, warehousing; protection to patents, trademarks, copyrights, and other industrial property rights; rights of commercial, industrial, or financial associations; and exterritorial jurisdiction.

Modern treaties, especially those concluded in the last 20 years, tend to be very broad in scope. They can include a wide variety of commercial issues, such as, for example, intellectual property rights, services, investment, public procurement, competition, electronic commerce, labour rights, environmental protection, and many others. Some of these issues will be analyzed separately using the available datasets.

2.3.2 Other Characteristics

Membership. When referring to membership in trade agreements, the main aspect relates to the number of members that a given agreement has. Historical treaties were predominantly bilateral. The GATT's main feature, as I show in the analysis below, is its expanded membership, and what we now define as 'multilateralism'. Other trade agreements existing parallel to the GATT can include a varying number of members. In addition, due to the varying levels of integration, some trade blocks themselves can conclude agreements with other countries or blocks[2].

One of the issues with regards to data organization is multiple membership. For

[2]Note that this type of membership does not necessarily imply accession to the existing agreement. An example of such a situation would be the customs union of the European Union with Turkey.

the historical treaties a very relevant feature is that the same country pair can sign multiple trade agreements over a span of years (superseding agreements). This can happen due the agreement's expiry (designated period for a treaty expiry was a common feature for the agreements in the 19th century), or due to the additions of new features, changes or amendments[3]. Modern preferential treaties also have modifications and amendments. In the analysis I will thus make a distinction between first-time treaties and superseding agreements. In addition to superseding agreements (i.e. those treaties that are made at distinct points in time for the same country pairs) modern trade agreements can also be overlapping: the same country pair is covered by two different agreements at the same time. For example, Colombia and Peru are both in Andean Group (Bolivia, Colombia, Ecuador, and Peru) and in Pacific Alliance (Chile, Colombia, Mexico and Peru). Since this happens relatively rarely in the data, I will count only the first concluded agreement between the two countries.

Length. Length of a commercial treaty is measured in two ways. First, it is defined as the number of distinct chapters in a given text. Since every chapter roughly corresponds to a separate commercial issue, the number of chapters can proxy for the scope (coverage) of the agreement. Although not precise, the advantage of using this variable is that it can be obtained for both historical and contemporary treaties. The second variable I will use to analyze the length of a commercial treaty is the average number of words per chapter. This variable is only available for digitized treaties, and comes from the machine-based text analysis (see Appendix A for details).

Legal enforceability. The degree of legal enforceability of a treaty is only available for the modern treaties, and is measured in two ways. First, the DTA dataset categorizes each provision into (i) provisions not mentioned in the agreement or not legally enforceable; (ii) provisions mentioned, legally enforceable but explicitly excluded by dispute settlement provision; and (iii) provisions mentioned and legally enforceable. Another measure used to analyse the extent of enforceability is constructed using text mining techniques. A trade lawyer identified words which can potentially lead to litigation or dispute using international arbitrage, as well as words that signal weak commitment. Using the counts of such words across treaties, I construct a measure of legal enforceability for each treaty (see Appendix A for details).

[3]See Appendix A on data collection for more details on the treatment of this issue.

2.4 The Evolution of Treaty Content

The types of policies included in commercial treaties have substantially evolved over the time horizon analyzed in this chapter. I analyse four periods, each broadly marked by the stylized facts regarding most relevant policy developments. These policy changes are, in turn, largely a response to the changing economic and political contexts.

The first period is called by Findlay and O'Rourke (2007) as 'the international economy of the 19th century,' and often referred to as 'the first wave of globalization,' spans the years from 1815 to 1914. The treaties of this period established a sound basis for the two principles of non-discrimination foundational for the contemporary multilateral system (the MFN and the national treatment), and paved the way to market access liberalization. The second period includes two World Wars and the interwar period. The tumultuous years brought important innovations to trade agreements of the time, mostly providing different safeguards and escape clauses. These harsh years are followed by the period covering roughly the second part of the 20th century, broadly characterized by relatively stable and expanding multilateralism. The GATT—the major institution governing commercial relations after World War II—also allowed for deeper economic integration among groups of countries. Roughly from the beginning of the 2000s starts the fourth period of analysis, marked by the shift in the types of treaties countries tend to sign. Expanding the multilateral rules in terms of their depth and scope, the generation of 'new trade agreements' started emerging during this period.

2.4.1 The International Economy of the 19th Century (1815-1914)

The Napoleonic wars had a disruptive effect on the international trade, and had a number of long-run implications for the political and economic context. Most importantly, the high costs of the wars of 1972-1815 brought a durable period of relative political stability to the European continent. The Congress of Vienna instituted the main features for a exceptionally long period of peace (until the Crimean war in the 1850s).

The wars brought an end to many mercantilist protectionist policies, with tariffs becoming more popular as a means of regulating foreign trade. In addition, the Congress of Vienna guaranteed the freedom of navigation in many rivers. Natu-

rally, many protectionist measures remained or were instituted in the aftermath of the wars, with the British Corn Laws passed in 1815, as the most prominent example[4]. In general, however, European trade was being gradually liberalized over the larger part of the 19th century (Findlay and O'Rourke (2007)).

One of the most outstanding developments in relation to trade in the 19th century is the remarkable advancement of the transportation technology. The increased use of steamships, the improvement of roads, the building of canals and the proliferation of railroads—all resulted in a sharp decline in transport costs. For example, the index of British ocean freight rates dropped by about 70% between 1840 and 1910 (Harley (1988)). The dramatic decline in transportation costs defined a somewhat secondary role for trade policy: even though the protectionist backlash of the 1870s stalled the process of integration, it was not able to revert the benefits of the trade cost reductions.

Historians also attach first-order importance to the British unilateral policy as a major factor in 19th century trade policy developments. A series of liberal reforms in 1820s and 1830s and the final repeal of the Corn Laws in 1846 made Britain a pioneer in trade liberalization efforts. The widely praised Cobden-Chevalier treaty between Britain and France is thus seen as a breakthrough in international commercial relationships which brought about a major episode of trade liberalization in Europe. This treaty between two major countries and large trading partners slashed import prohibitions in France, abolished export restrictions on British coal exports, lowered British tariffs on wine, and provided a detailed tariff schedule. The treaty is also praised for including the MFN provision which later served as a benchmark for the European trade treaty network formation. While there is no doubt on the role of the Cobden-Chevalier treaty in the formation of the resulting network of treaties, below I discuss that many of the features of the Cobden-Chevalier agreement were present also in the earlier treaties.

The period starting from the 1970s is generally coined as the 'protectionist backlash.' A severe worldwide economic crisis hit in 1873 and lasted until 1877. This period was also characterized by the strong nationalist rivalries among European countries (Brown (2009)), which erupted in a series of trade wars. At the same time,

[4]On 10 April 1815 a volcano called Mount Tambora in the present-day Indonesia erupted, in what is known as he most powerful volcanic eruption in recorded human history. This event resulted in world-wide reductions in temperatures, leading to extreme weather events and harvest failures in many areas around the world, and called for urgent policy measures.

the rise in tariff levels was not excessive. Bairoch (1993) estimates that the average level of import duties on manufactures was 13% in Germany, 20% in France and Sweden, 18% in Austro-Hungary and Italy, 9% in Belgium and Switzerland, and only 4% in the Netherlands, while tariffs on many agricultural products and raw materials were low or zero.

There were a total of 608 first-time commercial treaties signed over the period between 1815 and 1914, with another 272 renewing or substituting the pre-existing treaties (see Figure 2.4.1). All of the treaties were signed bilaterally, and therefore the total number of country pairs participating in treaties is dwarfed by the number of signatories of the commercial treaties nowadays (Figure 2.4.2). Approximately 8% of the treaties in the dataset before 1914 were not exclusively devoted to commercial issues. In general, the vast majority of the treaties of that period (77%) included an MFN provision, either in the form of port or duty privileges. Most of the treaties signed by the US were providing the MFN treatment in its conditional form, while European treaties were stipulating both types of provisions. Similarly, around half of all agreements included national treatment, and 38% of the treaties included market access provisions.

Despite little attention given in the literature to the treaties signed before the Cobden-Chevalier agreement, these agreements of the earlier period contained a number of important provisions that would later form the basis for the network. Table 2.4.1 breaks down the treaty sample into treaties concluded before and after the Cobden-Chevalier agreement. Out of 291 first-time treaties signed before 1860, 81% had an MFN provision of some form. Similarly, 61.9% of those treaties contained a national treatment provision related to commerce. Around a quarter of all agreements had some form of market access provisions included. Usually the market access provisions were in the form of specified rates for a number of distinct product categories. Only around 5.5% of all treaties before 1960, however, provided a detailed tariff appendix.

The literature studying the history of European trade emphasises the importance of the Cobden-Chevalier network, and highlights the unconditional MFN status granted to the signatories of the treaties as a major breakthrough. Table 2.4.1, however, clearly demonstrates that the main innovation prompted by the conclusion of the Cobden-Chevalier treaty was the inclusion of tariff schedules. Among the treaties signed between 1860 and 1914, almost a third of all treaties contained a detailed tariff schedule. In fact, 75 superseding agreements were signed to include a tariff specification

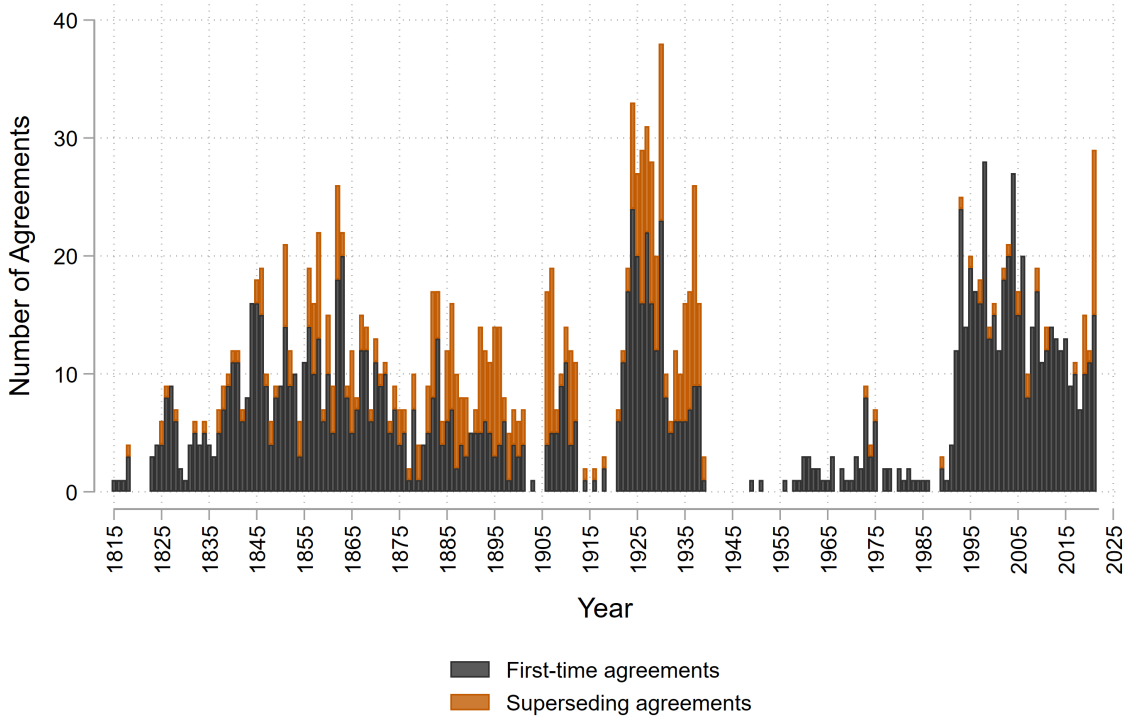

Figure 2.4.1: Number of first-time and superseding treaties entering into force by year, 1815-2022.

Note: The data before 1920 comes from the Consolidated Treaty Series (volumes 65-225); the data between 1920 and 1946 comes from the League of Nations Treaty Series (volumes 1-205); the data after 1947 comes from the DESTA dataset.

additionally to the existing provisions. Even if the remaining two thirds of the agreements still did not stipulate any detailed tariff appendix, slightly less than a half of all agreements contained market access provisions of some other form (usually specifying a list of goods and / or tariff rates explicitly in the text of the treaty). By some estimates, the tariff levels at the time were cut by about half as a result of the treaty network (Brown (2009)).

Indeed, when studying the effects of the Cobden-Chevalier network on aggregate trade, Accominotti and Flandreau (2008) find no effects. Instead, Lampe (2009) finds that tariff reductions are a key to driver of trade liberalization in the 1860s. Using disaggregated trade flows he shows that the effects of the Cobden-Chevalier network vary with tariffs across different commodity groups.

The Cobden-Chevalier treaty network is also believed to have introduced a greater

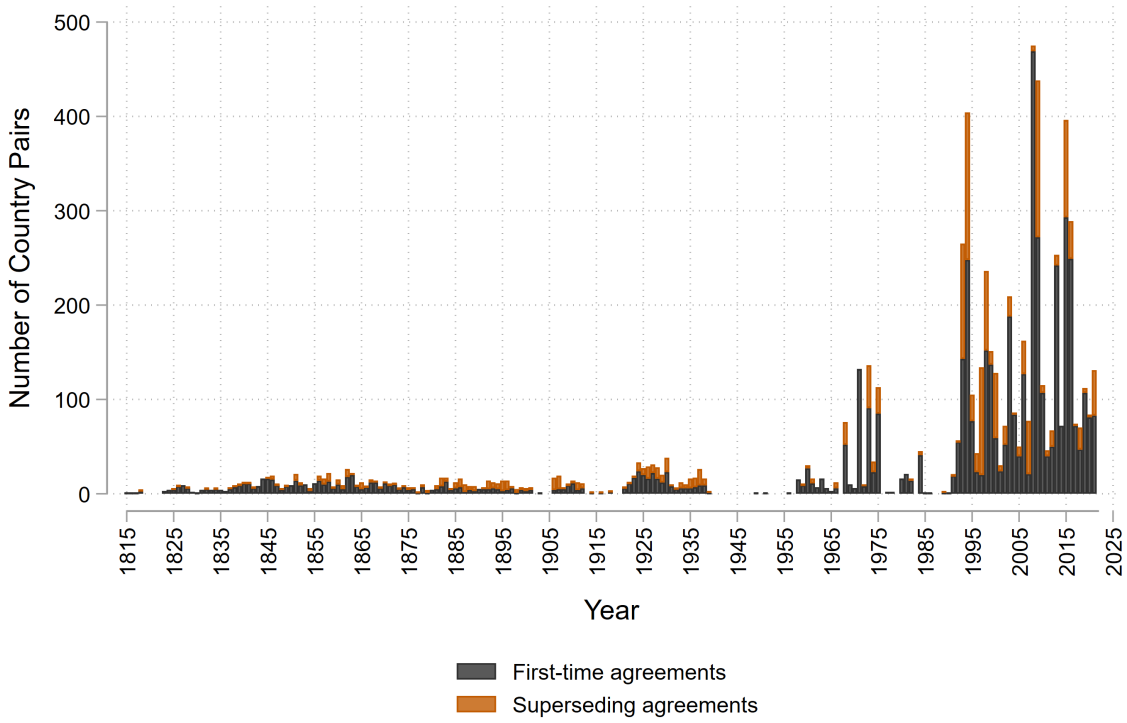

Figure 2.4.2: Number of country pairs in first-time and superseding treaties entering into force by year, 1815-2022.

