

Essays in International Trade

DISSERTATION
of the University of St.Gallen,
School of Management,
Economics, Law, Social Sciences,
International Affairs and Computer Science,
to obtain the title of
Doctor of Philosophy in Economics and Finance

submitted by

David Torun

from

Zurich

Approved on the application of

Prof. Dr. Reto Föllmi

and

Prof. Dr. Peter Egger

Dissertation no. 5255

Difo-Druck GmbH, Untersiemau, 2022

Essays in International Trade

DISSERTATION
of the University of St.Gallen,
School of Management,
Economics, Law, Social Sciences,
International Affairs and Computer Science,
to obtain the title of
Doctor of Philosophy in Economics and Finance

submitted by

David Torun

from

Zurich

Approved on the application of

Prof. Dr. Reto Föllmi

and

Prof. Dr. Peter Egger

Prof. Beatrix Eugster, PhD

Dissertation no. 5255

Difo-Druck GmbH, Untersiemaun, 2022

The University of St.Gallen, School of Management, Economics,
Law, Social Sciences, International Affairs and Computer Science,
hereby consents to the printing of the present dissertation,
without hereby expressing any opinion on the views herein
expressed.

St.Gallen, June 13, 2022

The President:

Prof. Dr. Bernhard Ehrenzeller

Acknowledgements

I wish to thank my supervisor, Reto Föllmi, for all his support. His encouragement, constructive feedback and relaxed aura have guided me throughout the past years. I have very much enjoyed being able to discuss both involved mathematical problems and favorite foods with him in one sitting. I am grateful to the faculty members and PhD students—especially Frank, Josef, Marc and Severin—for all the valuable feedback and advice they have given me.

Furthermore, I enjoyed collaborating with my co-authors, Christian, Piotr, Reto and Uli. Being able to rely on your knowledge and expertise has been a huge help (and complaining about comments on our work from the audience was therapeutic). I give special thanks to Uli, who has helped me through countless challenges ever since the course phase.

I am very grateful for my wonderful colleagues at the institute, and I have deeply enjoyed every moment we spent together over the past years. In particular, I want to thank Adrian, Albert, Björn, Enea, Franziska, Lea, Matthias, Michael, Mirjam, Noémie, Philine, Piotr and Stefan for intense discussions, great feedback and fun times.

Finally, I wish to thank my family and close friends for their unconditional support and friendship—before, during and after this journey. Grazie di tutto mamma, papà, Jasmine. Most importantly, I want to thank Nadine for her love and understanding, and for being so strong. I am extremely thankful to have you by my side.

Abstract

This dissertation theoretically and empirically investigates the international flow of goods between countries and the welfare consequences of globalization.

In Chapter 1, I build a quantitative Ricardian model of international trade that explains country-pair and production zeros at the industry level. I model zeros via a bounded productivity distribution and a non-homothetic final goods assembly function. Without productivity caps, trade shares reduce to a standard gravity equation. I develop a novel calibration strategy to fit data on trade flows and production. Counterfactual exercises suggest that welfare changes after trade-cost shocks are typically amplified when accounting for the extensive margin of trade. Industry-level bilateral zeros change by approximately half the shock size; e.g., a 10% rise in trade costs increases the number of zeros by 5%.

Chapter 2 investigates the so-called triangle inequality (TI) in international trade. When trade costs change, re-routing opportunities—as captured by the TI—might arise because a shipment through an intermediary becomes cheaper under adjusted trade costs. We show that the “hat algebra” approach is unable to measure potential gains from re-routing. We show that standard empirical estimates of iceberg trade costs often violate the TI. We propose a model-consistent estimation routine that respects the TI. The resulting estimates allow us to compute the welfare gains from re-routing; these gains are often substantial.

Chapter 3 analyzes how the history of trade liberalizations affects aggregate trade flows today. We develop a dynamic general equilibrium model of international trade with heterogeneous firms. Serving a market involves higher fixed costs upfront. As a consequence, there is less market exit in response to a trade shock compared to a model without sunk fixed costs. We derive a gravity equation and show that countries that liberalized their trade relationship earlier trade more today. Our theory can help explain a “home bias” in international trade and why historic events can have persistent consequences for trade. We provide supporting evidence for the underlying mechanism using data on average firm sales.

Product-level international trade flows are classified using the Harmonized System (HS). In Chapter 4, we inspect the regular HS classification updates and resulting intertemporal inconsistencies in trade statistics. These updates hinder the comparison across time of products covering up to 44% of world goods trade. Existing methods to standardize different HS vintages either (i) drop numerous product codes over time, or (ii) bulk adjusted codes into large synthetic groups. We develop a simple algorithm to convert trade flows between different HS vintages that avoids (i) and (ii). The conversion estimates are robust to year and country sample choices.

Zusammenfassung

Die vorliegende Dissertation untersucht theoretisch und empirisch internationale Güterströme zwischen Ländern und die Wohlfahrtsfolgen der Globalisierung.

In Kapitel 1 entwickle ich ein Ricardianisches Handelsmodell, das Länderpaar- und Produktionsnullen auf Industrieebene erklärt. Nullen entstehen durch eine beschränkte Produktivitätsverteilung und einen nicht-homothetischen Endprodukt-Aggregator. Ohne Produktivitätsgrenzen folgen Handelsströme einer Gravitationsgleichung. Meine Kalibrierung ermöglicht die Analyse von Handelsstrom- und Produktionsdaten. Wohlfahrtsveränderungen nach Handelskostenschocks werden typischerweise verstärkt durch die Berücksichtigung der Nullen. Die bilateralen Nullen auf Industrieebene ändern sich um etwa die Hälfte des Schocks; z.B. erhöht ein Handelskostenanstieg um 10% die Zahl der Nullen um 5%.

In Kapitel 2 wird die sogenannte Dreiecksungleichheit (DU) untersucht. Bei Handelskostenänderungen können sich Umleitungsmöglichkeiten ergeben, was von der DU eingefangen wird. Wir zeigen, dass der “hat algebra” Ansatz es nicht zulässt, potenzielle Gewinne aus Umleitungen zu messen. Standardschätzungen von Eisberg-Handelskosten verletzen oft die DU, wohingegen unsere Schätzungsroutine die DU berücksichtigt. Unsere Schätzwerte ermöglichen es, die Wohlfahrtsgewinne aus Umleitungen zu berechnen. Diese Gewinne sind oft erheblich.

Kapitel 3 analysiert, wie sich die Geschichte der Handelsliberalisierungen auf die heutigen Handelsströme auswirkt. Wir entwickeln ein dynamisches Handelsmodell mit heterogenen Firmen. Die Bedienung eines Marktes ist anfangs mit höheren Fixkosten verbunden. Als Reaktion auf einen Handelsschock resultieren dadurch weniger Marktaustritte als in einem Modell ohne versunkene Fixkosten. Unsere Gravitationsgleichung zeigt, dass Länderpaare, die ihre Handelsbeziehungen früher liberalisiert haben, heute mehr miteinander handeln. Unsere Theorie kann dazu beitragen, niedrige Handelsströme zu erklären und zu verstehen, warum historische Ereignisse anhaltende Folgen für den Handel haben können. Wir liefern Belege für den zugrundeliegenden Mechanismus, indem wir Daten zu Firmenexporten nutzen.

Handelsströme auf Produktebene werden anhand des “Harmonized System” (HS) klassifiziert. In Kapitel 4 untersuchen wir regelmässige Aktualisierungen der HS-Klassifizierung und die resultierenden Inkonsistenzen in der Handelsstatistik. Diese Aktualisierungen erschweren den intertemporalen Vergleich von Produkten, die bis zu 44% des weltweiten Warenhandels ausmachen. Bestehende Methoden zur Standardisierung verschiedener HS-Versionen lassen entweder (i) zahlreiche Produktcodes über die Zeit verschwinden oder (ii) fassen angepasste Codes zu grossen Gruppen zusammen. Wir entwickeln einen Algorithmus zur Umkodierung von Handelsströmen zwischen verschiedenen HS-Versionen, der (i) und (ii) vermeidet. Die Schätzungen sind robust gegenüber der Wahl der Jahre und Länder.

Contents

1	Quantifying the Extensive Margins of Trade and Production	1
1.1	Introduction	1
1.2	The Model	8
1.3	Calibrating the Model	20
1.4	Counterfactual Exercises	27
1.5	Conclusion	35
1.A	Appendices to Chapter 1	37
2	Triangle Inequalities in International Trade	60
2.1	Introduction	60
2.2	Theory	64
2.3	Methodology	68
2.4	Counterfactual Exercises	75
2.5	Conclusion	83
2.A	Appendices to Chapter 2	84
3	Gravity with History	111
3.1	Introduction	111
3.2	Motivating Evidence	117
3.3	A Simple Example	121
3.4	Model	122
3.5	Equilibrium	129
3.6	Empirical Application	134
3.7	Conclusion	144
3.A	Appendices to Chapter 3	145
4	Disappearing Products and Jumbo Clusters	178
4.1	Introduction	178
4.2	Related Literature	183
4.3	Methodology	187
4.4	Results	191
4.5	Descriptive Exercises	195
4.6	Heterogeneity and Robustness of Weights	198
4.7	Conclusion	202
4.A	Appendices to Chapter 4	204
	References of Chapters 1, 2, 3, 4	219

Chapter 1

Quantifying the Extensive Margins of Trade and Production

1.1 Introduction

Over the past several decades, research in international trade has increasingly focused on quantitative questions. Major advances in modeling techniques and access to new data sources have deepened our understanding of globalization and its consequences, as well as the welfare gains from trade. Numerous studies of why some trading partners are more attractive than others are able to explain the size of trade flows conditional on trading—or, the intensive margin of trade.¹ However, what is less well understood is whether two countries actually trade within an industry—the extensive margin of trade.

¹See, e.g., [Eaton and Kortum \(2002\)](#), [Anderson and van Wincoop \(2003\)](#), [Melitz \(2003\)](#), [Chaney \(2008\)](#), [Arkolakis et al. \(2012\)](#), and [Caliendo and Parro \(2015\)](#). For a review of the quantitative models and the “gravity literature” see, respectively, [Costinot and Rodríguez-Clare \(2014\)](#) and [Head and Mayer \(2014\)](#).

The workhorse multi-industry models of international trade predict that each country exports to (and imports from) all countries within every industry.² However, the data paint a starkly different picture, as is well-known in the literature on international trade. Table 1.1 illustrates that the median country exports in 72% of narrowly-defined industries, and that it serves only nine out of 149 potential destinations (conditional on exporting).³ Likewise, the median share of source countries that an importer involves is equal to 9%. There is, therefore, substantial variation in the data along the extensive margin that is neglected in standard trade models.