Note: The data before 1920 comes from the Consolidated Treaty Series (volumes 65-225); the data between 1920 and 1946 comes from the League of Nations Treaty Series (volumes 1-205); the data after 1947 comes from the DESTA dataset. The graph does not include the African Economic Community agreement, and African Continental Free Trade Area.

measure of certainty into trade relations (Shafaeddin (1998)), since most of the agreements were singed for a period of 10 years. Contrary to this belief, I find that earlier agreements were also stipulating similar provisions regarding expiry. For example, for the 260 treaty texts signed in 1815-1859, for which I find the explicit mention of the expiry period, the average period of stipulated treaty duration is 8.9 years. The same metric for the 215 treaties for which I have this data in years 1859-1914 is equal to to 9.5 years.

Table 2.4.1: The percentage of treaties including a given provision

	Treaties in 1815-1859		Treaties in 1860-1914	
	First-time treaties	All treaties	First-time treaties	All treaties
MFN (Port)	78.4	75.9	77.8	77.7
MFN (Duty)	74.2	72.5	75.9	76.6
National treatment	61.9	62.2	48.6	50.9
Market Access	24.1	25.8	42.3	46.3
Tariff schedule	5.5	6.6	26.6	30.1
Number of treaties	291	349	311	521

Note: The data before 1920 comes from the Consolidated Treaty Series (volumes 65-225); the data between 1920 and 1946 comes from the League of Nations Treaty Series (volumes 1-205). The number of other agreements for which provisions are recorded is slightly lower than the total number of agreements due to missing values (see Appendix A).

To understand what types of treaties were the most prevalent during this period, Table 2.4.2 counts the percentage of treaties with different combinations of the four most important provisions: MFN related to duties, MFN related to the treatment of commercial activities (most notably in the ports of entry), national treatment, and market access. Whenever a given provision is present in a treaty, such entry is coded as 1 in the left-hand side of the table. There are 16 possible combinations of provisions. The right-hand side of the table provides the information on the percentage of the treaties that have a given combination: the first column relates to years 1815 to 1859, while the second column provides information for the years 1860 to 1914.

During the early 19th century, almost half of all treaties (42.6%) included the three types of non-discrimination provision, with no mention of market access. Just the two MFN provisions were typically included in the next 16.2% of treaties. All the other combinations of provisions are present in less than ten percent of treaties. In the period after the Cobden-Chevalier treaty was signed, the type of the treaty including the three provisions for non-discrimination remained the most prevalent, however, it took up a much smaller proportion (27.7%). It is closely followed by the type of treaty with the two MFN provisions, but without the national treatment (22.8%). Notably, the most complete treaties (including all four provisions) now take up a significant share—16.7%.

A typical treaty of the time, however, was longer and much more complex than just including the four provisions coded in Table 2.4.2. During the period from 1815 to 1914 trade treaties would include around 19 chapters on average. Figure 2.4.3

shows the average number of commercial and non-commercial chapters per year[5]. There is a slight increase in the number of issues included in the agreements around 1850s-1870s.

Table 2.4.2: The percentage of treaties including a given combination of provisions, by time period

MFN (Duty)	MFN (Port)	National Treatment	Market Access	1815-1859	1860-1914	1915-1939
1	1	1	1	6.2	16.7	23.8
1	1	1	0	42.6	27.7	28.2
1	1	0	1	6.5	8.4	11.5
1	1	0	0	16.2	22.8	31.3
1	0	1	1	0	0	0
1	0	1	0	0.7	0	0
1	0	0	1	0.7	0.3	0
1	0	0	0	1.4	0	0
0	1	1	1	0.3	0	0
0	1	1	0	1.7	0	0
0	1	0	1	2.4	1.0	0
0	1	0	0	2.4	1.3	0
0	0	1	1	2.7	1.3	0
0	0	1	0	7.6	2.9	0
0	0	0	1	5.2	14.5	4.0
0	0	0	0	3.4	3.2	1.3

Note: The data before 1920 comes from the Consolidated Treaty Series (volumes 65-225); the data between 1920 and 1946 comes from the League of Nations Treaty Series (volumes 1-205). The left-hand side indicates with zero / one an absence / presence of a given provision, resulting in 16 different combinations for the four provisions.

The treaties would typically start with a preamble introducing the signatories (usually the heads of states). Most treaties would then describe the principles of non-discrimination applied to the citizens and the subjects of the treaty. The text would be describing the rights, privileges, and immunities for the nationals of the respective countries, including issues related to residency, dwelling, private property protection, legal representation, etc.

Later on in the text, the declaration of reciprocal freedom of commerce is made. The following parts of the treaties often describe in detail the treatment of merchants,

[5]The number of non-commercial issues is approximated at the half of the agreement.

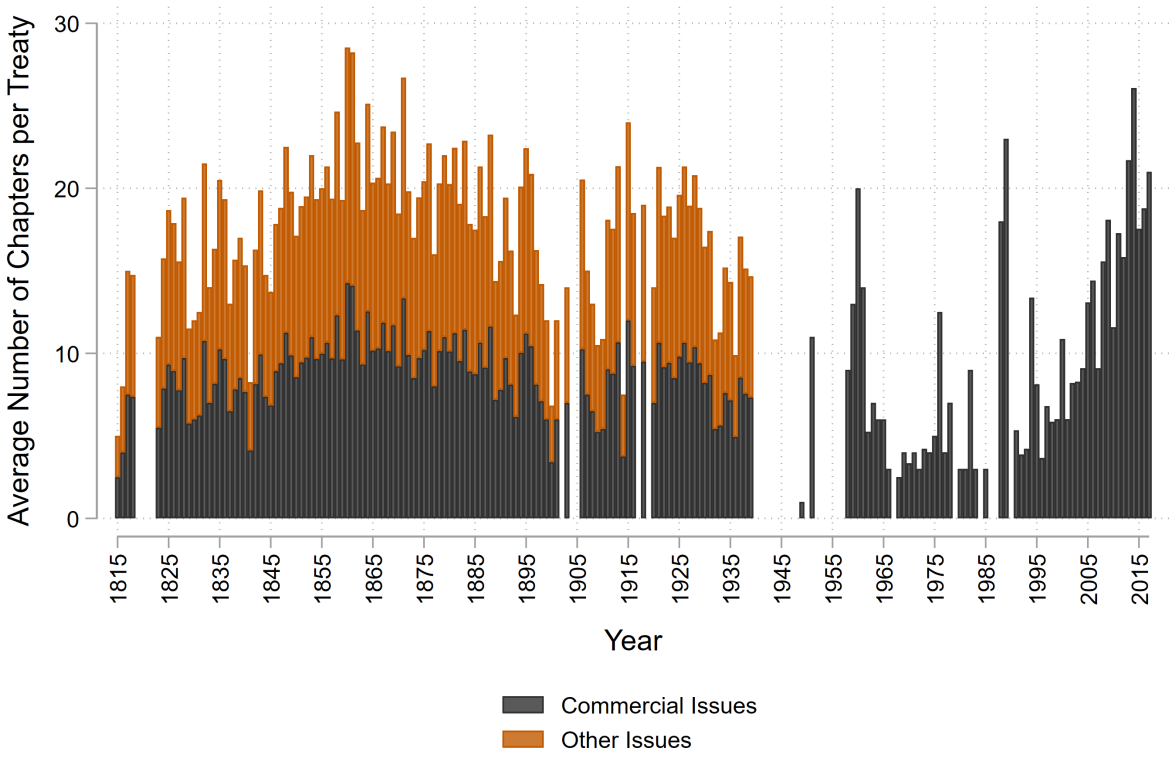

Figure 2.4.3: Average number of chapters by year, 1815-1939.

Note: The data before 1920 comes from the Consolidated Treaty Series (volumes 65-225); the data between 1920 and 1946 comes from the League of Nations Treaty Series (volumes 1-205); the data after 1947 comes from the ToTA dataset. The number of other issues in the earlier treaties is approximated at the half of the total number of chapters per agreement.

vessels and cargoes. The agreements would stipulate the perfect equality in the treatment of the vessels, freedom of importation of products allowed under the laws of the respective countries, the MFN provision related to dues and charges, and the national treatment with respect to dues of tonnage, harbour, lighthouse, and other payments. A chapter would be typically devoted to establishing no prohibitions on imports and exports, with an exception for weapons and gunpowder, or like products. Exceptions relating to coasting trade are mentioned in the vast majority of the treaties.

Following these more general rules, the treaties would go into the specifics for the rights of deposit and the unloading of the cargoes—usually providing a national treatment. Similarly, the texts would describe the rules to identify the nationality of

the vessels, and rules related to the vessels' flags. Most of the treaties also include provisions related to shipwreck, and the ownership of the property remaining after such events.

Finally, many treaties (especially the ones relating to amity and friendship, on top of commerce and navigation), would include a number of consular provisions. For example, some chapters could be devoted to the actions permissible in the event of war between the signatories. A considerable amount of chapters could be devoted to establishing consulates, appointing the consuls and the vice-consuls, describing their responsibilities and functions, and providing for non-discriminatory treatment of the official representatives. Finally, the last two chapters of every treaty are devoted to the treaty's duration, or the procedures to denounce the treaty; and ratification and its entry into force.

Besides the content, the regional distribution of the treaties is as an important feature of the treaties in a given time period. In the 19th century the European continent was the driver of world commerce and economic growth. The regional distribution of the trade policies embodied in the commercial treaties follows this overarching trend. Out of all the treaties concluded in this period 86.2% had a European country involved, and 39.3% of all treaties were concluded between two European countries. Table 2.4.3 shows the distribution of the number of treaties by broad geographical region in 1815-1914. The table clearly demonstrates the mass of agreements being concentrated among European countries. Counting the number of treaties can be misleading though, since European continent contained more countries and sovereign territories than other regions. Nevertheless, even when normalizing by the number of countries within a region, each European country had agreements with around four other European countries on average.

Table 2.4.3: Number of treaties concluded by region, 1815-1914

	Africa	Asia	Europe	North America	South America	Total
Africa	1	0	27	5	0	33
Asia	0	5	69	15	5	94
Europe	32	73	347	31	82	565
North America	0	1	12	3	5	21
South America	0	7	87	21	54	169

Note: The data before 1920 comes from the Consolidated Treaty Series (volumes 65-225); the data between 1920 and 1946 comes from the League of Nations Treaty Series (volumes 1-205).

The development of the provisions within the treaties is also largely driven by the developments on the European continent. Figure 2.4.4 plots the number of agreements with various types of provisions (MFN, national treatment, market access and tariff schedule) for three different regional groups: agreements between European countries, agreements of European countries with countries outside of Europe, and agreements of countries outside of Europe among themselves. The figure shows that the majority of treaties including an MFN or national treatment provisions were concluded among European countries in the decades preceding the Cobden-Chevalier treaty (dark grey bars are concentrated around 1830s-1850s). Market access and tariffs were, however, the focus of trade treaties after 1860. At the same time, European countries visibly sought to include the MFN and national treatment provisions in the treaties they were concluding in the middle of the 19th century with other countries (see the evolution of the orange bars). Similarly, there is an increase in the number of treaties containing market access provisions which European countries signed with the outsiders, following the popularity of those provisions within Europe.

The evolution of the treaty texts largely reflects the political and economic situation of this historical period. In the first part of the 19th century there was little innovation in terms of policies included in the treaties. At the same time, this is the period when the MFN and the national treatment provisions have firmly made their way into the treaty texts in agreements between the European countries, paving the way for their expansion through the future Cobden-Chevalier network. The treaty formation during this period reflected the relatively low desire of countries to engage into any new endeavours of trade liberalization in the period of relative peace and stability. With the development of technology and the British search for new markets, a more active period of liberalization began, following the signature of the Cobden-Chevalier agreement. The major innovation of the treaty network that formed afterwards is the inclusion of market access provisions, and the pursuit of more active trade liberalization through lowered tariffs. After the Franco-Prussian war of the 1870 and the onset of global recession, many unilateral protectionist measures were introduced, and a series of local trade wars erupted. This period of 'Long Depression' in the end of the 19th century, and the economic and political instability at the turn of the century was reflected in the reduced number of treaties signed, while introducing a series of amendments to the existing ones.

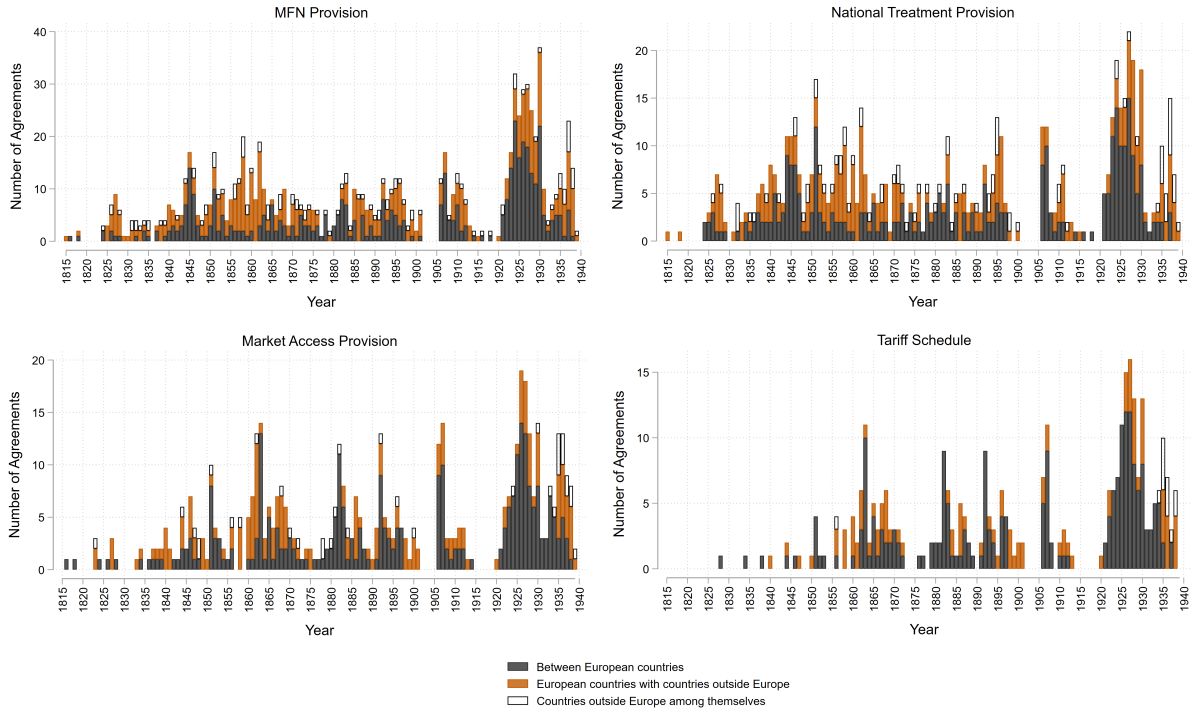

Figure 2.4.4: Number of agreements including a certain provision, by year and regional composition, 1815-1939.