In this paper, I incorporate (bilateral) industry-level zeros into a quantitative general equilibrium model of international trade. This model enables us to study how industry-level trade flows and production respond to changes in trade barriers and technology along both the extensive and intensive margins. It thus takes a further step towards understanding how the number of active trade relations changes after shocks to the trade environment, and whether countries choose to produce domestically in a given industry. In an application, I show that welfare changes in response to trade cost shocks are usually amplified when I allow for variation along the extensive margin. This is especially true for countries with a low to medium (per worker) income level. To my knowledge, this is the first study to quantify the welfare consequences of integrating the extensive margin into a trade model.

In Section 1.2, I build a flexible framework that includes many countries and industries, and that is able to capture the empirical variation in the number of industry-level partners. The environment is based on the multi-industry extension of the Eaton and Kortum (2002)—henceforth, EK—model (as used, e.g., by Costinot et al. (2012)). It preserves the baseline model’s ability to fit the intensive margin of trade, as well as much of its tractability. Households

²For instance, when bringing the canonical Eaton and Kortum (2002) model of international trade to the data, it either predicts that countries import from only one source (if the industry is set equal to the variety-level), or from every country (in a multi-industry version of the model à la Costinot et al. (2012)). The relation to the literature is discussed at the end of this section.

³These industries could also be called “product categories.” For instance, the HS4-digit-codes “0902” and “8701” report, respectively, trade in “tea” and “tractors.” For simplicity, I will stick to the label “industries” throughout the paper.

Table 1.1: The Extensive Margin of Trade at the Country-Industry Level

Share of...	Min.	Q1	Median	Q3	Max.
...Exported Industries	0.03	0.47	0.72	0.94	1.00
...Served Destinations ⁽ⁱ⁾	0.01	0.02	0.06	0.22	1.00
...Imported Industries	0.52	0.85	0.93	0.97	0.99
...Source Countries ⁽ⁱⁱ⁾	0.01	0.03	0.09	0.18	1.00
Nr. of Observations	27,212,168				

Notes. This table depicts the distribution of country-industry-level export and import extensive margins. The first row reports the shares of industries across countries with positive exports, while the second row refers to the share of destinations that are served within an industry by an exporter. Similarly, the third and fourth rows state, respectively, the shares of industries with positive imports and the corresponding share of source countries used by an importer within an industry. Industries are identified by HS4-digit codes. The sample consists of 149 countries and 1,234 industries in the year 2016.

(i) Conditional on exporting in a given industry.

(ii) Conditional on importing in a given industry.

Data source. [Atlas of Economic Complexity](#).

Results. Author's computations.

obtain utility from consuming a finite set of non-tradable final goods—the industry-level—which are composed out of a continuum of intermediate varieties. These intermediate inputs are tradable, and produced by perfectly competitive firms. I extend the canonical model in two ways: with a non-homothetic final goods assembly function, and an upper-bounded distribution of firms' productivities.⁴ In the assembly process, the marginal productivity of an intermediate variety at zero is finite. This yields a “choke price” above which an intermediate input is not used to assemble the final good.⁵ The country-industry-specific truncated distribution translates into a (positive) minimum price that an exporting country may offer to a specific destination within an industry. Naturally, if an exporter's lower-bound price is above the household's industry-specific choke price, this exporter will not serve that destination. As a consequence, a country may import from (and export to) only a few, and potentially zero, partners. A country pair may not trade at all, and a country

⁴More precisely, I solve the model for any productivity upper bound. The framework can thus accommodate arbitrarily large productivity cutoffs.

⁵This contrasts with standard CES aggregators, which would necessarily use all inputs.

does not necessarily produce intermediate varieties in every industry.⁶ In a nutshell, allowing for (i) the most productive firms to have finite productivity levels, and (ii) intermediate inputs to be substitutable along the extensive margin, is sufficient to embed the empirical facts above in an otherwise standard trade model.

The framework developed in this paper yields an expression for the level of bilateral trade flows that depends on exogenous parameters, as well as endogenously determined wages and choke prices. This allows me to fit bilateral industry-level trade flows without requiring information on domestic expenditures, which is often unavailable at narrow industry levels. By contrast, standard specifications only pin down trade shares (i.e., trade flows scaled by industry-level expenditure). The equation for trade flows resembles a “gravity” equation, but with a crucial difference.⁷ The “destination-size” component may vary across exporters as a result of the price lower bounds (i.e., the minimum price an exporting country can offer) being exporter-destination-specific. A major benefit of my framework is that the trade shares reduce to EK’s well-known expression when productivity upper bounds are infinite.

In Section 1.3, I develop a novel calibration strategy to fit data on aggregate production and industry-level bilateral trade flows. The main advantage of the closed-form expression for trade flows is that I can directly solve the model through explicit equilibrium conditions. This leads to a transparent calibration design that does not require additional assumptions when bringing the framework to the data. In an application of this routine, I fit data at the HS4-digit industry level for the top 80 countries in terms of total production. The model is able to replicate several important empirical features. The distribution of bilateral zeros is reproduced by the model. My main specification matches most empirical zeros (roughly 86%), and its intensive-margin fit is

⁶About 34% of exporter-importer dyads do not have a single positive flow across industries in the data used for Table 1.1. Moreover, production zeros are relevant, *inter alia*, for commodities. Fally and Sayre (2018) report that countries frequently have no domestic production of commodities despite having positive demand. To give an example, the HS4-digit-code “2709” records trade in “crude petroleum oil.”

⁷The gravity equation, in its typical form, relates trade flows to the product of the exporter’s size, the importer’s size, and a measure of bilateral trade barriers.

comparable to that of more standard frameworks. To create a reference point, I re-calibrate the model imposing an infinite productivity upper bound, which eliminates the extensive margin, and predicts the same trade shares as the EK model. I label the latter the “benchmark model,” while the version with a flexible productivity upper bound is the “extensive model.”

Section 1.4 puts the calibrated models to work and presents several counterfactual exercises. The focus lies on the response of trade flows and welfare to exogenous trade cost shocks. In the main counterfactual scenarios, I simulate both a drop and a rise in trade costs of 10%. First, the results reveal a significant reaction upon changes in trade barriers along the extensive trade margin. Industry-level bilateral zeros change by approximately half the shock size. For instance, a 10% rise in trade barriers increases the number of zeros by about 5%. This is a sizable change, since the share of bilateral industry-level zeros in the baseline equilibrium exceeds 60%. However, the impact of these dissolved trade links is modest in terms of trade lost along the intensive margin. There is substantially less movement along the extensive margin in terms of production. After a 10% drop in trade costs, there are only six (ten) instances where a country starts (stops) producing in a given industry.

Second, I find that welfare responses are generally larger in the extensive version than in the benchmark model, although changes are highly correlated between the two. The gains from trade liberalizations, and the losses from protectionist measures, tend to be amplified when allowing for variation along the extensive margin. The welfare gains from the liberalization exercise are at least 10% larger in the extensive model for a fifth of the sample. This primarily concerns low- to medium-income countries. By contrast, welfare responses are very similar across both versions of the model for high-income producers. This is a result of the productivity upper bounds being higher for richer economies. Welfare losses due to rising trade barriers are, if anything, slightly smaller in the extensive model for some high-income countries like the U.S. Overall, a similar picture arises when the shock size is equal to 5% or 15%.

Relation to the literature. This paper contributes to several strands of the literature.

Extensions of the EK model are go-to tools to quantitatively evaluate trade policy questions, and the consequences of globalization in general. For instance, the framework has been extended to include multiple industries (e.g., [Costinot et al., 2012](#)), sectoral input-output linkages (e.g., [Caliendo and Parro, 2015](#)),⁸ dynamics (e.g., [Caliendo et al., 2019](#); [Eaton et al., 2016](#)), and the current account ([Ravikumar et al., 2019](#)). I contribute to these studies by showing the quantitative importance of the extensive margins of trade and production at the industry level.

Perhaps most closely related to my study are the EK extensions by [Chor \(2010\)](#), [Fally and Sayre \(2018\)](#), and [Eaton and Fielser \(2019\)](#). [Chor \(2010\)](#) uses an upper- and lower-bounded Fréchet distribution, which generates bilateral industry-level zeros. Although the mechanism resembles that in my model, this specification loses tractability and cannot provide a closed-form expression for trade flows. This, in turn, implies that the model needs to be simulated, which is computationally expensive and limits the author to examining only a handful of countries and industries. [Fally and Sayre \(2018\)](#) show that the gains from trade are larger when we account for the low price elasticities of commodities. I take a broader approach by allowing countries to be very unproductive in certain industries. If a natural resource is scarce in a country, this translates into a very small productivity upper bound in the associated industry.⁹ [Eaton and Fielser \(2019\)](#) introduce an extensive margin of product-level trade that arises from a stochastic probability of (not) observing small shipments. I add to their contribution by incorporating industry-level productivity differences, allowing for aggregate bilateral zeros, and obtaining a closed-form expression for industry-level trade flows.

Models with heterogeneous firms that operate in a monopolistically competitive environment also explicitly model an extensive margin. In a [Melitz \(2003\)](#)-type model, firms may not serve a destination due to the fixed costs of

⁸Section 1.2.4 outlines an easy way to introduce sectoral linkages into my framework.

⁹In one specification they include a choke price above which a commodity is not used; this relates to my specification.

market access. [Helpman et al. \(2008\)](#) include a bounded Pareto distribution of productivities into such a framework to explain aggregate bilateral zeros in the data.¹⁰ However, a multi-industry extension of their model would require information on industry-level expenditure for a full calibration. [Eaton et al. \(2013\)](#) show that having only a finite number of firm—and thus productivity—draws generates aggregate bilateral trade zeros. The authors’ specification poses a challenge for performing a full calibration because of its data requirements.¹¹ I complement these approaches by developing a tractable framework that (i) generates zeros in the absence of fixed trade costs, and (ii) replicates important empirical regularities despite employing only readily available data.¹²

More broadly, my analysis relates to studies emphasizing the frequent occurrence of zeros in international trade data. A seminal contribution by [Santos Silva and Tenreyro \(2006\)](#) proposed using the by now standard Poisson pseudo-maximum-likelihood (PPML) estimator, which allows for zeros in the dependent variable. Empirical studies of the extensive margin of trade include [Hummels and Klenow \(2005\)](#) (who find that the extensive margin explains why larger economies export more in absolute terms), [Dutt et al. \(2013\)](#) (investigating the impact of WTO membership along the extensive margin), and [Baier et al. \(2014\)](#) (effect of economic integration agreements on the extensive margin of trade). I add to empirical ex-post analyses by proposing a mechanism that generates bilateral industry-level zeros and by simulating counterfactual scenarios.

¹⁰In line with [Helpman et al. \(2008\)](#), [Adão et al. \(2020\)](#) introduce an upper-bounded distribution to allow for bilateral zeros in a trade model with monopolistic competition. The key innovation is that their distribution of firm fundamentals has no parametric restriction.

¹¹More generally, there is a fundamental difference to perfectly competitive environments: when a firm-specific—and thus country-specific—variety becomes unavailable, there is no alternative source for this specific product. The utility function in these models is therefore defined over the set of *available* varieties. At the other extreme, in models with perfect competition (such as the one presented here) countries have the ability to produce every variety. Nevertheless, the baseline frameworks—without zeros—of both competition types yield similar welfare predictions (see, [Arkolakis et al., 2012](#); [Costinot and Rodríguez-Clare, 2014](#)).

¹²Taking a different approach, [Armenter and Koren \(2014\)](#) explain trade zeros using a statistical model that captures the sparsity of trade data.

1.2 The Model

I develop a multi-industry general equilibrium international trade model that allows for bilateral (industry-level) zeros, as well as export and import zeros. This static framework builds on the multi-industry version of [Eaton and Kortum \(2002\)](#) (henceforth, EK), as applied by, for example, [Costinot et al. \(2012\)](#). The key novelty is an upper bound of the productivity distribution, which, coupled with a non-homothetic final goods aggregator, can create zeros along several dimensions. An importer may import from only a few, and potentially zero, exporters in a given industry. Similarly, an exporter may serve only a few, and potentially zero, destinations in a given industry. A country pair may, as a consequence, not trade with each other at all. Countries are indexed by $i, j = \{1, 2, \dots, N\}$, while final goods are indexed by $k = \{1, 2, \dots, K\}$. I use “industry” to refer to the final consumption goods level, k , while “varieties” refers to tradable intermediate inputs which compose the final goods. Households obtain utility from consuming the final goods, which they assemble using industry-specific tradable intermediate inputs.¹³ They inelastically supply labor to competitive domestic intermediate producers which behave à la EK, and draw their productivities from an upper-bounded distribution. Below I explain each of these layers in detail.

1.2.1 Households

Households assemble final consumption goods by combining industry-specific tradable intermediate inputs. The assembled products are non-tradable, and prices are taken as given, a standard assumption in the literature.¹⁴

Demand function. Country j is populated by L_j symmetric households, each inelastically supplying one unit of labor to domestic intermediate producers. A household obtains utility from consuming a finite set of non-tradable

¹³In Section 1.2.4, I extend the model to allow for final goods to be composed of varieties from all industries.

¹⁴A modeling alternative would be to introduce competitive firms that import intermediate varieties and sell non-tradable assembled final products to households. While this would add another layer to the model, it would not yield any additional insight.

final goods

$$U_j = \frac{\sigma}{\sigma - 1} \sum_{k=1}^K (\xi_j^k)^{\frac{1}{\sigma}} (q_j^k)^{\frac{\sigma-1}{\sigma}},$$

where $\sigma > 0$ is the elasticity of substitution, and ξ_j^k are exogenous demand shifters satisfying $\sum_{k=1}^K \xi_j^k = 1$.¹⁵ Each final product k is composed of k -specific tradable varieties according to

$$q_j^k = \int_0^1 q_j^k(\omega^k) d\omega^k, \quad (1.1)$$

where $q_j^k(\omega^k) \in \{0, 1\}$ is the quantity of intermediate ω^k used for final product k . There are four important properties of the model that follow from this functional form assumption. First, in each industry k , there is a fixed measure of intermediates that is normalized to one, as is standard in EK-type models. Second, there is an intensive margin of consumption, since q_j^k is the total quantity of final consumption product k , even though the use of intermediates is modeled exclusively via the extensive margin. Third, marginal costs of final goods assembly are increasing—more on that below. Fourth, the marginal productivity of a variety at zero is finite, which implies that some intermediate varieties will not necessarily be used in assembling the final product. As I show in Section 1.2.2, this is a necessary assumption to allow for (bilateral) industry-level zeros in this framework.

The fact that intermediates enter the aggregator symmetrically implies that the order in which consumers add varieties follows the prices of intermediates. In other words, consumers first use the cheapest intermediate inputs. An intuitive way to think about q_j^k is thus in terms of quality or product complexity. The larger the mass of employed varieties to assemble the final good, the higher the utility from consuming it. Moreover, a larger mass of employed varieties leads to a higher average consumption price. I will, however, use q_j^k

¹⁵The derivation of trade flows in Section 1.2.2 does not rely on a specific (instantaneous) utility function. In this section, I derive the equilibrium with an explicit homothetic utility function, which will be used in the calibration (cf. Section 1.3). Appendix 1.A.A.c characterizes the equilibrium with a general additively separable utility function.

to refer exclusively to the final product's quantity.¹⁶

The household maximizes its utility subject to its budget constraint

$$E_j \geq \sum_{k=1}^K \left[\int_0^1 p_j^k(\omega^k) q_j^k(\omega^k) d\omega^k \right],$$

where E_j denotes its total income, and $p_j^k(\omega^k)$ the price of intermediate variety ω^k . Let varieties be sorted such that the intermediate's price, $p_j^k(\omega^k)$, is increasing in the index value—i.e., $\omega^k = 0$ ($\omega^k = 1$) is the cheapest (most expensive) intermediate input. This facilitates notation, and is without loss of generality. Let the “marginal variety” that is just used in the assembly process be $\bar{\omega}_j^k \in [0, 1]$. That is, varieties $\omega^k \in [0, \bar{\omega}_j^k]$ are used for aggregation, while varieties $\omega^k \in (\bar{\omega}_j^k, 1]$ are not. The household maximizes its utility with respect to $\{\bar{\omega}_j^k\}_k$, which yields the following demand functions

$$q_j^k = \xi_j^k (\lambda_j p_j^k(\bar{\omega}_j^k))^{-\sigma} \quad \forall k,$$

where λ_j is the Lagrange multiplier.¹⁷ Accordingly, I can define a “choke price” $\bar{p}_j^k := p_j^k(\bar{\omega}_j^k)$ above which intermediates are not used to assemble the final product k in country j . The assembly function implies that in equilibrium q_j^k equals $\bar{\omega}_j^k$. The average expenditure on intermediates from industry k is given by

$$P_j^k := \frac{1}{q_j^k} \int_0^{\bar{\omega}_j^k} p_j^k(\omega^k) d\omega^k. \quad (1.2)$$

Crucially, P_j^k also corresponds to the average price of varieties used in the

¹⁶For Ricardian models with explicit quality differentiation, see, e.g., [Eaton and Fielser \(2019\)](#); [Jaimovich and Merella \(2015\)](#).

¹⁷Specifically, the Lagrange function is given by

$$\mathcal{L}(\{\bar{\omega}_j^k\}_k, \lambda_j) = \frac{\sigma}{\sigma-1} \sum_{k=1}^K (\xi_j^k)^{\frac{1}{\sigma}} \left(\int_0^{\bar{\omega}_j^k} d\omega^k \right)^{\frac{\sigma-1}{\sigma}} + \lambda_j \left[E_j - \sum_{k=1}^K \int_0^{\bar{\omega}_j^k} p_j^k(\omega^k) d\omega^k \right].$$

Since, in equilibrium, there will be an increasing continuum of intermediate prices (cf. Section 1.2.2), there is a single crossing point between marginal utility and “marginal” prices.

assembly, as q_j^k is also the mass of employed varieties. Multiplying both sides of the demand function by P_j^k , and summing over products k yields

$$\sum_{k=1}^K P_j^k q_j^k = (\lambda_j)^{-\sigma} \sum_{k=1}^K \xi_j^k (\bar{p}_j^k)^{-\sigma} P_j^k.$$

This condition can be used to substitute for λ_j in the demand function above. The left-hand side of this expression corresponds to total expenditures, which are equal to E_j due to the budget constraint. Using this, we obtain

$$P_j^k q_j^k = E_j \frac{\xi_j^k (\bar{p}_j^k)^{-\sigma} P_j^k}{\sum_{k'=1}^K \xi_j^{k'} (\bar{p}_j^{k'})^{-\sigma} P_j^{k'}}. \quad (1.3)$$

These expenditure shares pin down the choke prices in equilibrium, as outlined in Section 1.2.3. To quantify welfare, I can simply plug the calibrated values of q_j^k , as well as ξ_j^k , and σ , into the utility function above (cf. Section 1.4).

The next section shows that the direct mapping between average prices and total expenditures is useful to characterize industry-level trade flows. I further show that if I used an *unbounded* productivity distribution for the efficiency draws in the intermediate production, predicted trade shares would be the same irrespective of whether I chose the “0-1” assembly function above or a standard CES aggregator. This is, however, not the case with the *upper-bounded* productivity distribution.

1.2.2 Production of Tradable Intermediate Inputs

Competitive firms that behave à la EK produce tradable intermediate varieties using labor as the only input into production.

Trade costs. Throughout the paper, i indexes exporters, while j indexes importers. Intermediate varieties can be traded subject to standard iceberg trade costs. For one unit of an intermediate variety ω^k to reach importer j from exporter i , $\tau_{ij}^k \geq 1$ units need to be shipped. As is standard in the literature, I normalize $\tau_{jj}^k = 1$, and assume that the triangle inequality holds, such

that $\tau_{ij}^k \leq \tau_{ih}^k \tau_{hj}^k \ \forall i, h, j$.¹⁸

Technology. Labor is assumed to be perfectly mobile across industries but immobile across countries. This results in a common wage across industries within a country. Intermediate producers operate in perfect competition with constant returns to scale, and labor productivity in country i to produce a variety ω^k is denoted by $\varphi_i^k(\omega^k)$. Hence, the cost of producing one unit of ω^k and shipping it to country j is equal to the marginal cost

$$c_{ij}^k(\omega^k) := \frac{w_i \tau_{ij}^k}{\varphi_i^k(\omega^k)},$$

where w_i is the wage rate in country i . As in EK, I assume that $\varphi_i^k(\omega^k)$ is independently drawn from a Fréchet distribution (Type II extreme value distribution).¹⁹ However, the distribution is truncated from above, that is,

$$F_i^k(\varphi) = \frac{\exp(-T_i^k \varphi^{-\theta})}{\exp(-T_i^k(\bar{\varphi})^{-\theta})}, \quad \text{for } \varphi \in [0, \bar{\varphi}], \quad (1.4)$$

where $\bar{\varphi}$ is the productivity upper bound, T_i^k captures the general state of technology in country i to produce the varieties of industry k , and θ is inversely related to the dispersion of productivities within a country-industry cell. The truncation, together with the assembly function in equation (1.1), is necessary to create (bilateral) industry-level zeros, as will become clear below.²⁰

Price distribution. By perfect competition, country j 's price to acquire variety ω^k from source country i , $p_{ij}^k(\omega^k)$, is equal to $c_{ij}^k(\omega^k)$. The corresponding price distribution is given by the probability of country i 's price being

¹⁸For details on the triangle inequality see Chapter 2.