Note: The data before 1920 comes from the Consolidated Treaty Series (volumes 65-225); the data between 1920 and 1946 comes from the League of Nations Treaty Series (volumes 1-205). The number of other agreements for which provisions are recorded is slightly lower than the total number of agreements due to missing values (see Appendix A).

2.4.2 The Wars and the Interwar Period (1915-1946)

The period between 1915 and 1946 is characterized by a series of dramatic historical events which manifested in the way countries conducted trade policies. World War I severely disrupted international trade relations. This disruption, however, did not result in the overall collapse of trade, but presented a more complex pattern: trade blockades had differential impacts across European countries, while countries in North America and Asia managed to substantially expand their exports (**?**). From the policy perspective, however, the war resulted in a series of restrictions, limitations, prohibitions and the tightening of government control over trade and shipping in many European countries.

Many of the restrictions remained in place after the war, and the Treaty of Versailles did not manage to lay down stable political foundations for the postwar po-

litical order, unlike the Congress of Vienna in 1815. A heavy legacy of war debts and reparations led to mounting political and social tensions in the 1920s. The failed attempts of the governments to return to the pre-1914 gold standard resulted in currency instability, leading to further disintegration of the world economy (Findlay and O'Rourke (2007)). The volatility of the exchange rates resulted in countries restoring to tariffs against "exchange dumping" (the fall in price of imported goods that was occasioned by the sudden depreciation of a trading partner's currency) (Brown (2009)). The new states that appeared after the war took largely a protectionist stance, while the Communist revolution in Russia was throwing it into an autarky.

While World War I and its immediate aftermath definitely stalled the process of integration, the protectionist tendencies were severely exacerbated by the onset of the Great Depression. The consequences of this unprecedented financial and economic crisis required radical measures to cope with the pressures on the balance of payments, and countries restored to currency depreciation and import controls. Quantitative restrictions became widespread—for example, estimates suggest that 58% of French imports were subject to some sort of quotas (Findlay and O'Rourke (2007)). Findlay and O'Rourke (2007) characterize this period as "the complete breakdown of the MFN principle of non-discrimination." As the treaty texts show, however, MFN was still a central element of the trade agreements signed at the time, but numerous exceptions which constitute the policy innovation of that period could be used to impose discriminatory prohibitions on imports. The same exceptions will later be mirrored in the GATT, and will become the foundations of GATT's institutional flexibility.

Another important event in trade policy of the time was the quiet shift of the US towards more liberal trade policy with the passage of the Reciprocal Trade Act, which provided for reciprocal reductions in trade barriers as the way to gain easier access to foreign markets. Together with the Fordney-McCumber Tariff Act of 1922, which incorporated the unconditional MFN policy, the US was set towards the multilateral cooperation path.

Given the largely negative historical context, perhaps one of the most surprising features of Figure 2.4.1 is the large number of treaties signed during the interwar period. There were a total of 367 treaties concluded between year 1915 and 1939, and 62% of those treaties were concluded for the first time, while the remaining ones were the superseding agreements. This expansion in the number of the first-time treaties can be explained by the creation of a number of new states in Europe that

quickly entered the existing network of agreements. For example, 29 of the treaties were concluded by Czechoslovakia, 27 by Latvia, 28 by Finland, 22 by Estonia and another 24 by Poland. Among the 140 superseding treaties, most were re-negotiated by Turkey (24), USA (19), and the UK (16). These treaties were still all concluded bilaterally, thus representing a relatively small part of the total number of country pairs covered (see Figure 2.4.2).

Table 2.4.4 describes the content of the agreements entering into force in 1915-1939. Nearly all of the agreements signed in that period contained a stylized form of the MFN principle. As I discuss later, although the MFN provision was widely present, the exceptions from it were extensively included in the treaty texts. More than half of the agreements also included a national treatment clause. Market access provisions have become widespread tools for foreign trade regulation during this time, with around half of the agreements containing them (see Table 2.4.4). Another important feature of the agreements in that period is the prevalence of tariff schedules—38.4% of the treaties included a detailed annex specifying the levels of duties charged for each specific item. The development of the detailed tariff schedules was also facilitated by the emergence of more uniform customs nomenclatures. In addition, the series of conferences convoked by the League of Nations in 1927 reached an agreement on a tariff truce, and encouraged countries to simplify their customs procedures (Brown (2009)).

Table 2.4.4: The percentage of treaties including a given provision

	Treaties in 1915-1939	
	First-time treaties	All treaties
MFN	94.7	90.2
National treatment	51.9	53.4
Market Access	39.2	47.1
Tariff schedule	33.0	38.4
Security and health carve-outs	54.6	56.1
Number of treaties	227	367

Note: The data before 1920 comes from the Consolidated Treaty Series; the data between 1920 and 1946 comes from the League of Nations Treaty Series (volumes 1-205). The number of other agreements for which provisions are recorded is slightly lower than the total number of agreements due to missing values (see Appendix A).

A few very important policy novelties emerged as a response to the dramatic events occurring during the interwar period. Most of the policies were stipulating

exceptions from the MFN principle, and would later pave their way into the GATT.

First, policies stipulating carve-outs for national security and protection of public health started to appear widely in the treaty texts (see Appendix B for the example text). This set of policies was later institutionalized in Articles 20 and 21 of the GATT. Table 2.4.4 shows that more that half of all the treaties signed in this period had some sort of exception from the principles of non-discrimination. This safeguarding mechanism, often praised as a flexible institutional feature of the GATT, was a direct response to the First World War. Moreover, although often overlooked, the global Spanish flu pandemic also manifested itself in the wording of the trade treaties. Shortly after the pandemic was over, the treaties started to include exceptions on import prohibitions on the grounds of public health.

The second set of exceptions introduced widely during this period was in relation to granting privileges to a certain set of countries. The agreement would specify, for example, that the MFN principle would not be granted to the same extent as to the member countries of a customs union or some other trade block. This sort of exemption would be later written into the GATT's Article 24. Appendix B demonstrates the type of language that would be used in treaties to provide this exception.

Next, various non-tariff barriers made their way into the treaty texts. The most prominent are the explicit quota schedules, akin to those used for tariffs. At least 17 treaties in the 1930s contained a detailed quota provision. Rules of origin and certificate requirements have also become an important tool for trade regulation (see Appendix B for the example text).

A typical commercial agreement of the interwar period would include 17 chapters on average (see Figure 2.4.3). Regarding the treaty type distribution, almost a third of all treaties included only two types of the MFN provisions, and did not provide for national treatment or market access (see Table 2.4.2). Almost 30% of all treaties included the three non-discrimination provisions, and around a quarter of all treaties were complete, including the MFN provisions, national treatment and market access rules.

Commercial treaties of the period were, in general, very comprehensive. Each chapter provided detailed rules and procedures for the application of commercial policies. Many treaties included extensive tariff annexes, with different lists compiled for different product categories. Some treaties also stipulated procedures for the changes in applied tariff rates and market access rules. The MFN provisions were very similar to the earlier treaties in their scope and language. This time, however,

specific exemptions were mentioned, such as those relating to border trade or customs union formation. Chapters prohibiting export or import restrictions were also detailing explicit carve-outs, such as policies related to public safety, sanitary measures, traffic of arms, state monopolies, etc. As before, treaties would often include extensive national treatment provisions with respect to natives, merchants, vessels, and the procedures on sales and consumption of the imported goods.

Substantial part of the treaty texts of the interwar period would be devoted to detailed procedures and paperwork for importation and exportation: customs declarations, customs valuation, certificates of origin, certificates for composition, purity, and sanitary conditions, etc. In addition, the treaties would often include rules for the operation of the joint stock companies, as well as rules for intellectual property protection, and protection against unfair competition.

With the overall development of the international law, commercial treaties have begun to 'outsource' a number of provisions to other international conventions, such as those regarding navigation, maritime transport, railroad usage, standardisation of the tonnage measurement system, dispute settlement, etc. As before, the treaties of this period would include provisions related to incidences of shipwrecks, freedom of transit, itinerant trading, as as well as consular appointments and functions. Ratification and rules for treaty denouncement were included as final provisions of the treaties. The vast majority of the treaties in the interwar period did not have a specific expiry date, and were technically meant to be indefinite. Instead, they stipulated the procedures that parties would need to carry out in case one of them decides to terminate the treaty at some point. Whether this feature was meant to bring stability is not clear, since the Second World War broke out, making many of these treaties obsolete.

As during the previous historical period, the interwar period treaties were geographically concentrated. 90% of all treaties included at least one European country, and 54.6% of all treaties were concluded between two European counterparts. Table 2.4.5 shows the number of treaties concluded by region in 1915-1939. Another notable development is the expansion of US treaties with Southern American countries. Figure 2.4.4 confirms that the number of treaties including the MFN, national treatment, market access provisions and tariff schedules was high with the European countries' participation, although over this period many agreements among countries outside of Europe start to adopt a similar language.

Overall, the commercial treaties concluded during the period from 1915 to 1939—amid the disruptive historical events—can be characterized by a series of important

Table 2.4.5: Number of treaties concluded by region, 1915-1939

	Africa	Asia	Europe	North America	South America	Total
Africa	0	1	2	1	0	4
Asia	0	3	21	5	3	32
Europe	1	55	200	16	10	282
North America	0	0	6	1	2	9
South America	0	2	20	12	5	39

Note: The data before 1920 comes from the Consolidated Treaty Series; the data between 1920 and 1946 comes from the League of Nations Treaty Series (volumes 1-205).

policy innovations. Many of those new features constituted some sort of exemption from the non-discrimination principles: either in the form of security regulations, health-related exceptions, or carve-outs for some more privileged trading partners. These innovations will later enter the GATT and become one of its important institutional features. Notably, the same exceptions that allowed to disrupt a large portion of trade for the treaty members in the interwar period, have become the foundations of GATT's institutional stability in the period of multilateralism that followed.

2.4.3 The Era of Multilateralism (1947-2000)

One of the most devastating military conflicts—World War II—had damaging consequences for the international trade. The effects were uneven across countries, and the trade disruptions created not just losers, but winners as well (Findlay and O'Rourke (2007)). In the aftermath of the War, political developments were dominating the world economic agenda. Soviet Union was consolidating the communist influence, and the Cold War was becoming a reality. Decolonization fueled nationalist moods in the newly created countries, and in the decades following the War, the developing world became more closed to international trade. On the contrary, the economies in Western Europe and North America began to slowly but steadily to liberalize their economies.

With political considerations leading the agenda, the situation after World War II was different compared to the aftermath of the Napoleonic wars. After the latter the world has seen an unprecedented advances in technology that were driving down the transportation costs. Despite some notable technological advances in the post World War II period, the evidence regarding the influence of these technologies on

freight rates is somewhat mixed. Hummels (1999) finds that the freight rates actually increased between 1950s and 1990s. The increase in the nominal rates, however, was a result of increasing prices of key inputs, such as fuel, and a variety of anti-competitive policies (Fink et al. (2002), Clark et al. (2004)). Ganapati and Wong (2023), on the contrary, show that transport usage by weight increased more than ten-fold from 1965 to 2020. They estimate that this increase was accompanied by a substantial fall in global transport costs between 1970 to 2014: the weight-based measure of transportation costs fell by 33-39%.

Another notable structural development in the world economy in the second half of the 20th century is the gradual industrialization of the developing world. Between 1960 and 2000, the share of manufacturing employment grew from 8 to 15% in East Asia, North Africa and Southwest Asia, and from 9 to 14% in South Asia (UNCTAD (2003)). These developments were accompanied by the accelerating pace of vertical specialization, with the manufacturing processes being subdivided into stages, with each stage located in a different country (Findlay and O'Rourke (2007)).

Even though the period between the end of World War II and the beginning of the new millennium was characterized by a rather uneven historical landscape with multiple disruptions, it is recorded as a period of unprecedented global economic growth. The world GDP per capita rose by 185% between 1950 and 2000, despite a 140% increase in the world population (Maddison (2003)).

On the policy side very important developments of the period resulted from the idea that economic and political cooperation and the newly created international institutions were instrumental to maintaining peace. The US lent support for the nascent European integration, and let the European countries decide how to share the Marshall Aid, enabling transition to functioning market economies (Eichengreen (2007)). The system of Bretton Woods institutions (the International Monetary Fund and the International Bank of Reconstruction and Development) was established. A series of conferences negotiated a setup for the International Trade Organization, which later was realized in its reduced form: the General Agreement on Tariffs and Trade. The development of the world trading system in the era of the GATT (and the early WTO) was a success story—although not without setbacks—that created a remarkably open world trading system.

The history of early international commercial treaties reveals that that the GATT's policies were largely drawn from the network of the non-discriminatory treaties formed before World War I. The key principles of reciprocity, nondiscrimination, and national

treatment were a long-standing part of the commercial relationships in the past. Part of the language was also borrowed from the Reciprocal Trade Agreements Act passed by the U.S. Congress in 1934. The flexibility that was built into the GATT, such as, for example, with respect to MFN exceptions for customs unions and free trade areas, or exemptions related to public health, national security or balance of payment crisis, were derived from the turbulent interwar experience. Similarly, the GATT has outlawed the quotas, which plagued and largely damaged international trade in the 1930s. Initially even the negotiating process within the GATT did not resemble a multilateral setup. The first GATT negotiating round resulted in 123 bilateral agreements which were generalized to the other member states according to the MFN principle (Findlay and O'Rourke (2007)). GATT's initial impact on the worldwide levels was also limited (Irwin (1995)).

The role of the GATT in the world trading system, however, cannot be understated. The most important innovation introduced by the GATT is its multilateral nature. In other words, the GATT provided the ground for negotiations to be held among several countries, balancing their diverse interests within an established regulatory framework. During the Kennedy Round of negotiations, the multilateral approach replaced a series of bilateral deals, and was finally established as a primary framework. Multilateralism was becoming more global with the number of GATT's members steadily growing—from 23 in 1949 to 127 by 1995, when the GATT was incorporated into the WTO (see Figure 2.4.5).

Another important innovation of the GATT is the dispute settlement system, which proved to be one of the most important functions of the WTO later on. While early bilateral commercial treaties included some sort of mechanisms to resolve disputes (largely regulated through consular relationships), the GATT included a fully institutionalized mechanism (Articles XXII and XXIII of GATT 1947). The principles and practices of the dispute settlement system have evolved over the half century and laid foundations for the decisions in court cases.