¹⁹I assume that there is a one-time, free efficiency draw per country and variety. If in equilibrium some varieties are not produced in country i , the efficiency realization represents country-variety-specific implicit knowledge about production. In equilibrium, firms anticipate whether selling a variety is profitable and refrain from doing so if it is not (without bearing any costs).

²⁰The entire derivation below would be analogous with country-industry-specific upper bounds, $\bar{\varphi}_i^k$, and industry-specific shape parameters, θ^k .

below a specific value p . Formally, the distribution of prices in country j for k -specific varieties from i , $G_{ij}^k(p) := \Pr[p_{ij}^k(\omega^k) \leq p] = \Pr[\varphi_i^k(\omega^k) \geq w_i \tau_{ij}^k/p]$, is given by

$$G_{ij}^k(p) = \begin{cases} 1 - \frac{\exp(-T_i^k (w_i \tau_{ij}^k)^{-\theta} p^\theta)}{\exp(-T_i^k (\bar{\varphi})^{-\theta})} & \text{for } p \in \left[\frac{w_i \tau_{ij}^k}{\bar{\varphi}}, \infty\right) \\ 0 & \text{otherwise.} \end{cases} \quad (1.5)$$

The corresponding probability density function, $g_{ij}^k(p)$, is given by $dG_{ij}^k(p)/dp$.²¹

With the distributions above, we can derive the overall price distribution in country j 's industry k . Let $\underline{p}_{ij}^k := w_i \tau_{ij}^k / \bar{\varphi}$, which is the minimum price that exporter i may offer to destination j . For a given importer j and industry k , I sort exporters such that

$$\underline{p}_{1j}^k \leq \underline{p}_{2j}^k \leq \dots \leq \underline{p}_{Nj}^k, \quad (1.6)$$

which comes in handy in the derivations below. To economize on notation, the rest of this section omits industry-superscripts k , since the derivation of trade flows is analogous across industries. However, note that trade flows will differ across industries due to industry-specific parameters.

Consumers shop globally and source input varieties from the cheapest available producer. The probability that a price is below p in country j is equal to $G_j(p) := 1 - \prod_{i=1}^N [1 - G_{ij}(p)]$. Let $\underline{p}_{(N+1)j} = \infty$ for notational convenience. Using the expression for $G_{ij}(p)$ above, and the fact that exporters are sorted by minimum prices in ascending order, we obtain

²¹There is a crucial difference between this and the well-known benchmark case where $\bar{\varphi} \rightarrow \infty$. The price distribution does not have positive support over every $p \geq 0$. But, more interestingly, even when the upper bound of the productivity distribution is not country-specific, the support differs across exporters due to differing input prices. This in turn implies that strong exporters—with low $w_i \tau_{ij}^k$ —can offer prices that less competitive source countries are unable to provide.

$$G_j(p) = \begin{cases} 1 - \gamma_j^{1:n} \exp(-p^\theta \Phi_j^{1:n}) & \text{for } p \in [\underline{p}_{nj}, \underline{p}_{(n+1)j}) \\ 0 & \text{otherwise,} \end{cases} \quad (1.7)$$

where

$$\begin{aligned} \gamma_j^{1:n} &:= \exp\left(\sum_{i=1}^n T_i \bar{p}^{-\theta}\right) \\ \Phi_j^{1:n} &:= \sum_{i=1}^n T_i (w_i \tau_{ij})^{-\theta} \end{aligned}$$

This piecewise definition of $G_j(p)$ arises from the exporter-specific lower bounds of the respective bilateral price distributions.²² Note that $\Phi_j^{1:n}$ captures the overall state of technology and trade barriers across producers i for a given destination j . However, within a specific price interval $p \in [\underline{p}_{nj}, \underline{p}_{(n+1)j})$, only the production capabilities and transportation costs of exporters $i \leq n$ are relevant; the remaining exporters $i > n$ do not operate in this price spectrum. As I show next, this property of the price distribution, coupled with the choke prices from Section 1.2.1, can result in bilateral industry-level trade zeros.

Mass of input varieties. We can now characterize the mass of intermediate varieties that country j sources from i , which I denote by M_{ij} . Let m_{ij} be the per capita equivalent, i.e., $m_{ij} = M_{ij}/L_j$. Only varieties with prices smaller than or equal to $p_j^k(\bar{\omega}_j^k)$ will be used for the composite final good, following the household setup in Section 1.2.1. Let this “choke price” be denoted by $\bar{p}_j$ (again suppressing the k -superscript), which will be endogenously determined in equilibrium. Because there is a continuum of varieties, m_{ij} corresponds to the joint probability of i ’s prices being below $\bar{p}_j$ and below competitors’ prices, where potentially $i = j$. Therefore, as detailed in Appendix 1.A.A.a, the mass

²²Importantly, $\gamma_j^{1:n}$ is j -specific even with a common productivity upper bound across countries, because the sorting of exporters is destination-specific. Moreover, we have

$$1 - \gamma_j^{1:n} \exp\left(-\underline{p}_{(n+1)j}^\theta \Phi_j^{1:n}\right) = 1 - \gamma_j^{1:(n+1)} \exp\left(-\underline{p}_{(n+1)j}^\theta \Phi_j^{1:(n+1)}\right),$$

implying that $G_j(p)$ is a continuous function.

of varieties per capita that i exports to j equals

$$m_{ij} = T_i (w_i \tau_{ij})^{-\theta} \times \sum_{n=i}^N \frac{\gamma_j^{1:n}}{\Phi_j^{1:n}} \left[\exp \left(-\min\{\underline{p}_{nj}, \bar{p}_j\}^\theta \Phi_j^{1:n} \right) - \exp \left(-\min\{\underline{p}_{(n+1)j}, \bar{p}_j\}^\theta \Phi_j^{1:n} \right) \right], \quad (1.8)$$

where $\gamma_j^{1:n}$ and $\Phi_j^{1:n}$ are as defined in equation (1.7), and exporters are sorted according to inequalities (1.6). Note that due to the “0-1” assembly function, M_{ij} is also the total *quantity* flow from i to j .

The expression for m_{ij} nicely illustrates how zeros arise in this framework. The mass of intermediates that j sources from i is equal to zero if $\underline{p}_{ij} \geq \bar{p}_j$. This means that country i is unable to serve market j whenever the destination’s choke price is lower than the exporter’s minimum attainable production cost. Hence, $\bar{p}_j < \infty$ together with $\bar{\varphi} < \infty$ can lead to (i) country i not selling to destination j , and (ii) country i not exporting or importing any variety within industry k in equilibrium. This enables me to explain zeros at the importer-industry, exporter-industry, and bilateral(-industry) levels without requiring the introduction of ad hoc solutions like infinite trade barriers. The overall mass of intermediates that country j uses in its assembly of a final good is given by $M_j := \sum_{i=1}^N M_{ij}$, with $m_j = M_j/L_j$. By summing the above expression over i , it is easy to see that m_j corresponds to the price distribution in country j evaluated at $\bar{p}_j$, that is,

$$m_j = G_j(\bar{p}_j). \quad (1.9)$$

Note that $\bar{p}_j \rightarrow \infty$ yields $m_j = 1$, implying that in such a case every variety in this industry is employed to assemble the final good.

Trade flows. The central outcomes of the model that stem from the intermediates layer are bilateral trade flows, i.e., expenditures on varieties from

foreign countries. Recall that to assemble q_j^k only one (marginal) unit of a variety is added, if that variety is used in the aggregator. Therefore, following equation (1.2), to obtain expenditures we need to integrate over the prices of employed varieties. In Appendix 1.A.A.b, I show that within each industry this results in per capita trade flows from exporter i to importer j equal to

$$x_{ij} = T_i (w_i \tau_{ij})^{-\theta} \times \sum_{n=i}^N \frac{\gamma_j^{1:n}}{\Phi_j^{1:n}} (\Phi_j^{1:n})^{-1/\theta} \tilde{\Gamma} \left(\frac{\theta+1}{\theta}, \min\{\underline{p}_{nj}^\theta, \bar{p}_j^\theta\} \Phi_j^{1:n}, \min\{\underline{p}_{(n+1)j}^\theta, \bar{p}_j^\theta\} \Phi_j^{1:n} \right), \quad (1.10)$$

where $\tilde{\Gamma}(s, \underline{x}, \bar{x}) := \int_{\underline{x}}^{\bar{x}} x^{s-1} \exp(-x) dx$.²³ By symmetry of the households, total expenditures on intermediates from exporter i are given by $X_{ij} = L_j x_{ij}$. Note that $\tilde{\Gamma}(\frac{\theta+1}{\theta}, p, p) = 0$, showing that equation (1.10) allows for zeros in (bilateral) trade flows. The mechanism is analogous to that for $m_{ij} = 0$ discussed above.

There are three main insights from the expression for x_{ij} . First, equation (1.10) highlights that, loosely speaking, $\bar{\varphi}$ predominantly affects the *extensive* margin of trade (via $\underline{p}_{ij} = w_i \tau_{ij} / \bar{\varphi}$), while T_i governs the *intensive* margin.²⁴ Second, the model yields an expression for the level of trade flows as opposed to trade shares, which would be scaled by total expenditure. This has the major advantage of not requiring information on domestic industry-level trade flows to fit the data, which is often unavailable at narrow industry levels. Third, the model generates something akin to a gravity equation with a crucial difference. Although the terms within the sum are the same across ex-

²³ $\tilde{\Gamma}(s, \underline{x}, \bar{x})$ can be directly evaluated using the so-called lower incomplete gamma function, which is defined as $\Gamma(s, \bar{x}) := \int_0^{\bar{x}} x^{s-1} \exp(-x) dx$. Thus, $\tilde{\Gamma}(s, \underline{x}, \bar{x}) = \Gamma(s, \bar{x}) - \Gamma(s, \underline{x})$.

²⁴For instance, if two exporters i and i' have the same lower-bound prices, $\underline{p}_{ij} = \underline{p}_{i'j} < \bar{p}_j$, their exports to j are exactly proportional to T , that is,

$$\frac{X_{ij}}{X_{i'j}} = \frac{T_i}{T_{i'}}.$$

porters, the number of these elements to be added up is exporter-dependent. Therefore, this “market size” component varies by exporter, which impedes the running of a standard log-linear gravity regression with exporter and importer fixed effects and a proxy for τ_{ij} . A more elaborate calibration strategy is necessary, as I outline in Section 1.3. Finally, country j ’s (industry-level) per capita expenditures on intermediate varieties are given by $x_j = \sum_{i=1}^N x_{ij}$, yielding

$$x_j = \sum_{n=1}^N \gamma_j^{1:n} (\Phi_j^{1:n})^{-1/\theta} \tilde{\Gamma} \left(\frac{\theta+1}{\theta}, \min\{\underline{p}_{nj}^\theta, \bar{p}_j^\theta\} \Phi_j^{1:n}, \min\{\underline{p}_{(n+1)j}^\theta, \bar{p}_j^\theta\} \Phi_j^{1:n} \right), \quad (1.11)$$

and total (industry-level) expenditures are given by $X_j = L_j x_j$.