Over the long course of GATT's history, naturally, negative economic and political events influenced its functioning and efficiency. For example, in the early days of the GATT, trade barriers were much less relevant than the acute scarcity of convertible currencies that most countries experienced after the war. Tight exchange and import controls limiting market access created reluctance to move forward on the GATT's negotiating mandate. In these circumstances, the formation of the European Economic Community (EEC) in 1958 served as an impetus to continue trade negotiations. Both

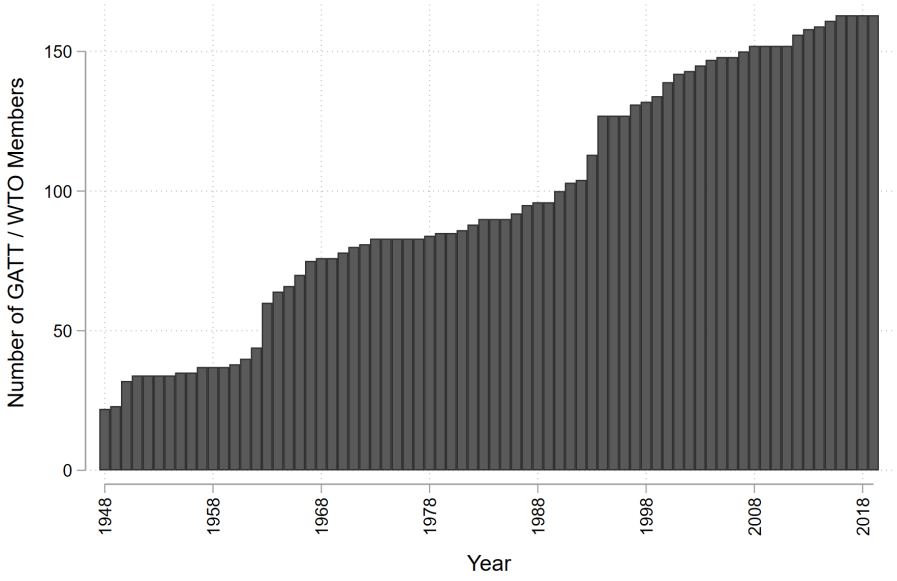

Figure 2.4.5: Number of GATT / WTO members, 1948-2019.

Note: The data comes from CEPII's Gravity Dataset.

the Dillon and the Kennedy Rounds were responses to the European integration, and the Kennedy Round in particular was a milestone in the postwar history of trade. Its consequence was a substantial reduction in the tariffs on manufactured goods in the more developed countries.

Another setback for the multilateral trading system came in the 1970s, when the fixed exchange rate regime collapsed, and a turbulent macroeconomic environment installed. In the policy domain, developed countries introduced measures that were both protectionist and discriminatory. The resulting multilateral agreement (the Multi-Fiber Arrangement) under the auspices of the GATT stipulated quantitative restrictions and discriminated against textile exports from a number of countries. A number of discriminatory trade practices, such as voluntary export restraints and orderly marketing agreements, lay outside of the GATT, and were convenient tools for protectionism. The early GATT dispute settlement system was addressing a number of issues that were brought to the negotiating table, but was some policies were outside of its reach, since the countries did not invoke the discussions (WTO (2017)).

After the creation of the GATT, which initially united 23 members under one agreement, the number of additional trade agreements signed was small (see Figure 2.4.1). Naturally, this is explained by the fact that these agreements were not

anymore only bilateral deals (compare the number of agreements to the number of country pairs covered by commercial treaties plotted in Figure 2.4.2). The early wave of regionalism was centered around Europe in the 1950s and 1960s, with the formation of the European Coal and Steel Community in 1951, followed by the European Economic Community (EEC) agreement in in 1957. Besides reinvigorating the GATT negotiations, as mentioned earlier, the EEC prompted the formation of the European Free Trade Association (EFTA) in 1957, for the countries which remained outside of the ECC deal. Moreover, the EEC sparked regional integration ideas among the developing countries in Africa, the Caribbean, Central and South America (most of which realized later on).

Regional commercial deals continued to be the focus of attention throughout the 1980s and the 1990s. The path towards a single market in Europe let to the transformation of the EEC into the European Community (EC) after the Maastricht Treaty was signed in 1993. The EC began a new push for bilateral agreements with the Central and Eastern European countries after the collapse of the Soviet Union, and also with the countries outside Europe. The US, frustrated with the delays in the Uruguay Round of multilateral negotiations, also embraced regionalism. The US-Canada trade agreement signed in 1988 evolved into the North American Free Trade Agreement (NAFTA) by 1994 with the inclusion of Mexico. Interestingly, the issues that the US was seeking to discuss in the multilateral setting (trade in services, investment, intellectual property rights and government procurement) manifested themselves in those regional talks, prior to being embraced in the Uruguay Round negotiations (WTO (2011)).

In other regions of the world regional integration was advancing too. In South America MERCOSUR (Southern Common Market) negotiations aimed at forming a customs union among Argentina, Brazil, Paraguay and Uruguay. In Africa a number of regional groupings were under formation[6]. The Association of Southeast Asian Nations (ASEAN) negotiated a free trade area uniting ten countries in the region. In general, the regional pattern of trade deals before 2000 was relatively evenly distributed among different regions (in proportion to the number of countries comprising each region), as shown in Table 2.4.6.

Regional commercial treaties in the second half of the 20th century were develop-

[6]The Common Market for Eastern and Southern Africa (COMESA), the East African Community (EAC), the Economic Community of West African States (ECOWAS) and the Southern African Development Community (SADC).

Table 2.4.6: Percentage of treaties concluded by region, 1948-2021

	1948-2000	2001-2021
Africa	7.5	2.5
Americas	30.1	18.8
Asia	8.9	16.6
Europe	27.8	16.8
Oceania	1.7	0.3
Intercontinental	24.0	44.9

Note: The data comes from DESTA database (see Appendix A for details of data cleaning).

ing alongside the major advances in the multilateral trading system. The Uruguay round of trade negotiations was successfully concluded in 1994, resulting in the creation of the WTO, which incorporated all trade and trade-related agreements within a single legal framework. Such issues as trade in services, investment, and intellectual property rights, have become an integral part of the regulatory landscape. The round also substantially improved the existing mechanism for dispute settlement, eased market access for services industries, revised subsidy rules, and outlawed voluntary export restraints, among other achievements.

Against this background, the regional deals were mostly focused on granting additional deeper market access to the signatories, rather than expanding the scope of the commercial negotiations. There were a total of 237 agreements concluded between 1949 and 2000 (95% of them can be characterised as the first-time treaties). Table 2.4.7 shows the percentage of treaties with different properties. Around a quarter of treaties were not included in the WTO list of trade agreements. Nearly one fifth of all treaties were formed as customs unions—a substantially higher proportion than in the following period. Given the universal and unconditional nature of the MFN treatment stipulated by the GATT, the type of the MFN provision included in these preferential trade deals is referred to as 'third party MFN'[7]. Around 40% of the treaties in the dataset provide some form of third-party MFN. The MFN treatment for services and investment, however, was a rare instance for the treaties concluded before year 2000 (7.6 and 5.9% respectively).

Focusing on the agreements included in the WTO list of preferential treaties con-

[7]In a preferential trade agreement, most-favoured nation (MFN) treatment requires parties to the agreement to grant to other parties the most favourable tariff and regulatory treatment granted to any third party.

Table 2.4.7: The percentage of treaties with a given property or including a given provision, 1948-2021

	1948-2000	2001-2021
Notified under GATT/WTO	76.4	84.3
Free trade areas	80.2	96.2
Customs unions	19.8	3.8
MFN goods	41.7	36.2
MFN services	7.6	39.2
MFN investment	5.9	40.6

Note: The data comes from DESTA database (see Appendix A for details of data cleaning).

firms the intuition that most of the treaties concluded before 2000 were rather conservative in scope. Figure 2.4.6 plots the number of agreements containing a certain provision under the mandate of the GATT/WTO. If included, these provisions would usually reaffirm the commitment to the GATT principles, and deepen the level of liberalization. A striking regularity in Figure 2.4.6 shows that these provisions were rarely a topic for trade agreements concluded during this period. Similarly, Figure 2.4.7 plots the number of treaties including a regulation outside of the scope of the GATT/WTO treaties. Again, very few treaties before year 2000 included provisions related to investment, movement of capital, competition policy, labor market, intellectual property, and data protection. Regulation of standards was also not at the top of the agenda: among all the treaties signed (both on and outside of the WTO list) few agreements before 2000 included standard harmonization as a general aim or ensured full or selective harmonization (see Figure 2.4.8).

The time period reviewed in this section is characterized by a continuous development of the multilateral rules framework and the system of regional integration. Therefore, it is hard to identify an exact date to finish the description of this period, and year 2000 clearly does not mark the end of the era of multilateralism. In fact, the modern trading system of the past 20 years has the multilateral system at its foundation, and continues to build on it. Rather, looking at the nature of the trade treaties signed in the 2000s (described in the next section), it becomes clear that there is a certain structural shift in the content of trade treaties.

Even with the blurred end date, commercial policy of the second half of the 20th century can be characterized by a number of stylized facts. The revolutionary innovation is the GATT's ability to establish a multilateral system of trade relations, and

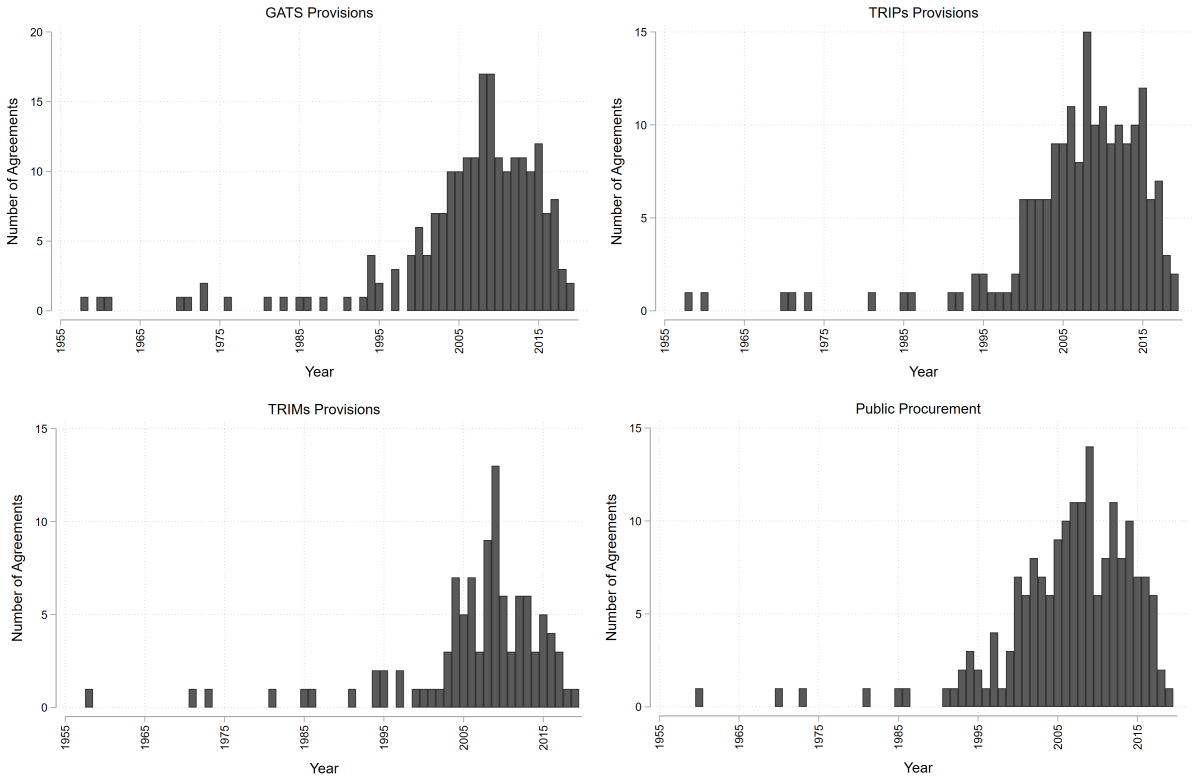

Figure 2.4.6: Number of agreements including a given provision under the current mandate of the WTO, 1958-2019

Note: The data comes from DTA database (see Appendix A for details of data cleaning).

provide an ever-expanding and unifying institution for regulating trade policy. Not without flaws, this system, however, managed to survive and advance through tumultuous decades. The multilateral trading system, consolidating the best practices and policies developed in the previous decades, proved to be highly successful in reducing trade barriers, while maintaining enough flexibility to accommodate a diverse set of interests. Another important addition that sustained the success of the system was the dispute settlement provision, which later resulted in a full-scale WTO court.

GATT's structure allowed regional integration initiatives and agreements to develop alongside the multilateral negotiations. While some of the most important trade deals were signed and evolved over this period, the number of treaties and their participants under MFN exceptions from the GATT (Article 24) was still relatively low. The majority of the treaties concentrated on deepening market access and

cutting tariffs, while including a conservative number of issues, in sharp contrast to the next period of the development of commercial treaties.

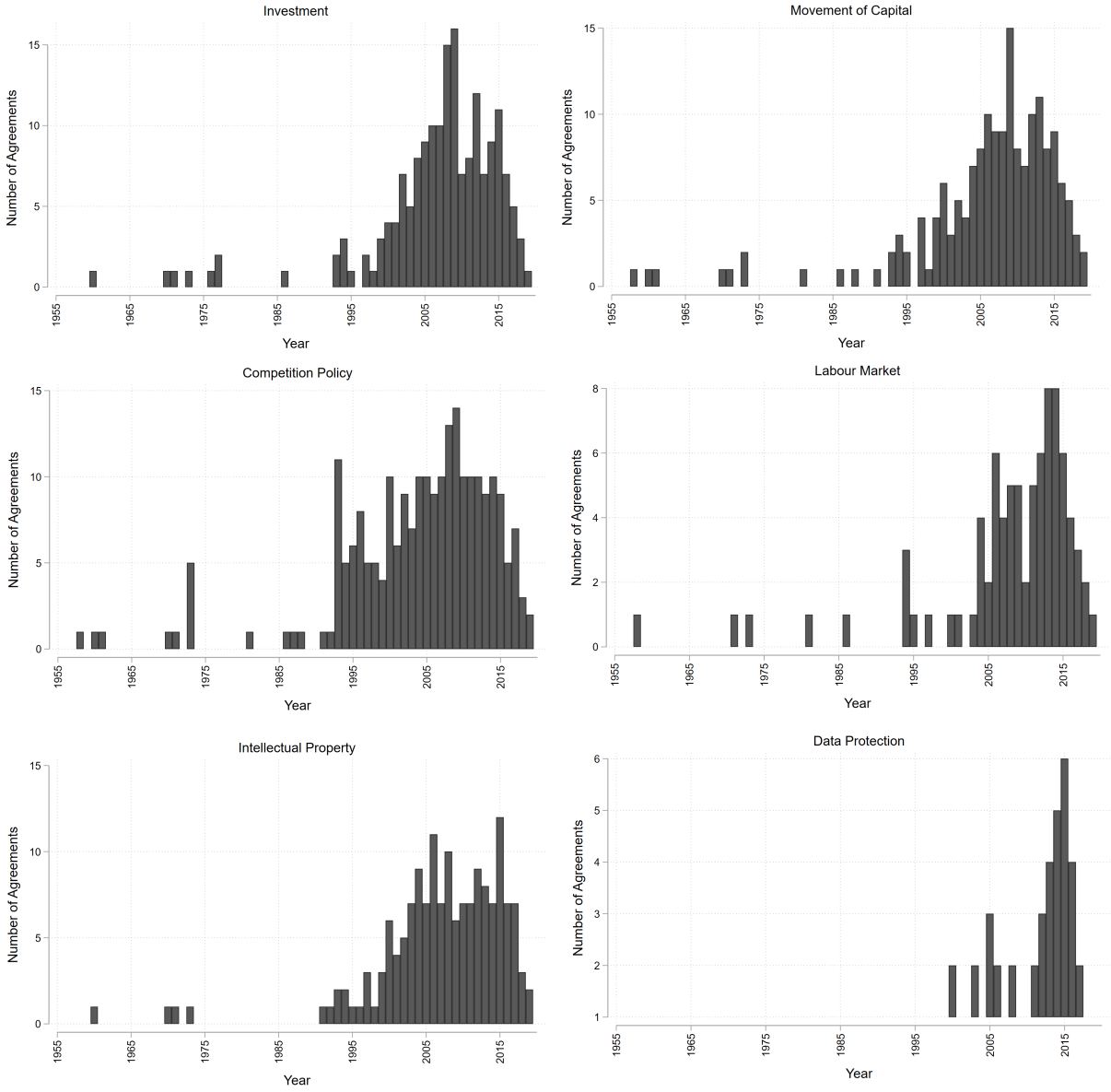

Figure 2.4.7: Number of agreements including a given provision outside of the current mandate of the WTO, 1958-2019

Note: The data comes from DTA database (see Appendix A for details of data cleaning).

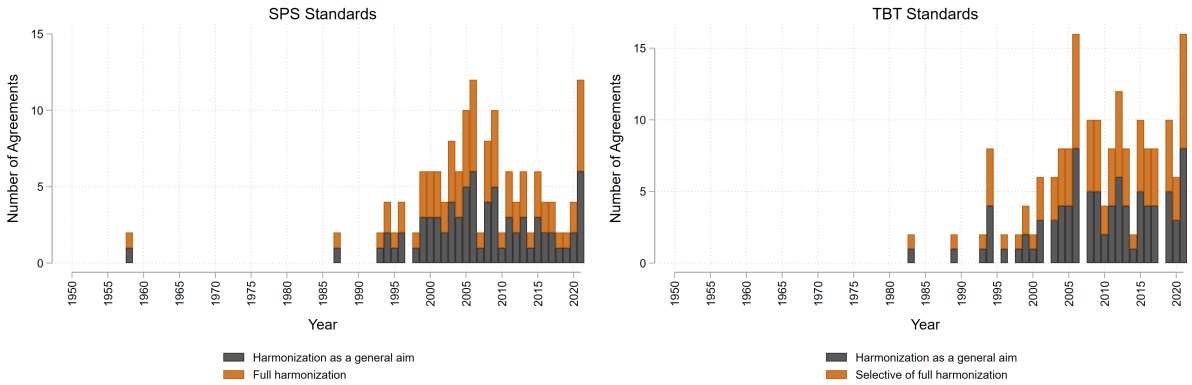

Figure 2.4.8: Number of agreements including standards harmonization provisions, 1949-2021.

Note: The data comes from DESTA database (see Appendix A for details of data cleaning).

2.4.4 The Modern Trading System (2000-2023)

One of the major structural shifts in the world economy from the beginning of the 2000s is attributed to the unprecedented growth of China and India. While economic growth and export expansion was a continuous process over the preceding decades, the early 2000s mark China's entry into the WTO, prompting discussions about the 'China shock' (Autor et al. (2013))—the negative impact of rising Chinese exports on manufacturing employment in the United States and Europe. The economic growth and trade expansion of the two large emerging economies lifted a large proportion of the population out of poverty, but also lead to environmental degradation and political tensions, which would manifest themselves in the increased rivalry between the US and China later on.

The global economic crisis and the following recession in 2008-2009 prompted a profound collapse in international trade and a led to rising protectionist moods. Once again, trade policy has become a tool to respond to the shocks of a different nature. And as before, protectionism did not play a role in explaining the trade collapse itself (Bems et al. (2012)), but most certainly disrupted trade in particular sectors and industries.

During this period, the number of trade agreements formed remained persistently high, comparing in the historical perspective (see Figure 2.4.1), and the number of participating countries was indeed exploding (see Figure 2.4.2). While most of the trade deals were regional in nature in the preceding decades, now almost half of all

trade agreements have become intercontinental (see Table 2.4.6). The tendency to expand the global membership and geographical scope of the treaties also manifests itself in a sharp decline in the proportion of customs unions formation: now 96.2% of all newly concluded trade treaties are free trade areas (see Table 2.4.7).

From the treaty content point of view, some major innovations are becoming evident: the generation of the so-called 'new trade trade agreements' started to emerge from the beginning of the 2000s. Grossman et al. (2021) relate the development of this new generation to the successes of the multilateral trade liberalization, and the shift of attention to addressing the numerous non-tariff barriers impeding trade. In addition, soon after the start of the Doha round of multilateral trade negotiations, it became clear that the ambition was lost, and the expanded WTO membership ran into problems with balancing its diverse commercial interests, prompting countries to conclude preferential deals (Bhagwati (2008)).

One distinct characteristic of this generation of trade agreements is their increased length. Figure 2.4.9 plots the average number of words per agreement, which went from 3.4 thousand in 1949-2000 to 12.5 thousand in 2001-2017—and almost 4-fold increase in length. This tendency is also reflected in the average number of chapters each agreement includes (see Figure 2.4.3). While the mean number of chapters was around 7 in the second half of the 20th century, it doubled for trade agreements signed after year 2000. Naturally, these increases in length of a typical treaty reflect the increased scope and coverage.

The content expansion of the new treaties was happening across the board: both deepening the cooperation on the existing issues and adding new ones. Figure 2.4.6 shows the number of agreements including a given provision under the current WTO mandate (relating to trade in services, intellectual property rights protection, investment and public procurement). Similarly, Figure 2.4.7 plots the number of treaties containing issues outside of the WTO's mandate, such as certain aspects of investment regulation and intellectual property, movement of capital, competition policy, labour market regulation, and data protection. Besides the examples provided in the figure, a similar expansion can be found for other issues, such as environmental protection, trade defense instruments, regulatory cooperation, dispute settlement, global value chains, etc. Both figures show a striking feature: the number of trade agreements including these types of issues exploded somewhat around year 2000.

The novelty of the trade agreements signed after the 2000s also manifests in the mechanisms described in Grossman et al. (2021), relating to standards harmoniza-

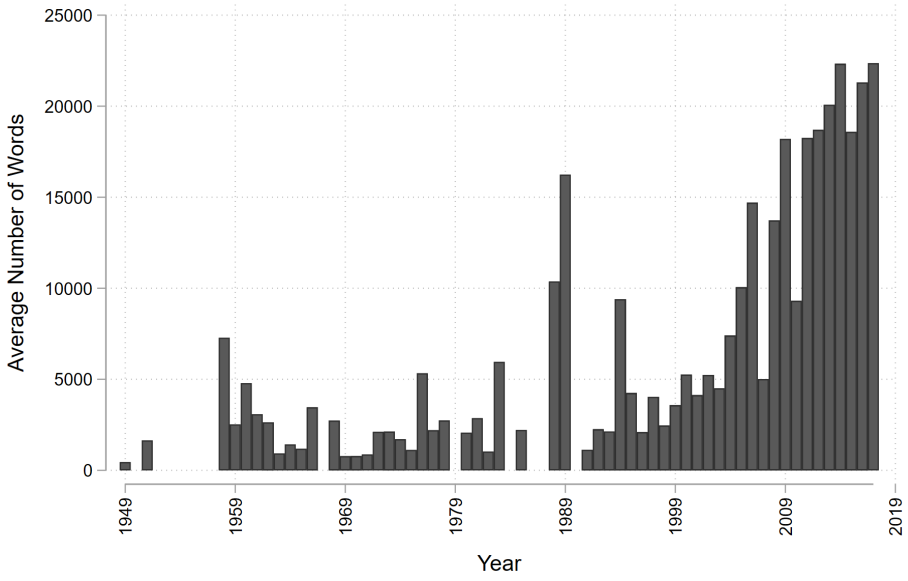

Figure 2.4.9: Average number of words per agreement, 1949-2017

Note: The data comes from ToTA database (see Appendix A for details of data cleaning).

tion. Figure 2.4.8 plots the number of agreements including provisions on standards harmonization of sanitary and phytosanitary regulations (left) or technical regulations (right). The number of agreements including standards harmonization as a general aim, or actually providing full standards harmonization, is almost negligible before the beginning of 2000s, while almost one in tree agreements signed after year 2000 include these types of provisions.

Finally, another feature worth noting about the evolution of trade agreements in this period is related to the fact that the length of the text and the number of provisions included do not necessarily reflect the extent to which these treaty provisions are actually implemented. As mentioned in Section 3.2 and Appendix A, the wording of the texts matters: including some provisions does not automatically lead to compliance or enforcement of sanctions in case members do not act in accordance with the provisions.

Figure 2.4.10 plots the proportion of legally enforceable provisions, as as well as the provisions formally excluded from the dispute settlement, across different types of issues included in the agreements. Naturally, the proportion of legally enforceable and 'punishable' provisions is high for issues under the current WTO mandate (left panel). At the same time, roughly from the beginning of the 2000s there is a tendency

to exclude a number of provisions on these traditional issues from the dispute settlement. The right panel of Figure 2.4.10 plots an even more striking pattern for issues outside of the WTO mandate: very few of them are mentioned in conjunction with the dispute settlement provisions. The new generation of trade agreements, however, started to explicitly exclude these provisions from the dispute settlement procedures.

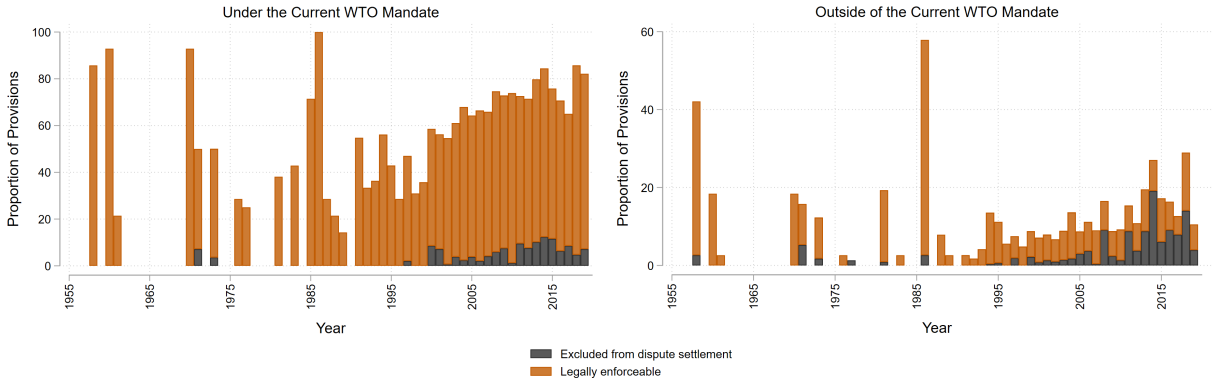

Figure 2.4.10: Proportion of provisions excluded from dispute settlement and legally enforceable provisions, 1958-2019

Note: The data comes from DTA database (see Appendix A for details of data cleaning).

A similar pattern emerges when using words and their semantic meanings to measure the extent of legal enforceability in a given treaty. The measure plotted in Figure 2.4.11 is the average ratio of strong to weak words (see Appendix A for definitions), normalized by the total amount of words per agreement (Baker et al. (2016)). This ratio is declining since the beginning of 2000s, reflecting the decreasing number of issues that contain provisions which are not legally enforceable.

The system of commercial relationships in 2000s can be broadly characterized by a decline in multilateral negotiations, and the development of a dense network of treaties governing around half of the total world trade. These treaties spread widely in terms of their geographical coverage, while simultaneously expanded in their scope. Modern commercial treaties include a large number of issues, including those outside of the current WTO mandate. At the same time, the extent to which these treaties can credibly enhance cooperation is unclear: many of the issues discussed in these treaties do not represent a legally enforceable provision.

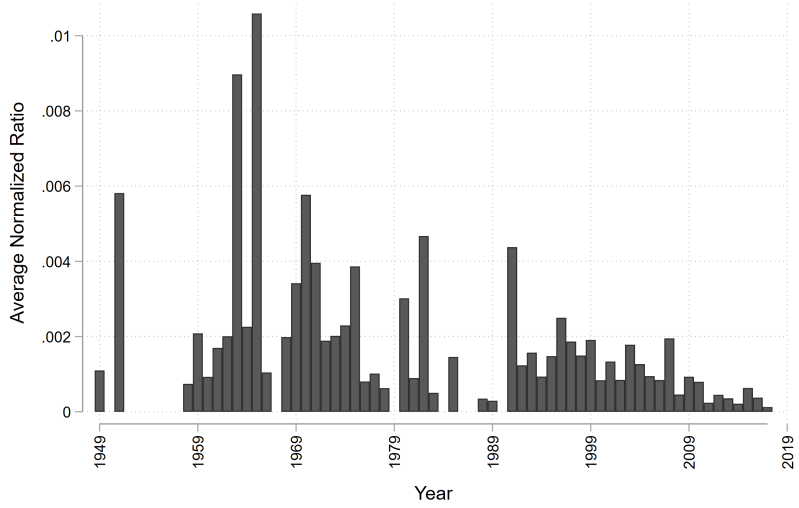

Figure 2.4.11: Average value of normalized ratio of strong words to weak words by year, 1949-2017.

Note: The data comes from ToTA database (see Appendix A for details of data cleaning) and measure construction.

2.5 Conclusions and the Future of Trade Treaties

Looking at the development of the content of the commercial treaties from the perspective of the long run reveals an important regularity: many of the policy innovations are introduced in response to the major world events and shocks. Trade policy often follows the geopolitical and economic agenda. The expansion of market access provisions in response to economic expansion of the 1860s, the inclusion of exemptions and exceptions following the turbulent interwar period, the creation of a functioning multilateral trading system following World War II, and the spread of the new generation of trade agreement following the slowdown of the WTO negotiation function—all reflect the complimentary role for trade policy and trade agreements. The future commercial policies and treaties are thus likely going to be a response to the events and shocks we are living through today.