Comparison to benchmark. It is instructive to compare the expressions above to the results of EK. For instance, suppose the productivity distribution is unbounded (i.e., $\bar{\varphi} \rightarrow \infty$). Equation (1.10) then reduces to

$$X_{ij} = \frac{T_i (w_i \tau_{ij})^{-\theta}}{\sum_{i'=1}^N T_{i'} (w_{i'} \tau_{i'j})^{-\theta}} X_j,$$

which, after dividing by total expenditure X_j , is equal to EK’s well-known expression for trade shares.²⁵ Therefore, in terms of trade flows, my model nests the canonical EK version as a special case. This result is due to an established regularity in the benchmark model. Namely, with an *unbounded* productivity distribution, the distribution of prices conditional on i being the lowest-cost provider is equal to $G_j(p)$ for every source i . This, in turn, implies that the average price provided by i to j is the same for every i . Note that this holds even if we have a finite choke price $\bar{p}_j < \infty$.²⁶ Accordingly, with

²⁵Note that in this case we would have

$$\frac{M_{ij}}{M_j} = \frac{X_{ij}}{X_j},$$

which follows from setting $\bar{\varphi} \rightarrow \infty$ in equation (1.8), and using $M_j = G_j(\bar{p}_j) L_j$.

²⁶This is in line with the main finding by [Hepenstrick and Tarasov \(2015\)](#).

$\bar{\varphi} \rightarrow \infty$, I obtain the same formula for trade shares when replacing the “0-1” assembly function in equation (1.1) with a standard CES aggregator.

What happens if instead every variety had to be used in the final goods assembly, i.e., if we let $\bar{p}_j \rightarrow \infty$? Equation (1.10) then becomes

$$x_{ij} = T_i (w_i \tau_{ij})^{-\theta} \sum_{n=i}^N \frac{\gamma_j^{1:n}}{\Phi_j^{1:n}} (\Phi_j^{1:n})^{-1/\theta} \tilde{\Gamma} \left(\frac{\theta+1}{\theta}, \underline{p}_{nj}^\theta \Phi_j^{1:n}, \underline{p}_{(n+1)j}^\theta \Phi_j^{1:n} \right),$$

where $\underline{p}_{(N+1)j} = \infty$ for notational convenience. The expression remains similar but does not allow for zeros whenever $\underline{p}_{nj} < \infty \forall n$. In other words, a homothetic aggregator in the final goods tier of the model would be unable to explain any of the industry-level zeros in the data as long as we have positive productivity upper bounds and finite trade costs.

1.2.3 Equilibrium

In equilibrium, trade is balanced, that is,

$$\sum_{i \neq j} \sum_{k=1}^K X_{ij}^k - X_{ji}^k = 0, \quad (1.12)$$

which pins down the set of wages in equilibrium.²⁷

Choke prices. To characterize the equilibrium in this economy, it is key to identify the choke prices of intermediate varieties, $\bar{p}_j^k$ (i.e., the maximum price a variety in industry k can have to just be used in the final goods aggregator). Let Q_j^k be the total quantity consumed in industry k , i.e., $Q_j^k = L_j q_j^k$. Goods market clearing implies $Q_j^k = M_j^k = G_j^k(\bar{p}_j^k) L_j$ (cf. equation (1.9)). I use $P_j^k = X_j^k / Q_j^k$ to substitute for the final good’s average price—or, the aver-

²⁷Therefore, I do not allow for trade deficits—a prevalent feature of trade data—, as is frequently the case in the literature. An easy way to allow for trade imbalances would be to add an additional parameter that matches net exports in the initial equilibrium (see Dekle et al., 2007). I address this issue again in Section 1.4.2. Furthermore, the model could easily incorporate tariff income (see, e.g., Caliendo and Parro, 2015), from which I also abstract as it is in no way essential to illustrate the main innovations of my framework.

age expenditure on intermediates from industry k . Further, note that $E_j = w_j$, since labor is the only source of household income. Using these equilibrium conditions in equation (1.3), we obtain

$$\frac{w_j L_j}{G_j^k(\bar{p}_j^k)} = (\xi_j^k)^{-1} (\bar{p}_j^k)^\sigma \left[\sum_{k'=1}^K \xi_j^{k'} (\bar{p}_j^{k'})^{-\sigma} \frac{X_j^{k'}}{G_j^{k'}(\bar{p}_j^{k'})} \right]. \quad (1.13)$$

For a given set of wages and parameters, equation (1.13) implicitly determines $\bar{p}_j^k$, since $x_j^k = X_j^k/L_j$ is also a function of $\bar{p}_j^k$, as depicted by equation (1.11). Note that $w_j L_j = \sum_{k=1}^K X_j^k$.

Set of equations. To sum up the equilibrium in this economy, for given parameters θ , σ , $\{L_i\}_i$, $\{\xi_i^k\}_{i,k}$, $\bar{\varphi}$, $\{T_i^k\}_{i,k}$, $\{\tau_{ij}^k\}_{i,j,k}$, the equilibrium is determined by a set of wages and choke prices, $\{w_i\}_i$, $\{\bar{p}_i^k\}_{i,k}$, which satisfy the equilibrium conditions in equations (1.9)–(1.13). Section 1.3 details the calibration strategy.

1.2.4 Extensions

Before turning to the calibration strategy, I discuss two possible extensions of the model. First, I outline how to introduce consumption zeros. Second, I allow for final products to be composed out of varieties from all industries.

Consumption zeros. In Appendix 1.A.A.c, I derive the equilibrium with a general additively separable utility function. The major difference to the specification in the main text is that preferences may be such that the marginal utility at zero is finite. In turn, this implies that a household does not necessarily consume every final product in equilibrium. Therefore, import zeros in an industry need not imply the purchase of domestic varieties in this industry. This specification is interesting to investigate how the range of consumed products might vary with trade barriers. Such an extension of the model would thus be well-suited to studying low-income countries or income inequality within countries. With income inequality, and thus heterogeneous households, one

would obtain an income-group-specific choke price. Apart from this, the model would look as specified in Appendix 1.A.A.c.

Sectoral linkages. In Appendix 1.A.A.d, I outline an easy way to allow for sectoral input-output linkages. Specifically, in this extension a final product k is composed of varieties from all industries, as opposed to exclusively out of k -specific intermediates. The main technical difference from the version above is that this version introduces industry-pair-specific choke prices. The intuition behind the derivation of this equilibrium is analogous to that outlined in the main text.

1.3 Calibrating the Model

In this section, I outline the calibration strategy, and discuss the model's performance in fitting the data.

1.3.1 Calibration Strategy

The calibration strategy follows naturally from the equilibrium conditions in Section 1.2.3. To capture the extensive margin in the data, I allow for country-industry-specific productivity upper bounds, $\bar{\varphi}_i^k$. Moreover, to reduce dimensionality, I set $\tau_{ij}^k = \tau_{ij}\tau_j^k$. The latter step introduces an asymmetric country-pair-specific trade cost component, along with an importer-industry-specific shifter which contains across-industry differences in import barriers. In what follows, I briefly describe the data used for the calibration. Next, I explain how to solve the equilibrium for a given set of parameters. Finally, I outline how I calibrate the parameters such that the model mimics the data.²⁸

²⁸The influential “exact hat algebra” approach developed by Dekle et al. (2007), which allows to infer welfare gains after shocks to the trade environment without requiring a calibration of baseline parameters, cannot be applied in my model. The main reason being that trade flows are not characterized by a log-linear equation. This in turn implies that constant parameters do not cancel out when I divide counterfactual trade flows by their baseline values. Also, by construction, the hat algebra method predicts a zero in the counterfactual scenario whenever the baseline value is equal to zero.

Data. To calibrate the model, I use bilateral goods trade data at the HS4-digit level from the [Atlas of Economic Complexity](#), which is a cleaned version of the widely-used UN Comtrade data. In particular, I compute five-year averages of trade flows using the years 2012–2016. By taking a single-year cross section one potentially overestimates the number of “true” zeros in that some products are shipped very infrequently (e.g., airplanes). Each HS4 product, of which there are 1,240 in total, corresponds to an index k in my model. The algorithm below targets the empirical values of X_{ij}^k for $i \neq j$.²⁹

To add gross total production, I use the International Trade and Production Database for Estimation ([ITPD-E](#)). The ITPD-E provides production data retrieved from several administrative data sources (see [Borchert et al. \(2021\)](#) for details). Production data is used to measure total expenditure, $X_j := \sum_{k=1}^K X_j^k$. I restrict the sample to the top 80 producers in 2016 (i.e., with the largest gross output values) in order to reduce the large number of parameters that need to be calibrated. As is standard in the literature on international trade, I first merge China and Hong Kong, as well as Belgium and Luxembourg. To measure L_j , I use the total labor force from the World Development Indicators ([WDI](#)). This yields a baseline value for the wage rate, since $w_j = X_j/L_j$. Aggregate wages are the numéraire. That is, I normalize wages to sum up to one, $\sum_{j=1}^N w_j = 1$.

Each demand shifter, ξ_j^k , is set equal to product k ’s overall trade share (i.e., total trade in industry k divided by overall trade). Accordingly, $\xi_j^k = \xi^k \forall j$, with the normalization $\sum_{k=1}^K \xi^k = 1$. This is done because I do not have consumption data at such a disaggregated level. An alternative specification where ξ is j - k -specific is presented in Section 1.4.2. I set $\theta = 4.14$, the well-known preferred estimate from [Simonovska and Waugh \(2014\)](#), and $\sigma = 2$,

²⁹Recall that on the left-hand side of equation (1.10) we have levels rather than trade shares. The theory does therefore not require to scale trade flows by total expenditure, X_j^k , which instead is necessary in standard models. This is a significant advantage of the assembly function in equation (1.1), since production data at such a disaggregated level is unavailable for most countries. [French \(2016\)](#), using a standard industry-level EK framework, offers an alternative that builds on [Anderson and van Wincoop \(2003\)](#) to circumvent the issue of unobserved industry-level production.

a common value in macroeconomic models.³⁰ The remaining parameters and endogenous variables are calibrated applying the algorithm below.

Solving the equilibrium. For a given set of parameters, θ , σ , $\{L_i\}_i$, $\{\xi_i^k\}_{i,k}$, $\{\bar{\varphi}_i^k\}_{i,k}$, $\{T_i^k\}_{i,k}$, $\{\tau_{ij}^k\}_{i,j,k}$, I solve for the sets of equilibrium wages and choke prices, $\{w_i\}_i$, $\{\bar{p}_i^k\}_{i,k}$. Let sets be denoted by bold letters; e.g., $\mathbf{T} := \{T_i^k\}_{i,k}$. Regarding trade costs, let $\boldsymbol{\tau} := \{\tau_{ij}^k\}_{i,j,k}$, $\boldsymbol{\tau}^1 := \{\tau_{ij}\}_{i,j}$, and $\boldsymbol{\tau}^2 := \{\tau_j^k\}_{j,k}$. A crucial benefit of the conditions in Section 1.2.3 is that they allow me to compute the equilibrium with a transparent algorithm which makes use of the expressions' monotonicity in the target variables. This algorithm consists of three steps. The steps are first listed for ease of reference, and then explained in more detail below.

A.1 Guess a set of wages $\mathbf{w}$, and choke prices $\bar{\mathbf{p}}$.

A.2 Use equation (1.13) to pin down $\bar{\mathbf{p}}$.

A.3 Verify equation (1.12). If this condition is satisfied, an equilibrium is found. Otherwise, update $\mathbf{w}$ and return to step A.2.

Step A.1 requires no further explanation.

Step A.2 pins down intermediate choke prices, $\bar{\mathbf{p}}$, using equation (1.13). More precisely, I compute

$$\frac{w_j L_j}{G_j^k(\bar{p}_j^k)} - (\xi_j^k)^{-1} (\bar{p}_j^k)^\sigma \left[\sum_{k'=1}^K \xi_j^{k'} (\bar{p}_j^{k'})^{-\sigma} \frac{X_j^{k'}}{G_j^{k'}(\bar{p}_j^{k'})} \right], \quad (1.13')$$

where trade flows are calculated according to equation (1.10). The first element of equation (1.13') is weakly decreasing in $\bar{p}_j^k$, since $g_j^k(\bar{p}_j^k) \geq 0$. Moreover, note that $X_j^k / G_j^k(\bar{p}_j^k)$ is weakly increasing in $\bar{p}_j^k$, because it is the average intermediate price in industry k and country j —multiplied by L_j —conditional on prices

³⁰There is a vast literature that delivers estimates of related elasticities (see, e.g., Broda and Weinstein, 2006; Caliendo and Parro, 2015; Giri et al., 2021; Simonovska and Waugh, 2014). Although the quantitative results depend on the specific values of θ and σ , the qualitative implications are the same across several specifications, as I show in Section 1.4.2. I therefore refrain from separately identifying these elasticities.

being capped at $\bar{p}_j^k$. Therefore, the right-hand side of equation (1.13) is weakly increasing in $\bar{p}_j^k$, ceteris paribus. Accordingly, to satisfy equation (1.13), I increase (decrease) $\bar{p}_j^k$ if equation (1.13') is positive (negative). The relative size of this change is determined by the expression in equation (1.13') divided by $w_j L_j / G_j^k(\bar{p}_j^k)$. Hence, larger deviations from equation (1.13) lead to larger changes in $\bar{p}_j^k$.

Step A.3 pins down the set of wages, $\mathbf{w}$, using equation (1.12). For every country j , I compute

$$1 - \frac{\sum_{i \neq j} \sum_{k=1}^K X_{ij}^k}{\sum_{i \neq j} \sum_{k=1}^K X_{ji}^k},$$

i.e., net exports scaled by total exports. If these deviations are sufficiently small, an equilibrium is found. Otherwise, $\mathbf{w}$ is altered, and we return to step A.2. I increase (decrease) w_j if the expression above is positive (negative). Intuitively, if country j 's exports are larger than its imports, the wage rate needs to increase.³¹ The relative size of this change is determined by the expression above. Next, I outline how the technology parameters, $\bar{\varphi}$ and $\mathbf{T}$, and trade costs, τ , are calibrated.

Calibration. The main parameters of the model are calibrated by an algorithm consisting of three steps.³² The steps are first listed for ease of reference, and then explained in more detail below.

B.1 Guess a set of productivity upper bounds $\bar{\varphi}$, technology shifters $\mathbf{T}$, and trade cost components τ^1 and τ^2 .

B.2 Use equation (1.13) to pin down $\bar{\mathbf{p}}$.

B.3 If empirical trade flows are matched well, stop. Otherwise, update $\bar{\varphi}$, $\mathbf{T}$, τ^1 , τ^2 , and return to step B.2.

³¹This procedure mirrors the tâtonnement algorithm by Alvarez and Lucas (2007).

³²Recall that θ , σ , $\mathbf{L}$, $\boldsymbol{\xi}$, and $\mathbf{w}$ are set using data and values from the literature. In particular, since wages are directly retrieved from the data, the algorithm to calibrate the technology parameters and trade costs does not use the balanced trade condition (1.12).

Step B.1 requires no further explanation. Step B.2 works exactly as step A.2, which is explained above.

Step B.3 seeks to minimize deviations from observed trade data by adjusting $\bar{\varphi}$, $\mathbf{T}$, $\boldsymbol{\tau}^1$, and $\boldsymbol{\tau}^2$, while holding the remaining parameters, as well as wages and choke prices constant. Bilateral industry-level trade flows are calculated using equation (1.10). The algorithm exploits a specific kind of variation in the data to identify each of these four parameter sets. While I use the intensive margin of trade to fix $\boldsymbol{\tau}^1$, $\boldsymbol{\tau}^2$, and $\mathbf{T}$, the extensive margin determines $\bar{\varphi}$. To update $\bar{\varphi}_i^k$, I use the extensive margin of an exporter i in industry k across importers j . An increase in $\bar{\varphi}_i^k$ implies a drop in $\underline{p}_{ij}^k$, which in turn raises the likelihood of observing $X_{ij}^k > 0$, ceteris paribus (cf. equation (1.10)). Therefore, I increase (decrease) $\bar{\varphi}_i^k$ if the count of positive flows within this i - k cell is larger (smaller) in the data than suggested by the model.

To update T_i^k , I use total exports within this i - k cell. I increase (decrease) T_i^k if country i 's exports in industry k are larger (smaller) in the data than predicted by the model.³³ The calibration of $\boldsymbol{\tau}^2$ works analogously. I increase (decrease) τ_j^k if country j 's imports in industry k are lower (greater) in the data than predicted by the model, as trade flows are inversely related to trade costs. Similarly, to calibrate $\boldsymbol{\tau}^1$, I increase (decrease) τ_{ij} if total exports from i to j are lower (greater) in the data than the model's flows. As above, the magnitude of these changes is larger the greater the relative deviation between the two respective figures. The algorithm converges when the sum of relative deviations between empirical and predicted trade flows does not change by sufficiently much compared to the previous iteration. After each update of parameters, the algorithm returns to step B.2 to pin down the choke prices.³⁴

³³Note the resemblance to an exporter-industry fixed effect in a regression, which would capture exporter i 's relative size in this industry.

³⁴Recall that I do not match industry-level home expenditure, $\{X_{jj}^k\}_{j,k}$, as this is not observed in the data. However, since there are fewer unknown parameters than observations, this is not an issue for the calibration. The calibrated values of $\boldsymbol{\tau}^1$, $\boldsymbol{\tau}^2$, $\mathbf{T}$, and $\bar{\varphi}$, together with the normalization $\tau_{jj}^k = 1 \forall j, k$, yield a "fitted" value of $\{X_{jj}^k\}_{j,k}$, and thus of $\{X_j^k\}_{j,k}$. The latter is used in step B.2.

1.3.2 The Model’s Performance in Matching the Data

Before targeting real data I tested the algorithm using simulated data. First, I randomly generated the model parameters. I then solved for the equilibrium following steps A.1–A.3. This generates trade flows, from which I omitted domestic flows to mimic what is observed in the data. To conclude, I calibrated τ^1 , τ^2 , T , and $\bar{\varphi}$ following steps B.1–B.3 while observing the remaining parameters. The results are reassuring in that the simulated flows are matched very closely, suggesting that the algorithm exploits the correct figures in the data to recover a specific parameter.

When bringing the model to the data, it performs well along several dimensions. Table 1.2 shows the distribution of the extensive margin in the data (Panel A) and the model (Panel B), analogous to the figures presented in Table 1.1. There are only a few country-industry-level export and import zeros in this subset of the data because I restricted the sample to the top 80 producers. These export and import zeros are often not picked up by the calibration. However, there is substantial variation at the bilateral extensive margin. The latter is matched very well. Recall that I did not explicitly target these moments, but tried to directly fit the bilateral extensive and intensive margins. The share of trade zeros in the data is equal to 63.2%, while that in the model is slightly lower at 61.4%. The model matches roughly 86% of the empirical zeros, which means that around 11% of the model’s zeros are “false negatives.”³⁵

To create a reference point, I re-calibrate the model using steps B.1–B.3 but imposing $\bar{\varphi}_i^k \rightarrow \infty \forall i, k$. I refer to this version as the “benchmark model,” given that, in this case, trade shares reduce to EK’s well-known expression (cf. Section 1.2.2). I label the version of the model with a flexible productivity upper bound the “extensive model.” All country pairs have positive flows across all industries in the benchmark model. However, in the literature exploring the extensive margin of trade, trade values are frequently set to zero if they

³⁵ Alternatively, one could perfectly fit the data by letting τ absorb the residual variation. I opted for the specification above because it is more transparent.

Table 1.2: The Extensive Margin of Trade – Data vs. Model

Panel A: Data	Min.	Q1	Median	Q3	Max.
Sh. Exported Industries	0.59	0.89	0.97	0.98	1
Sh. Served Destinations ⁽ⁱ⁾	0.01	0.1	0.3	0.67	1
Sh. Imported Industries	0.94	0.97	0.98	0.98	1
Sh. Source Countries ⁽ⁱⁱ⁾	0.01	0.2	0.35	0.53	1
Nr. of Observations	7,836,800				
Panel B: Model	Min.	Q1	Median	Q3	Max.
Sh. Exported Industries	0.99	1	1	1	1
Sh. Served Destinations ⁽ⁱ⁾	0.01	0.09	0.27	0.65	1
Sh. Imported Industries	0.95	1	1	1	1
Sh. Source Countries ⁽ⁱⁱ⁾	0.01	0.19	0.37	0.58	1
Nr. of Observations	7,836,800				

Notes. This table depicts the distribution of industry-level export and import extensive margins. Panel A reports the distribution in the data, Panel B the distribution in the model. The first row in each panel reports the shares of industries across countries with positive exports, while the second row reports the share of destinations that are served within an industry by an exporter. Similarly, the third and fourth rows state, respectively, the shares of industries with positive imports and the corresponding share of source countries used by an importer within an industry. An industry corresponds to the HS4-digit level. See Section 1.3.2 for details. Section 1.3.1 describes the data.