Naturally, identifying the future trends for trade treaties is a speculative exercise. There are, however, important trends in the global economic agenda that suggest potential policy responses. "Deglobalization" has been a central concept in the discussions about the current trends in the global economy. While not yet showing in the data (Goldberg (2019), Antràs (2020)), a series of recent events and policy narra-

tives have suggested a shift in the way globalization will likely proceed in the future (Goldberg and Reed (2023)).

The second half of the 2010s marks a shift in commercial policy, with the growing fears about the impact of China's import competition and the refugee flows threatening the economic stability of more developed countries. The rising negative public sentiment and the government actions resulted in a series of trade wars between China and the US. This period is also marked by the UK's vote to terminate its membership in the European Union. The arguments about the threats to national security and sovereignty have become widespread. To make matters worse, the only functioning institution to resolve commercial disputes—the WTO's judicial branch—was paralyzed by the US's block on the appointment of the Appellate Body judges. The COVID-19 pandemic in 2020 exacerbated the concerns about the lack of resilience in supply chains, and Russia's invasion of Ukraine in 2022 led to wide-spread appeals to 'decouple' and creates a more fragmented system relying on political alliances ("friendshoring").

The question on whether these events are going to lead to a structural shift in the way countries conduct trade policy, or this is going to be a temporary setback (akin to those in 1870s and the 1970s), remains open. This far, the world remains as globalized as it has ever been in history, and there are no dramatic changes in the global trade yet noticeable in the short run. There was a major spike in the number of trade treaties signed in 2021 due to Brexit and the UK's re-signing of trade treaties. According to the WTO's early announcement database, there are currently 35 treaties under negotiations, major part of which are, however, pending for more than five years. One of the most important trade deals was concluded by China, Japan and Korea, Australia and New Zealand, together with 10 ASEAN countries in 2020, in the hopes to boost trade cooperation in the region.

The number of trade agreements signed has already slowed down in the beginning of the 2020s (not counting the UK's trade deals), and this deceleration is natural, given the extensive treaty network that already formed. The coverage of these existing treaties is also quite comprehensive in terms of the issues covered. Yet existing regulation is severely lagging behind the development of the new technologies: the development of electronic commerce, the crucial role of digitization and data, and the rapid development of artificial intelligence—all pose challenges in defining the global commercial rules. Due to the unprecedented novelty of these issues and the extraordinary pace of their development, some time may pass until countries define

a set of policies to regulate these areas internally, and even more time is necessary for these policies to be agreed internationally. New negotiations could then promote the deepening of the existing treaties, or even their renegotiation and updating. What we do know for sure, however, is that there is room for commercial policy to address the mounting challenges, as this was the case in preceding historical periods.

Appendix

2.A Data Construction

Consolidated Treaty Series

Consolidated Treaty Series is a collection of multilateral and bilateral treaties in facsimile concluded between 1648 and 1918, published between 1969 and 1980, under the editorship of Professor Clive Parry. The 243 volumes (circa 500 pages each) compile all available international treaties, including commercial agreements, peace treaties, colonial agreements, telegraph conventions, etc. The volumes are not available in any digital form, and were accesses at the library of Universitat Pompeu Fabra in April and May 2023.

The volumes do not contain a separate index for commercial treaties (unlike, for example, colonial treaties), and each volume was skimmed through for the presence of commercial treaties. I would then access the text of every treaty directly, read the treaty, and codify the textual information into a series of data points.

Treaties were considered to be 'commercial' if they had a mention of the word 'commerce' or 'trade' in their title. Many treaty names included different combinations of the following words: amity, friendship, alliance, commerce, navigation, establishment, customs, and peace. Some examples of common treaty names are:

- Commercial convention
- Treaty of commerce and navigation
- Treaty of peace commerce and navigation
- Treaty of amity, navigation and commerce
- Convention of commerce and navigation
- Treaty of friendship, navigation and commerce
- Treaty of commerce

There are some some types of treaties that did mention commerce, but are not included in the dataset. These would include all types of supplements, amendments, additions, explanatory notes, and provisional agreements. In addition, declarations were generally excluded, since they proclaim general goals and are usually not broken up by chapters. Moreover, all the treaties related to the formation of Zollverein

are also not included in the dataset, as this process is considered to be different in nature, compared to other types of commercial treaties. The agreements of the parts of Zollverein, before their unification, with other European countries, is considered as part of the dataset. Similarly, the agreements of the different parts of contemporary Italy with other European countries is included in the dataset.

The final database includes 888 treaties, spanning the years from 1815 to 1919. These treaty texts are located across volumes 65 to 225 of the *Consolidated Treaty Series*. Each treaty is typically presented in one or two languages (sometimes there are translations in more than two languages). There are 388 treaties available in English, 446 in French, 32 in Spanish, 12 in Italian, and 10 in German. Due to translation limits, I could code all the treaties except for the 10 treaties in German.

The dataset on the historical treaties (also including the treaties coded from the dataset described next—the League of Nations Treaty Series) includes the following variables:

- **name**: an abbreviated name of the treaty (the decoding of each treaty name is provided in the same Excel file).
- **source**: the two sources for historical treaties ("CTS" refers to *Consolidated Treaty Series*, "LNTS" refers to *League of Nations Treaty Series*.
- **volume**: the volume in a given treaty source, in which the treaty text locates.
- **page**: the page in a given treaty source and volume, in which the treaty text starts.
- **first**: an indicator variable which takes a value 1 if the treaty is the first agreement concluded between a given country pair since 1815; and takes value 0 if a given country pair had concluded an agreement before the year of signature and after 1815.
- **signed**: the year of the signature of the agreement
- **year**: the year of ratification of the agreement, if explicitly stated in the description or preamble of the treaty, and is different from the year of signature.
- **party 1** and **party2**: the names of the sovereign units signing the agreement, as provided by the *Consolidated Treaty Series* treaty description. The names of

the countries are in alphabetical order: for example a treaty of Belgium and the Netherlands would always have Belgium as **party1** and Netherlands as **party2**.

- **reg1** and **reg2**: broad geographical regions for the parties of the treaty, including Africa, Asia, Europe, North America and South America.
- **lang1** and **lang2**: the languages in which treaties are provided
- **mfn_duty**: an indicator variable taking a value 1 if the treaty has a provision on MFN treatment related to the payment of duties and import taxes (see example language in Appendix B).
- **mfn_port**: an indicator variable taking a value 1 if the treaty has a provision on MFN treatment of vessels in the port (see example language in Appendix B).
- **nt**: an indicator variable taking a value 1 if the treaty has a national treatment provision related to commerce (see example language in Appendix B).
- **ma**: an indicator variable taking a value 1 if the treaty has a market access provision of some form, i.e. if it mentions the sizes of import duties on all goods, or provides a list of goods with the duties explicitly specified for those goods (see example language in Appendix B).
- **ntb**: an indicator variable taking a value 1 if the treaty has exceptions from the freedom of commerce, usually in the form of prohibitions on imports and exports with the purposes of defending public health, national security, monopolies, etc. (see example language in Appendix B).
- **numchap**: numer of chapters in the main treaty text (not counting additional protocols and annexes).
- **expiry**: the number of years that the treaty is agreed to be in force from the date of ratification.
- **terminated**: the year in which a treaty was terminated or denounced by one of the parties, in case such information is explicitly provided in the treaty description.
- **prolonged**: the year in which a treaty was prolonged or extended, in case such information is explicitly provided in the treaty description.

- **standalone**: an indicator variable that takes value 1 if the treaty's major purpose is to regulate commerce and navigation. If a treaty includes provisions on the treatment of natural persons or ships of war, but the majority of the treaty is devoted to commercial issues, such a treaty will still be characterized as stand-alone. If a major part of the treaty is related to, for example, colonial relationships, military alliances or immigration, then such treaty would not be considered a stand-alone.

- **note**: relates to a treaty having a mention of specific goods or having a tariff schedule.

- **quota**: indicates if a given agreement explicitly specified quotas.

League of Nations Treaty Series

League of Nations Treaty Series collection is structured in the same was as the *Consolidated Treaty Series*. The 250 volumes include all agreements concluded between countries between years 1920 and 1946 (a total of 4,834 agreements). I use the same criteria to identify treaties related to commerce and use the same coding as for the *Consolidated Treaty Series*. The texts were accessed online in May 2023.

Handbook of Commercial Treaties

Handbook of Commercial Treaties: Digests of Commercial Treaties, Conventions, and Other Agreements of Commercial Interest Between All Nations is a book describing treaties related to commerce, written by Herman Gustav Adolph Brauer. The book includes the summaries of treaties concluded between 1654 and 1922.

The collection of treaties covered is extensive: there are a total of 3121 treaties mentioned. Not all of them, however, are provided with a digest. 3035 of these treaties are signed after year 1815. There is a variety of forms that the treaties are represented in: agreements, conventions, declarations, acts, articles, exchanges of notes, notes, protocols, as well as additional and supplementary treaties. They also cover a wide variety of topics, which I code into 6 broad categories:

1. **Territorial treaties**: treaties related to accessions of territories, annexation, colonial application, adhesion to colonies, traffic through certain territories, etc.

2. **Additional treaties**: treaties correcting errors, interpreting other treaties, modifying existing agreements, providing additional articles, amending the existing treaties, etc.

3. **Commercial treaties**: these are the treaties that largely use the same words in their titles, as the previous two datasets, i.e. treaties focusing broadly on commerce and navigation.

4. **Separate commercial issues**: treaties regarding issues, such as trade in cattle, coasting trade, application of duties, customs procedures, rules of fisheries, navigation, trade marks, etc.

5. **General treaties**: agreements related to consular affairs, economic unions, peace, etc.

6. **Other**: all the treaties which cannot be identified as belonging to the categories outlines above.

The final dataset includes the following variables:

- **num**: number of the treaty in the book's classification.
- **party1** and **party2**: parties to the treaties as provided by the book.
- **type**: type of the document, such as convention, declaration, agreement, etc.
- **name**: the extended name of the treaty, typically describing the object of the treaty.
- **category**: one of the six categories identified above.
- **signed**: year of the signature of the agreement
- **ratification**: year of treaty's ratification, if explicitly provided.
- **page**: the page of the summary of the corresponding treaty in the book.

Figure X below shows the total number of treaties listed in the Handbook, by year. The figure confirms the earlier finding that there was an increase of treaties singed in the beginning of the 20th century, and during the interwar period. **Figure Y**

breaks down the number of treaties by category and plots the commercial treaties, the treaties related to specific issues in commercial policy, as well as additional treaties.

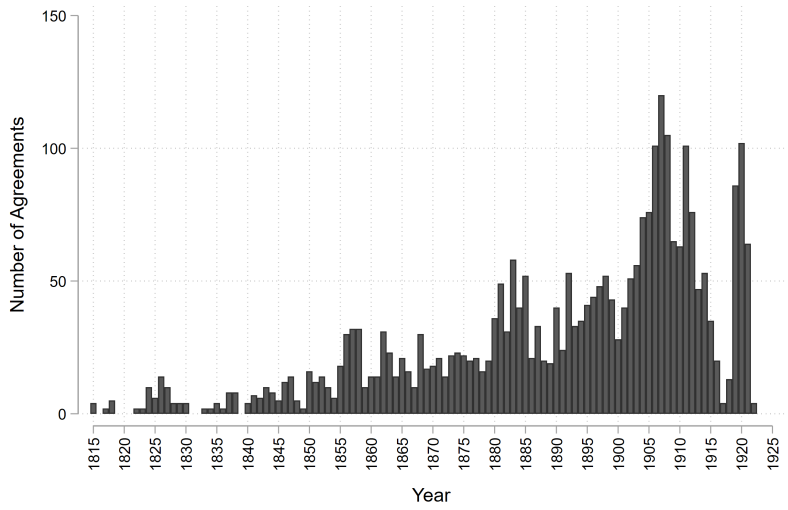

Figure 2.A.1: Total number of treaties listed in the *Handbook of Commercial Treaties*, 1815-1922.

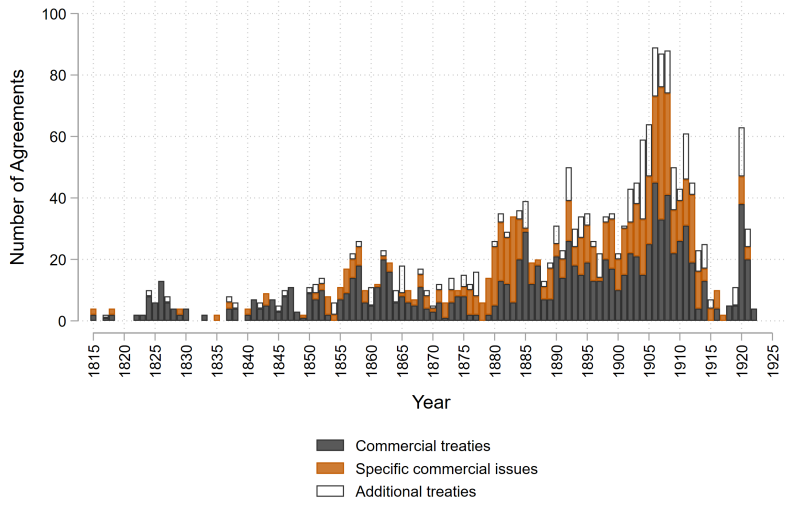

Figure 2.A.2: Number of treaties listed in the *Handbook of Commercial Treaties*, by category, 1815-1922.

Design of Trade Agreements Database

Design of Trade Agreements Database **(Dür et al. (2014))** includes 781 agreements signed from 1948 to 2021. The cleaning protocol used deletes the superseding treaties, amendments and additional protocols, as well as partial scope treaties and framework agreements[8]. The final sample includes 556 treaties. DESTA includes an extensive list of variables indicating the presence or absence of certain provisions in the treaty texts, and a detailed codebook is provided on the dataset's website.

Deep Trade Agreements Database

World Bank's *Deep Trade Agreements (DTA) Database* **(Hofmann et al. (2017))** includes 318 agreements signed from 1958 to 2019. For each agreement the dataset codes 52 commercial provisions, dividing them into items under the current WTO mandate, and provisions outside of the WTO. Other features as well as detailed description of the database content can be found on the website.

Texts of Trade Agreements Database

UNCTAD's *Text of Trade Agreements (ToTA)* database is a collection of PTA texts coded in unbalanced `XML` format. The database contains 450 texts with metadata on participating signatories, the type of agreement, date of signature, date of entry into force, status of the agreement as of year 2017, its composition, region, language and text source, which I extract into a matrix. In addition, I extract information on the number of chapters in a given agreement.

The cleaning procedure deletes agreements in languages other than English (24 Spanish and two in French); as well as the agreements under non-reciprocal concessions, with coding mistakes in the XML files, and amendments: 1st Convention of Lome, 2nd Convention of Lome, 3rd Convention of Lome, Generalized System of Preferences, Yaounde Convention I, Yaounde Convention II, EC-Syria, Arusha Agreement, EU-Overseas Territories, and Croatia-Serbia-Montenegro Agreement.

The text corpus is cleaned using text mining techniques, stripped of punctuation and the so-called "stopwords" that do not carry any semantic meaning, and transformed into a Document-Term Matrix (DTM) with each cell representing a particular

[8]Partial scope agreements liberalize only some part of trade, while framework agreements have no specific trade liberalization provisions.

word count per agreement.

Trade texts contain words (mostly verbs) that can proxy for a degree to which signatories are obliged to implement the measures and policies according to a given agreement. These words can indicate "high" and "low" levels of legal enforceability. In order to identify which words have can be classified into the two commitment categories, information on the intended semantic meanings was drawn from trade law practice. A trade lawyer identified which words, generally defined, can potentially lead to a litigation or a dispute using international arbitrage, or trigger the process of suspension of parties' concessions. The words that correspond to "strong" enforcement (those which have a higher binding power in trade law) include: "shall," "commit," "require," "compliance" (often represented in a bi-gram "non-compliance"), "penalty," "accord," "adopt," "ensure." The second group of words, representing "weak" enforcement includes words such as "may," "endeavor," "aim," "cooperate," "dialogue," "possible" (often used as a part of the world collocation "to the extent possible").

The following example illustrates the differences transmitted by semantic structures. In the chapters of the agreement related to early release of goods from customs, the comprehensive agreement between US and Korea (KORUS) states: "Each Party *shall* adopt and maintain procedures providing for the expeditious release of goods admitted under this Article." On the same issue, Korea's agreement with Vietnam reads: "Each Party shall *endeavor* to adopt and maintain procedures providing for the expeditious release of goods admitted under this Article." In the former agreement, the obligation to release goods is prescriptive, while in the latter it is indicative. In other words, a signatory that "endeavors" to adopt and maintain a trade measure does not necessarily has to actually implement it.

Following the initial identification of word groups, I construct an index which serves as a proxy for the legal enforceability of a given agreement.

The constructed variable takes the ratio of the total sum of "strong" words indexed from $\{1, ..., i\}$ to the sum of "weak" words indexed from $\{1, ..., j\}$ in a given document d, normalizing it by the total word count per agreement agreement, indexed by $\{1, ..., w\}$:

$$\text{Legal Enforceability Index}_d = \frac{\sum_i \text{Strong}_{i,d}}{\sum_j \text{Weak}_{j,d}} \times \frac{1}{\sum_w \text{Words}_{w,d}}$$

2.B Examples of Commercial Provisions Texts

Reciprocity. "There shall be reciprocal liberty of commerce between all the dominios and possessions of the two High Contracting Parties; and the subjects of each of them shall, throughout the whole extent of the territories and possessions of the other, enjoy the same rights, privileges, liberties, favours, immunities, and exemptions, in matters of commerce and navigation, which are or may be enjoyed by native subjects."

Treaty of Commerce between Belgium and Great Britain, 23 July 1862 (volume 126, page 141, Consolidates Treaty Series), Article I

MFN (general). "If in time to come and during the period that the present Treaty shall remain in force, one of the High Contracting Parties shall grant any special favour to another nation in point of navigation and of commerce, that favour shall be immediately acquired by the other Contracting Party, without any charge or impediment, if it was granted gratuitously to the other nation, or on equivalent conditions if the grant was for a considerations and correlative."

Treaty of Commerce and Navigation between Greece and Tuscany, 22 April 1856 (volume 115, page 17, Consolidates Treaty Series), Article XXVII

MFN (natural persons). "Each of the High Contracting Parties assures to the subjects of the other the right of traveling and residing freely in its dominions, saving the police precautions adopted, or which may be adopted, by the Government of each country, and applied to the subjects of the most favoured nation; of occupying houses and warehouses, and disposing their personal property of whatever kind or denomination acquired by sale, exchange, donation, will or in any other manner, without offering them the least impediment; they shall not, under any pretext, be obliged to pay more imposts or contributions than are paid, or may be paid, besides the natives, by the subjects of the most favoured nation; [...]"

Treaty of Commerce and Navigation and in Relation to Consuls between Spain and the Two Sicilies, 16 March 1856 (volume 114, page 357, Consolidates Treaty Series), Article III

MFN (commercial issues). "The Subjects of the two Sovereigns, respectively, shall not pay in the ports, harbours, roads, cities, towns, or places whatsoever in either Kingdom, any other higher duties, taxes, or imposts, under whatsoever names des-

ignated or included, than those that are paid by the subjects of the most favoured Nation; and the subjects of each of the High Contracting Parties shall enjoy the same right, privileges, liberties favours, immunities, and exemptions, in matters of commerce and navigation, that are granted, or may hereafter be granted, in either Kingdom, to the subjects of the most favoured Nation."

Treaty of Commerce and Navigation between Great Britain and the Netherlands, 27 October 1837 (volume 87, page 95, Consolidates Treaty Series), Article I

MFN (port). "The vessels and subjects of the High Contracting Parties shall, by the present Treaty, enjoy reciprocally all the advantages, immunities, and privileges, within the Ports of their respective States and Possessions, which now enjoyed by the Navigation and Commerce of the most favoured Nations."

Treaty of Commerce and Navigation between Austria and Great Britain, 3 July 1838 (volume 88, page 13, Consolidates Treaty Series), Article XI

MFN (duty). "No duty of Customs or other impost shall be charged upon any goods produce of one country, upon importations, by sea or by land, from such country into the other, higher than the duty or impost charged upon goods of the same kind, the produce of , or imported from, any other country."

Treaty of Commerce and Navigation between Great Britain and the Netherlands, 27 October 1837 (volume 87, page 95, Consolidates Treaty Series), Article I

"Neither of the two High Contracting Parties shall impose upon goods the produce or manufacture of the other party, other or higher duties of importation than such as are or may be imposed upon the same goods the produce of any other foreign country.

Each of the two parties engages to extend to the other any favour or privilege, or reduction in the tariff duties of importation or exportation, on articles mentioned, or not mentioned, in the present Treaty, which either of them may grant to any third Power. They engage, moreover, not to establish against each other any duty or prohibition of importation or exportation, which shall not, at the same time, be applicable to all other nations."

Treaty of Commerce between Belgium and Great Britain, 23 July 1862 (volume 126, page 141, Consolidates Treaty Series), Article XIV

Unconditional MFN. "The High Contracting Parties agree that in all matters relating to commerce and industry any privilege, favor, or immunity whatever which either High Contracting Party has actually granted or may hereafter grant to any other foreign State shall be extended immediately and unconditionally to the subjects or citizens of the other Contracting Party; it being their intention that the commerce and industry of each country shall be placed, in all respects, by the other on the footing of the most favored nation."

Treaty of Commerce between Great Britain and Bolivia, 1 August 1911 (volume 214, page 181, Consolidates Treaty Series), Article V

Conditional MFN. "Except as otherwise expressly provided in this Treaty, the High Contracting Parties agree that, in all that concerns commerce and navigation, any privilege, favor, or immunity which either Contracting Party has actually granted, or may hereafter grant, to the citizens or subjects of any other State shall be extended to the citizens or subjects of the other Contracting Party gratuitously, if the concession in favor of that other State shall have been gratuitous, and on the same or equivalent conditions, if the concession shall have been conditional."

Treaty of Commerce between United States and Japan, 21 February 1911 (volume 213, page 98, Consolidates Treaty Series), Article XIV

National Treatment. "No duties of tonnage, harbour, lighthouses, pilotage, quarantine, or other similar corresponding duties, of whatever nature or under whatever denomination, shall be imposed in either country upon the vessels of the other, in respect of voyages between 2 countries, if laden, or in respect of any voyage, if in ballast, which shall not be equally imposed, in the like cases, on national vessels; and in neither country shall any duty, charge, restriction, or prohibition be imposed upon, not any drawback, bounty, or allowance be withheld from, any goods imported from, or exported to the other country, in the vessels of that other country, which shall not be equally imposed or withheld from such goods, when so imported or exported in national vessels."

Treaty of Commerce and Navigation between Great Britain and the Netherlands, 27 October 1837 (volume 87, page 95, Consolidates Treaty Series), Article I

Market Access. "It is further agreed that if sea salt refined in Belgium should obtain a deduction of more than 7 per cent from the general duty of excise, British salt refined

in Belgium shall enjoy, at the same moment, a deduction from the excise which shall not be inferior by more than 7 per cent to the deduction granted to sea salt."

Treaty of Commerce between Belgium and Great Britain, 23 July 1862 (volume 126, page 141, Consolidates Treaty Series), Article XIV

"His Majesty the King of the Two Sicilies promises, moreover, that while this Treaty is in force, all the merchandize and productions of the Kingdom of Spain, of the Peninsula and the adjacent islands, as well as of her possessions beyond sea, imported into his dominions in Spanish or Sicilian vessels, shall enjoy a reduction of 10 per cent from the dues established in the Customs' tariff, [...]"

Treaty of Commerce and Navigation and in Relation to Consuls between Spain and the Two Sicilies, 16 March 1856 (volume 114, page 357, Consolidates Treaty Series), Article X

Tariff schedules.

TARIFF (A).—Annexed to the Treaty of Commerce between the German Empire and Italy.

DUTIES on Entry into Germany.

Denomination of the Merchandize.	Duty per 100 kilog.
	Marks.
Sulphur, raw or refined	Free.
Tartar, raw or refined	„
Liquorice juice	„
Fresh grapes	10
Poultry of any sort, not living	12
Fresh oranges, citrons, lemons†	4
Fresh Seville oranges†	4
Fresh pomegranates†	4
Fresh dates or almonds	4
Dried dates or almonds	10
Dried Seville oranges or pomegranates	10
Olives	30
Rice, in the husk or cleaned	4
Edible oil, in bottles or flasks	10
Olive oil, in casks	4

TARIFF (B).—Annexed to the Treaty of Commerce between the German Empire and Italy.

DUTIES on Entry into Italy.

Denomination of the Merchandize.	Per—	Duty.
		Lire.
Alkaloids—		
(*a.*) Salts of quinine	Kilog.	5
(*b.*) Not named, and their salts	„	5
Zinc—		
(*a.*) In blocks and lumps	Quintal	1
(*b.*) In plates and sheets	„	4
(*c.*) Manufactured in other forms, not gilt	„	12
(*d.*) Manufactured in other forms, gilt	„	58
Optical and mathematical instruments, instruments of precision and of investigation, chemical, physical, surgical instruments, &c.	„	30
Hops		Free.

Treaty of Commerce and Navigation between Germany and Italy, 16 March 1883 (volume 162, page 47, Consolidates Treaty Series), Annex

Articles.	Quantities.	Duty.			
		T.	M.	C.	C.
Gambier	Per 100 catties	0	1	5	0
Gamboge..	,,	1	0	0	0
Ginseng, American, crude	,,	6	0	0	0
,, ,, clarified	,,	8	0	0	0
Glass, window	box of 100 sq. ft.	0	1	5	0
Glue	100 catties	0	1	5	0
Gold thread, real	catty	1	6	0	0
,, imitation	,,	0	0	3	0
Gum Benjamin	100 catties	0	6	0	0
,, ,, oil of	,,	0	6	0	0
,, dragon's blood	,,	0	4	5	0
,, myrrh	,,	0	4	5	0
,, olibanum	,,	0	4	5	0
Hides, Buffalo and Cow	,,	0	5	0	0
,, Rhinoceros	,,	0	4	2	0
Horns, Buffalo	,,	0	2	5	0
,, Deer	,,	0	2	5	0
,, Rhinoceros	,,	2	0	0	0
Indigo, liquid	,,	0	1	8	0
Isinglass	,,	0	6	5	0
Lacquered-ware	,,	1	0	0	0
Leather	,,	0	4	2	0
Linen, fine, as Irish or Scotch, not exceeding 50 yds. long	piece	0	5	0	0
,, coarse, as linen and cotton, or silk and linen mixtures, not exceeding 50 yds. long..	,,	0	2	0	0
Lucraban seed	100 catties	0	0	3	5
Mace	,,	1	0	0	0
Mangrove bark	,,	0	0	3	0
Metals:					
Copper, manufactured, as in sheets, rods, nails	,,	1	5	0	0
,, unmanufactured as in slabs ..	,,	1	0	0	0
,, yellow metal, sheathing, and nails..	,,	0	9	0	0
,, Japan	,,	0	6	0	0
Iron, manufactured, as in sheets, rods, bars, hoops	,,	0	1	2	5
,, unmanufactured, as in pigs	,,	0	0	7	5
,, kentledge	,,	0	0	1	0
,, wire	,,	0	2	5	0
Lead, in pigs	,,	0	2	5	0
,, in sheets	,,	0	5	5	0
Quicksilver	,,	2	0	0	0
Spelter (saleable only under regulation appended)	,,	0	2	5	0
Steel	,,	0	2	5	0
Tin	,,	1	2	5	0
Tin plates	,,	0	4	0	0
Mother-o'-pearl shell	,,	0	2	0	0
Musical boxes	5 per cent.	ad valorem.			
Mussels, dried	100 catties	0	2	0	0
Nutmegs	,,	2	5	0	0
Olives, unpickled, salted, or pickled	,,	0	1	8	0
Opium	,,	30	0	0	0
Pepper, black	,,	0	3	6	0
,, white	,,	0	5	0	0
Prawns, dried	,,	0	3	6	0
Putchuk	,,	0	6	0	0
Rattans	,,	0	1	5	0
Rose maloes	,,	1	0	0	0
Salt fish	,,	0	1	8	0

Treaty of Peace, Friendship and Commerce between China and Great Britain, 26 June 1858 (volume 119, page 163, Consolidates Treaty Series), Annex

Import and export prohibitions. "The Contracting Parties will not hamper the commercial relations between the two countries by any special measures, more particularly by import or export prohibitions.

The Contracting Parties nevertheless reserve the right to prohibit or restrict imports and exports in the following cases, provided that the said prohibitions or restrictions are simultaneously applicable to all other countries or to countries in which similar conditions prevail:

(i) For reasons of public security;

(2) For moral or humanitarian motives;

(3) In respect of the traffic in arms, ammunition, and war material, or, under exceptional circumstances, of that in all other materials needed in war;

(4) In respect of the protection of public health or the protection of animals and plants against disease, insects or noxious parasites;

(5) For the protection of national treasures of an artistic, historical or archaeological nature;

(6) In respect of prohibitions or restrictions applicable to gold, silver, coin, paper money or securities;

(7) In respect of prohibitions or restrictions designed to extend to foreign goods the regime imposed or hereafter to be imposed within the country itself on the production of, traffic in, transport and consumption of home products of the same kind;

(8) In respect of products which are or may hereafter be the subject of State monopolies or of monopolies controlled by the State."