⁽ⁱ⁾ Conditional on exporting in a given industry.

⁽ⁱⁱ⁾ Conditional on importing in a given industry.

Data source. [Atlas of Economic Complexity](#).

Results. Author's computations.

fall below a fixed threshold.³⁶ Therefore, in Table 1.A.C.1 of Appendix 1.A.C, I reproduce Table 1.2 using the benchmark model and a fixed cutoff of 1,000 USD (i.e., trade flows below this value are set equal to zero). The number of positive bilateral flows is substantially higher in the benchmark model across the first to third quartiles. About 43.8% of the benchmark model's trade flows are equal to zero when imposing the cutoff, of which 93% are empirically correct. For completeness, Table 1.A.C.2 uses this cutoff in the extensive model. This produces results similar to those presented in Table 1.2.

In terms of the intensive margin, both versions of the model perform in a similar way. To compare the two, I compute the sum of absolute deviations between empirical trade flows and those generated by the models. There is a slight increase in this measure by roughly 1.8% when I move from the extensive

³⁶See [Kehoe and Ruhl \(2013\)](#) for a detailed discussion of issues related to using such cutoffs.

to the benchmark model. A better fit of the extensive version is unsurprising, given that it entails a larger number of parameters being calibrated.³⁷

In sum, the extensive model replicates most of the empirical zeros, and fits the intensive margin slightly better than the benchmark model. The section below presents counterfactual exercises that use the calibrated models as a starting point.

1.4 Counterfactual Exercises

In this section, I study changes in trade flows, production, and welfare that arise from increases and reductions of trade barriers. In every counterfactual scenario, the new equilibrium is computed via the algorithm steps A.1–A.3 described in Section 1.3.1. Moreover, since trade is not balanced in the data, I first solve the equilibrium using the calibrated parameters. This serves as the baseline comparison in the counterfactual experiments to remove potentially misleading conclusions from closing trade imbalances (more on that in Section 1.4.2).

1.4.1 Shocks to Trade Barriers

The two main counterfactual exercises decrease and increase, respectively, global bilateral trade costs by 10%. This shock size corresponds to the main simulation conducted in Eaton et al. (2013). I leave domestic trade barriers unchanged, and do not allow trade costs to fall below one. To put welfare changes into perspective, I compare them to those predicted by the benchmark model. Additionally, the extensive model enables me to analyze how the

³⁷As a further test, I evaluated the extensive model against a standard PPML fit. Specifically, I ran a PPML regression with absolute trade flows as the dependent variable, and exporter-industry- and importer-industry-specific fixed effects, as well as country-pair fixed effects, as explanatory variables. This specification is in line with the expression for trade flows from the benchmark model. Table 1.A.C.3 reproduces Table 1.A.C.1 using the PPML fit (and the 1,000 USD cutoff) and shows that the match of the extensive margin is slightly worse than that of the benchmark model. By contrast, the sum of absolute deviations decreases by about 7% when moving from the extensive model to PPML. Considering that the latter, *inter alia*, does not impose a specific structure on the destination-specific component or restrictions on τ , the extensive framework appears to perform reasonably well.

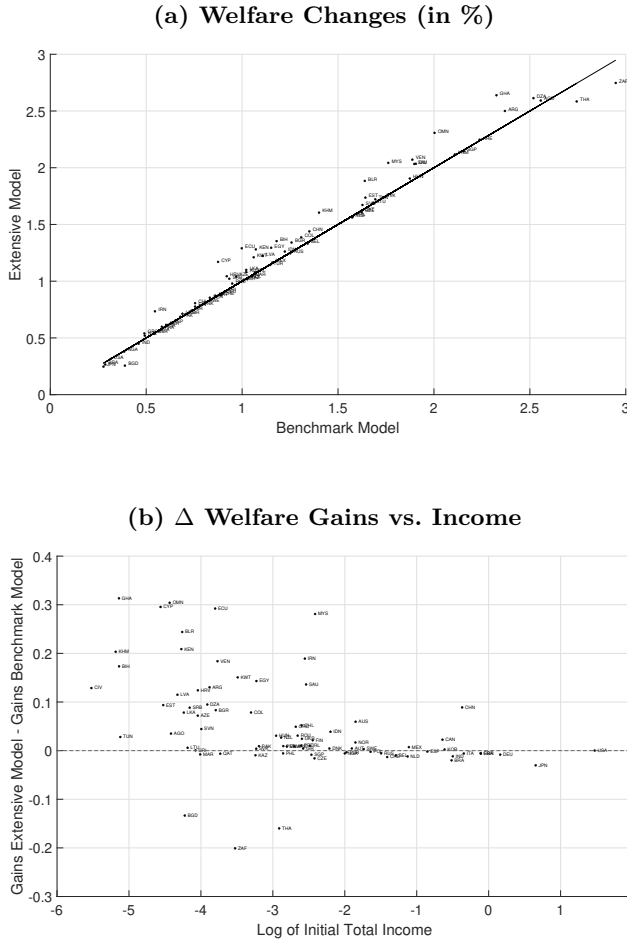
share of bilateral zeros changes in response to trade cost shocks.

Decrease in τ . A drop in global trade barriers should—analogueous to the workhorse models—positively impact trade flows and welfare. The latter is confirmed by Figure 1.1a, which depicts relative welfare changes (in percent) after a 10% drop in trade costs. Welfare increases relative to the baseline equilibrium for all countries, with growth rates ranging from 0.25% to 2.75% in the extensive model.^{38,39}

The welfare responses implied by the two versions of the model are highly correlated, while, more interestingly, the extensive model predicts larger gains than in the benchmark version for most countries. This is intuitive, since low productivity upper bounds broadly capture industry-level constraints in production, which translates into a larger benefit from trade liberalizations. These limits in productivity chiefly matter for less developed countries. For instance, regressing countries’ average $\log(\bar{\varphi}_i^k)$ on log-wages yields an elasticity of 0.82. In line with this, Figure 1.1b shows that additional welfare gains tend to be largest for countries with low to medium levels of production. By contrast, the differences are minor for most of the largest producers. This reflects the fact that the higher the productivity upper bounds become, the more the extensive version approximates the benchmark model. Therefore, if a country, and its most important trading partners, have low productivity constraints, the two models tend to predict similar changes. Figure 1.A.B.1a in Appendix 1.A.B highlights that, for lower income countries, the differences are often sizable relative to the benchmark gains. For a fifth of the sample, consisting exclusively of low- to medium-sized producers, the welfare gains are at least 10% higher in the extensive version. Figure 1.A.B.1b shows that a similar picture arises when I relate differences in welfare changes to the initial wage rate instead of the total production level. I therefore use “larger” and “richer” interchangeably in the discussion of the results.

³⁸Recall that welfare changes are not computed as real income changes, but by directly evaluating the utility function (cf. Section 1.2.1). Importantly, preferences are exactly the same in the extensive and the benchmark model.

³⁹The welfare gains suggested by trade models with a common elasticity across sectors tend to be rather low in general (see, e.g., Costinot and Rodríguez-Clare, 2014).

Figure 1.1: Welfare Changes after 10% Drop in Global Trade Costs

Notes. This figure depicts welfare changes (in %) after a drop in global trade costs of 10%. The productivity upper bound is flexible in the extensive model, and infinite in the benchmark model. Panel (b) relates the (log) initial production level, $\log(X_j)$, to the difference between the two models' welfare gains. The results are discussed in Section 1.4.

Data. See Section 1.3.1.

Graph. Author's representation.

I re-compute the welfare predictions for both a 5% and a 15% drop in trade costs. The main insights from the analysis above are confirmed: welfare gains tend to be higher in the extensive model, especially for countries with a low- to medium-sized aggregate production level (cf. Figure 1.A.B.2 in Appendix 1.A.B), although welfare implications are highly correlated between the two models. The ranking of countries in terms of their welfare differences between the two models is relatively stable across the three simulations.

How do industry-level trade flows respond along the extensive margin? The first column of Table 1.3 reports how bilateral zeros change relative to the baseline equilibrium. There is a significant reaction along the extensive margin. When global trade costs drop by 10%, the number of bilateral zeros decreases by about 6%. This is sizable, as the share of bilateral industry-level zeros in the baseline trade flow matrix exceeds 60%. In absolute terms, this corresponds to an average of 235 new positive trade links per industry. However, this hides considerable heterogeneity; the number of newly established relations ranges across industries from 60 to 420 (SD=63). In 972 cases a country starts exporting in a new industry—or, there is a 1% growth in active exporter-industry cells. In terms of trade creation along the intensive margin, the contribution from these new flows is modest (second column of Table 1.3), which is in line with the findings of Eaton et al. (2013). Finally, in terms of production, there is little movement. The extensive model predicts that in six (ten) cases a country starts (stops) producing in a given industry.

Increase in τ . In recent years, several protectionist measures have been adopted by countries that impede the free flow of goods across borders (see, e.g., Fajgelbaum et al. (2020) for a detailed analysis of the 2018 surge in U.S. import tariffs). It is, therefore, relevant to assess the stability of (industry-level) trade relations and the consequences for welfare of rising trade barriers. I thus repeat the exercises above in the opposite direction.

Figure 1.2a shows a contraction of welfare in response to a 10% increase in trade costs. Although the implied changes are similar in both models, Figure 1.2b reveals that welfare losses in the extensive version of the model are

Table 1.3: Bilateral Zeros after Changes in Global Trade Costs

%-Change in τ	%-Change in Bil. Zeros Rel. to Baseline	%-Change in Trade (through Zeros)
-5%	-3%	0.09%
-10%	-5.9%	0.35%
-15%	-8.9%	0.8%
+5%	2.9%	-0.09%
+10%	5.4%	-0.26%
+15%	7.7%	-0.48%

Notes. This table reports how bilateral industry-level trade zeros change upon a shock to trade barriers. The first column shows the percentage change in the number of zeros relative to the baseline equilibrium. For a trade cost decrease, the second column depicts trade among newly established relations relative to total bilateral trade in the baseline equilibrium. For a trade cost increase, the second column depicts trade among relations that stopped trading in the counterfactual equilibrium, scaled by total bilateral trade in the baseline equilibrium. An industry corresponds to the HS4-digit level. The results are discussed in Section 1.4.1.

Data source. See Section 1.3.1.