Treaty of Commerce and Navigation between the Kingdom of Serbs, Croats and Slovenes and Czechoslovakia, 14 November 1928 (volume 97, page 9, League of Nations Treaty Series), Article VII

MFN Exemptions. "The provisions of the present Convention shall not apply:

(1) To treatment which is or may hereafter be granted by one of the Contracting Parties in frontier traffic with contiguous countries;

(2) To special privileges resulting from a Customs union;

(3) To special privileges and benefits which are or may hereafter be established in respect of Customs tariffs, and generally in all other commercial transactions between Turkey and the territories detached from the Ottoman Empire in 1923;

(4) To privileges and benefits which Finland has granted or may hereafter grant to Estonia with a view to preserving her traditional trade relations with that country;

(5) To privileges which Finland has granted or may hereafter grant to neighbouring States in respect of navigation in the Baltic Sea and its gulfs north of Lat. 580 N.;

(6) To privileges which Finland has granted or may hereafter grant in respect of the importation of wines and alcoholic beverages."

Convention of Commerce and Navigation between Finland and Turkey, 2 June 1926 (volume 70, page 329, League of Nations Treaty Series), Article XXIV

Quotas. "Supervision of the wine quota.

In order to obtain the advantages set forth in the present Agreement, each consignment of wine with a natural alcoholic content, in receptacles of a capacity of 50 litres or more (including tank waggons and reservoir waggons), made by the same consignor to the same consignee, must be accompanied on importation into German Customs territory by a quota certificate issued by the French Ministry of Agriculture.

Each application for a certificate must be presented in duplicate and must be made out in accordance with the attached form.

The French Ministry of Agriculture shall affix a visa on the applications up to the amount of the quota. After the visa, together with the signature of the competent official of the Ministry of Agriculture, have been obtained, the application shall serve as a quota certificate, and shall be returned to the applicant as soon as the quantity has been entered in the quota register.

The duplicate of the quota certificate shall be sent to the German Embassy in Paris with the word 'Copy' written thereon.

The quota certificate shall authorise the importer to import through one of the Customs offices through which wine is allowed to be imported, the quantities of goods mentioned in the certificate. The quota certificate must accompany the goods when they cross the frontier and must be presented for clearance together with the Customs declaration.

Quota certificates shall not be transferable; they shall only be valid in respect of the quantities for which they are issued. Nevertheless, the goods may be imported in several lots up to the amount of the said quantities, on the express condition that the total quantity covered by the same certificate is declared and presented for final clearance at the same office within a period of one month and in any case before the expiry of the present Agreement.

The quantities of goods imported into German Customs territory shall be entered

in the quota register at the weight at which the duties payable are computed. The German Government shall send a communication to the French Government as soon as 75% of the quota has been applied for.

The German Government shall communicate to the French Government each month the balance outstanding of the quantities covered by the quota certificates registered with the French Ministry of Agriculture, and new quota certificates may be issued up to this amount.

After the quota has been exhausted, these goods can only be admitted at the general tariff rate.

Wine in bottles is not included in the quota."

Commercial Agreement with Protocol of Signature and Annexes, 17 August 1927 (volume 76, page 5, League of Nations Treaty Series), Ad No. 18o of the German tariff

Certificate of Origin. "As regards the application of Articles I to 8, the High Contracting Parties may require that products and goods imported into their territory shall be accompanied by a certificate of origin attesting:

(i) In the case of raw materials properly so called or natural products, that they originate in the other country;

(2) In the case of a manufactured product, that as regards the raw materials incorporated in it or the labour expended upon it, it satisfies the conditions required by the importing country for the recognition of nationality, as mentioned in Article 17 above.

Certificates of origin shall be issued either by the Customs authorities or by the competent Chambers of Commerce of each of the High Contracting Parties they shall be drawn up in accordance with the forms adopted by the Customs Administration or officially recognised Chambers of Commerce in the exporting country; they shall be made out either in the language of the country of origin or in the language of the country of destination. In the former case, the two countries reserve the right to require a translation. [...]"

Commercial Agreement with Protocol of Signature and Annexes, 17 August 1927 (volume 76, page 5, League of Nations Treaty Series), Article XXII

Certification. "If, in the territory of one of the Contracting Parties, the importation of goods, in view of the requirements of its internal legislation, or their clearance at a reduced Customs tariff depends on special technical conditions concerning their

composition, purity, sanitary conditions, place of origin or similar matters, the Customs authorities of the country importing these goods shall accept the certificates drawn up by a competent authority in the exporting country. [...]"

Commercial Agreement with Protocol of Signature and Annexes, 17 August 1927 (volume 76, page 5, League of Nations Treaty Series), Article XXIV

Bibliography

Accominotti, O. and M. Flandreau (2008). Bilateral treaties and the most-favored-nation clause: The myth of trade liberalization in the nineteenth century. *World Politics 60*(2), 147–188.

Alschner, W., J. Seiermann, and D. Skougarevskiy (2018). Text of trade agreements (tota)—a structured corpus for the text-as-data analysis of preferential trade agreements. *Journal of Empirical Legal Studies 15*(3), 648–666.

Anderson, J. E. and E. van Wincoop (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review 93*(1), 170–192.

Antràs, P. (2020). De-globalisation? Global value chains in the post-COVID-19 age. Working Paper 28115, National Bureau of Economic Research.

Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine 28*, 3083 – 3107.

Autor, D. H., D. Dorn, and G. H. Hanson (2013). The China syndrome: local labor market effects of import competition in the United States. *American Economic Review 103*(6), 2121–68.

Bagwell, K. and R. W. Staiger (1999). An economic theory of GATT. *American Economic Review 89*(1), 215–248.

Baier, S. and J. Bergstrand (2004). Economic determinants of free trade agreements. *Journal of International Economics 64*(1), 29–63.

Baier, S. and J. Bergstrand (2009). Estimating the effects of free trade agreements on international trade flows using matching econometrics. *Journal of International Economics 77*(1), 63–76.

Baier, S. L. and J. H. Bergstrand (2007). Do free trade agreements actually increase members' international trade? *Journal of International Economics 71*(1), 72–95.

Bairoch, P. (1993). *Economics and World History. Myths and Paradoxes*. The University of Chicago Press.

Baker, S., N. Bloom, and S. Davis (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics 131*(4), 1593–1636.

Becker, S. and A. Ichino (2002). Estimation of average treatment effects based on propensity scores. *Stata Journal 2*(4), 358–377.

Bems, R., R. C. Johnson, and K.-M. Yi (2012). The great trade collapse. Working Paper 18632, National Bureau of Economic Research.

Bhagwati, J. (2008). *Termites in the Trading System: How Preferential Agreements Undermine Free Trade*. Oxford University Press.

Bown, C. P. and D. A. Irwin (2015). The GATT's starting oint: Tariff levels circa 1947. NBER Working Papers 21782, National Bureau of Economic Research.

Brauer, H. G. A. and M. A. Kasten (1922). *Handbook of commercial treaties: Digests of commercial treaties, conventions, and other agreements of commercial interest between all nations*. U.S. Government Printing Office.

Brown, A. G. (2009). *Reluctant partners. A history of multilateral trade cooperation, 1850-2000*. University of Michigan Press.

Clark, X., D. Dollar, and A. Micco (2004). Port efficiency, maritime transport costs, and bilateral trade. *Journal of Development Economics 75*(2), 417–450. 15th Inter American Seminar on Economics.

Costinot, A. and A. Rodríguez-Clare (2014). Chapter 4 - trade theory with numbers: Quantifying the consequences of globalization. In G. Gopinath, E. Helpman, and K. Rogoff (Eds.), *Handbook of International Economics*, Volume 4 of *Handbook of International Economics*, pp. 197–261. Elsevier.

Crump, R. K., V. J. Hotz, G. W. Imbens, and O. A. Mitnik (2009). Dealing with limited overlap in estimation of average treatment effects. *Biometrika 96*(1), 187–199.

Dawid, A. P. (1979). Conditional independence in statistical theory. *Journal of the Royal Statistical Society: Series B (Methodological) 41*(1), 1–15.

Dekle, R., J. Eaton, and S. Kortum (2008). Global rebalancing with gravity: Measuring the burden of adjustment. *IMF Staff Papers 55*(3), 511–540.

Dür, A., L. Baccini, and M. Elsig (2014). The design of international trade agreements: Introducing a new dataset. *The Review of International Organizations 9*(3), 353–375.

Egger, H., P. Egger, and D. Greenaway (2008). The trade structure effects of endogenous regional trade agreements. *Journal of International Economics 74*(2), 278–298.

Egger, P. and M. Larch (2008). Interdependent preferential trade agreement memberships: An empirical analysis. *Journal of International Economics 76*(2), 384–399.

Egger, P., M. Larch, K. E. Staub, and R. Winkelmann (2011). The trade effects of endogenous preferential trade agreements. *American Economic Journal: Economic Policy 3*(3), 113–43.

Egger, P., M. Larch, and Y. Yotov (2020). Gravity-model estimation with time-interval data: Revisiting the impact of free trade agreements. School of Economics Working Paper Series 2020-11, LeBow College of Business, Drexel University.

Egger, P. H. and F. Tarlea (2021). Comparing apples to apples: Estimating consistent partial effects of preferential economic integration agreements. *Economica 88*(350), 456–473.

Eichengreen, B. (2007). *The European Economy since 1945: Coordinated Capitalism and Beyond*. Princeton University Press.

Feenstra, R. (2004). *Advanced international trade: theory and evidence*. Princeton University Press.

Findlay, R. and K. H. O'Rourke (2007). *Power and plenty: Trade, war, and the world economy in the second millennium*. Princeton University Press.

Fink, C., A. Mattoo, and I. C. Neagu (2002). Trade in international maritime services: How much does policy matter? *The World Bank Economic Review 16*(1), 81–108.

Ganapati, S. and W. F. Wong (2023, April). How far goods travel: Global transport and supply chains from 1965-2020. Working Paper 31167, National Bureau of Economic Research.

Ghosh, S. and S. Yamarik (2004). Are regional trading arrangements trade creating? an application of extreme bounds analysis. *Journal of International Economics 63*(2), 369–395.

Goldberg, P. K. (2019). The future of trade. Technical report, IMF, Finance Development.

Goldberg, P. K. and T. Reed (2023). Is the global economy deglobalizing? And if so, why? And what is next? Working Paper 31115, National Bureau of Economic Research.

Grossman, G. M., P. McCalman, and R. W. Staiger (2021). The "New" Economics of Trade Agreements: From Trade Liberalization to Regulatory Convergence? *Econometrica 89*(1), 215–249.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis 20*(1), 25–46.

Harley, C. K. (1988). Ocean freight rates and productivity, 1740-1913: The primacy of mechanical invention reaffirmed. *The Journal of Economic History 48*(4), 851–876.

Head, K. and T. Mayer (2014). Gravity equations: Workhorse,toolkit, and cookbook. Volume 4, Chapter Chapter 3, pp. 131–195. Elsevier.

Head, K., T. Mayer, and J. Ries (2010). The erosion of colonial trade linkages after independence. *Journal of International Economics 81*(1), 1–14.

Hofmann, C., A. Osnago, and M. Ruta (2017). Horizontal depth: A new database on the content of preferential trade agreements. World Bank Policy Research Working Paper No. 7981.

Hofmann, C., A. Osnago, and M. Ruta (2019). The content of preferential trade agreements. *World Trade Review 18*(3), 365–398.

Hummels, D. (1999). Toward a Geography of Trade Costs. Technical report.

Imbens, G. (2014). Matching methods in practice: Three examples. NBER Working Papers 19959, National Bureau of Economic Research, Inc.

Imbens, G. W. and D. B. Rubin (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press.

Irwin, D. A. (1995). The GATT in historical perspective. *The American Economic Review 85*(2), 323–328.

Irwin, D. A. and K. H. O'Rourke (2011). Coping with shocks and shifts: The multilateral trading system in historical perspective. Oxford Economic and Social History Working Papers 92, University of Oxford, Department of Economics.

King, G. and R. Nielsen (2019). Why propensity scores should not be used for matching. *Political Analysis 27*(4), 435–454.

Lampe, M. (2009). Effects of bilateralism and the mfn clause on international trade: Evidence for the cobden-chevalier network, 1860–1875. *The Journal of Economic History 69*(4), 1012–1040.

Limão, N. (2016). Chapter 6 - preferential trade agreements. Volume 1 of *Handbook of Commercial Policy*, pp. 279–367. North-Holland.

Maddison, A. (2003). *The World Economy: Historical Statistics*. Development Centre Studies, OECD Publishing, Paris.

Magee, C. S. (2003). Endogenous preferential trade agreements: An empirical analysis. *The B.E. Journal of Economic Analysis Policy 2*(1), 1–19.

Magee, C. S. (2008). New measures of trade creation and trade diversion. *Journal of International Economics 75*(2), 349–362.

Olivero, M. and Y. Yotov (2012). Dynamic gravity: endogenous country size and asset accumulation. *Canadian Journal of Economics 45*(1), 64–92.

Parry, C. (1969). *The Consolidated Treaty Series. Volumes 1-243*. Oceana Publications.

Rosenbaum, P. R. and D. B. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika 70*(1), 41–55.

Rosenbaum, P. R. and D. B. Rubin (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association 79*(387), 516–524.

Rubin, D. B. (1990). Formal modes of statistical inference for causal effects. *Quality Engineering 36*, 185–188.

Rubin, D. B. (2005). Causal inference using potential outcomes. *Journal of the American Statistical Association 100*(469), 322–331.

Santamaría, M., J. Ventura, and U. Yeşilbayraktar (2020). Borders within europe. Working Paper 28301, National Bureau of Economic Research.

Santos Silva, J. and S. Tenreyro (2006). The log of gravity. *The Review of Economics and Statistics 88*(4), 641–658.

Shafaeddin, M. (1998). How did developed dountries industrialize? The history of trade and industrial policy: the cases of Great Britain and the USA. UNCTAD Discussion Papers 139.

Soloaga, I. and A. L. Wintersb (2001). Regionalism in the nineties: what effect on trade? *The North American Journal of Economics and Finance 12*(1), 1–29.

UNCTAD (2003). Trade and development report 2003. Technical report.

WTO (2011). World Trade Report 2011—The WTO and preferential trade agreements: from co-existence to coherence. Working papers.

WTO (2017). *A Handbook on the WTO Dispute Settlement System* (2 ed.). Cambridge University Press.

Yang, D. and J. E. Dalton (2012). A unified approach to measuring the effect size between two groups using SAS.

Yotov, Y. (2012). A simple solution to the distance puzzle in international trade. *Economics Letters 117*(3), 794–798.

Yotov, Y., R. Piermartini, J.-A. Monteiro, and M. Larch (2016, 11). *An Advanced Guide to Trade Policy Analysis : The Structural Gravity Model*.