Results. Author's computations.

larger for most countries. This is primarily true for small- to medium-sized economies, which is intuitive in light of the discussion on the implications of production constraints above. On the contrary, some large producers appear to experience somewhat lower welfare losses when I account for the extensive margin, although the differences are mostly small.⁴⁰ Analogous to the liberalization experiment from above, Figure 1.A.B.3a in Appendix 1.A.B reveals that the relative additional losses (or gains) from accounting for the extensive margin can be meaningful. Three out of four countries have larger losses in the extensive model, of which about a fifth experiences losses that are at least 8% larger than in the benchmark version. Figure 1.A.B.3b shows that the implications are similar when the differences are related to per worker income levels. These insights are confirmed when I instead increase trade costs by 5%

⁴⁰The fact that large producers may experience lower welfare losses in the extensive model becomes less surprising when recalling the expression for trade flows in equation (1.10). Higher productivity upper bounds result in larger destination-size terms (the summation). Therefore, the productivity limits imposed by $\bar{\varphi}$ on the top producers are countervailed by the fact that exporters with lower productivity upper bounds are unable to compete in certain parts of the price spectrum (cf. Section 1.2.2 for a detailed discussion). For some large countries, this cushioning appears to slightly outweigh the negative effect of productivity caps.

or 15% (cf. Figure 1.A.B.4 in Appendix 1.A.B).

The number of bilateral industry-level zeros increases by 5.4% (cf. Table 1.3), which corresponds to an average of 215 dissolved trade relations per industry. Therefore, numerous industry-level connections appear to be relatively weak. However, as the second column in Table 1.3 emphasizes, these links tend to be small in terms of trade flows. In 1,897 cases, a country stops exporting in a given industry (representing 2% of active exporter-industry cells). Note that the effects of increasing or reducing τ on the extensive margin do not have to be symmetric. To gain intuition, consider an extreme scenario where all countries trade with each other in every industry. Reductions in trade costs would then have zero effects along the extensive margin. This is, however, not true for an increase in trade barriers.

The key takeaway is twofold. First, industry-level bilateral zeros change by approximately half the shock size. Second, welfare responses upon trade cost shocks are often amplified when I account for the extensive margin. This is primarily true for small- to medium-income economies. The richest countries have roughly the same or, if anything, slightly smaller welfare changes in the extensive version compared to the benchmark model.

1.4.2 Alternative Specifications

I alter the calibration strategy in several ways to test the robustness of the main insights presented in Section 1.4.1.

As opposed to a single-industry EK model, the multi-industry version requires a calibration of σ to assess welfare gains. Therefore, welfare implications do not only depend on the so-called trade elasticity, θ , but also on the elasticity of substitution. I re-run the calibration and main counterfactual exercises using $\sigma = \{1.5, 3\}$ instead of $\sigma = 2$. Welfare tends to respond stronger when σ is higher. However, the main insight that welfare changes are generally amplified when I account for the extensive margin is preserved in both specifications (cf. Figures 1.A.B.5 and 1.A.B.6 in Appendix 1.A.B). Analogous to the main specification, this primarily holds for small- to medium-sized economies.

I further test whether the main conclusions are driven by the choice of θ .

Figure 1.2: Welfare Changes after 10% Rise in Global Trade Costs

Notes. This figure depicts welfare changes (in %) after a rise in global trade costs of 10%. The productivity upper bound is flexible in the extensive model, and infinite in the benchmark model. Panel (b) relates the (log) initial production level, $\log(X_j)$, to the difference between the two models' welfare losses (a negative number implies larger losses in the extensive model). The results are discussed in Section 1.4.

Data. See Section 1.3.1.

Graph. Author's representation.

The baseline value $\theta = 4.14$ is taken from [Simonovska and Waugh \(2014\)](#), who estimate it using data on prices and trade flows, and a single-sector EK framework. [French \(2016\)](#) proposes a value of $\theta = 6$ in his multi-industry EK model to match an aggregate elasticity of 4.1.⁴¹ Using $\theta = 6$ in the calibration, the main counterfactuals yield the same conclusions as the baseline setting (cf. Figure 1.A.B.7 in Appendix 1.A.B). The magnitude of the welfare changes is comparable but slightly higher with $\theta = 6$. An intermediate value, $\theta = 5$, is also commonly used (see, e.g., [Costinot and Rodríguez-Clare, 2014](#)). For completeness, Figure 1.A.B.8 illustrates the results of the main simulation with $\theta = 5$. The outcome is largely consistent with the insights from Section 1.4.1.

The main specification equalizes the preference shifters, ξ_j^k , across countries, due to a lack of consumption data at the HS4-digit industry level. In an additional specification, I set ξ_j^k equal to industry k 's import share in country j (i.e., imports in industry k divided by total imports). If country j has zero imports in industry k , ξ_j^k is set equal to the minimum positive value of ξ_j^k across importers within this industry. Otherwise, countries with zero imports would never consume product k . The results in Figure 1.A.B.9 show that the main findings are robust to varying the preference shifters. Welfare changes tend to be somewhat larger in this alternative specification, which partly explains why the differences between the benchmark and the extensive model are greater compared to, e.g., Figure 1.1b. The relative differences between the benchmark and the extensive version are closer to those found in the main specification, but are slightly larger when ξ_j^k is importer-specific.⁴²

My main specification starts from a baseline equilibrium where trade is balanced to be consistent with the theory, and to avoid potentially misleading

⁴¹Still, when I regress the log of aggregate bilateral trade flows on the log of average bilateral trade costs (averaged across industries within country pairs), I obtain an elasticity equal to -3.98 in the extensive model.

⁴²I have also tried omitting ξ_j^k . In this case, there are challenges in fitting the data, probably due to the fact that the algorithm targets the trade flows in levels. The sum of absolute deviations increases by almost 30% compared to the baseline fit. Moreover, the model did not converge when I tried to compute the (counterfactual) equilibrium. A version without preference shifters would therefore likely require a different parameter to absorb scale differences across industries (e.g., an industry-specific measure of existing intermediate varieties). Nevertheless, in a recent study, [Arkolakis et al. \(2021\)](#) argue that local preference shifters are crucial in explaining variation in firm-level product exports.

conclusions from closing imbalances. In an additional exercise, I start from the calibrated baseline that does not impose zero net exports. A shock to trade costs then has two effects: it alters the trade cost matrix, and it closes trade imbalances (because the new equilibrium is found using steps A.1–A.3). The latter effect does not need to be the same across the two versions of the model, and it is not present in the theoretical framework. Still, Figure 1.A.B.10 shows the results from this specification. The main insights from the liberalization experiment are confirmed in this alternative specification (cf. Figure 1.A.B.10a). By contrast, Figure 1.A.B.10b suggests that most countries have lower welfare losses from a 10% rise in trade costs in the extensive version than in the benchmark model. However, due to the discussion before on trade imbalances, this outcome should be taken with a grain of salt. For instance, three countries (AGO, BGD, QAT) even gain in terms of welfare from the increase in trade barriers.⁴³

Finally, Table 1.A.C.4 in Appendix 1.A.C shows that the implications in terms of trade zeros are very similar across all specifications. Intuitively, a higher trade elasticity, θ , leads to larger changes.

1.5 Conclusion

It is well-known that most countries trade with only a few, and sometimes zero, partners within narrowly-defined industries. To capture this, I build a quantitative Ricardian model of international trade that explains both country-pair and production zeros at the industry level. Zeros are modeled via a bounded productivity distribution and a non-homothetic final goods assembly function. Without productivity caps, trade shares reduce to Eaton and Kortum’s (2002) familiar gravity equation.

I develop a novel calibration strategy to fit data on industry-level bilateral trade flows and aggregate production. The framework replicates the em-

⁴³Properly evaluating the impact of trade imbalances in the context of a model with extensive-margin variation would require an extension of the theory. In this regard, an interesting direction could be the model by Ravikumar et al. (2019), which introduces a current account into a dynamic EK framework. As this is beyond the scope of the present paper, I preferred abstracting from trade imbalances.

pirical distribution of bilateral zeros, and matches the intensive margin in a way comparable to more standard specifications. In counterfactual exercises, industry-level bilateral zeros change by approximately half the shock size. For instance, a 10% rise in trade costs increases the number of zeros by 5%. Welfare changes after trade-cost shocks are typically amplified when I account for the extensive margin of trade. This is primarily true for low- to medium-income countries. Instead, most high-income countries have similar welfare changes as in a version of the model with infinite productivity upper bounds.

Several extensions of the analysis would likely generate additional insights. A first step would be to use a narrower industry level (HS6-digit) and a larger sample size, especially since the welfare responses from modeling the extensive margin are strongest for low- to medium-sized countries. Moreover, the relatively rare instances in my sample where countries do not import or export in an industry at all are often not matched in the calibration. Hitting these targets would bring the model closer to the data. The share of value added in production is assumed to be one, which could certainly be refined. Finally, quantifying a version of the model with sectoral linkages, along with introducing sector-specific elasticities, would yield a more complete picture of the consequences of modeling the extensive margin. The derivation of the model allows for sector-specific elasticities, and in the paper I propose a simple way to introduce sectoral linkages into the framework. The implementation of these extensions is left open for future work.

1.A Appendices to Chapter 1

1.A.A Detailed Derivations

This appendix outlines the steps to obtain the expressions presented in Section 1.2, as well as extensions of the theory.

1.A.A.a Mass of Input Varieties

To derive equation (1.8) (i.e., the mass of intermediates per capita that country j sources from country i), I start off by calculating the probability that i is the lowest-cost provider to destination j . This derivation suppresses the superscript k , as it is analogous across industries. The probability that i offers j a price equal to p while the remaining countries have prices above p is given by $\prod_{i' \neq i} [1 - G_{i'j}(p)] g_{ij}(p)$, where $G_{ij}(p)$ is depicted by equation (1.5), and $g_{ij}(p) = dG_{ij}(p)/dp$, that is,

$$g_{ij}(p) = \begin{cases} \frac{\exp(-T_i (w_i \tau_{ij})^{-\theta} p^\theta)}{\exp(-T_i \bar{\varphi}^{-\theta})} T_i (w_i \tau_{ij})^{-\theta} \theta p^{\theta-1} & \text{for } p \in \left[\frac{w_i \tau_{ij}}{\bar{\varphi}}, \infty \right) \\ 0 & \text{otherwise.} \end{cases}$$

Hence, the probability that country i is the lowest-cost provider, π_{ij} , can be calculated as

$$\begin{aligned} \pi_{ij} &:= \int_{p_{nj}}^{\infty} \prod_{i' \neq i} [1 - G_{i'j}(p)] g_{ij}(p) dp \\ &= \sum_{n=i}^N \left[\int_{p_{nj}}^{p_{(n+1)j}} \left(\prod_{i' \leq n, i' \neq i} \frac{\exp(-T_{i'} (w_{i'} \tau_{i'j})^{-\theta} p^\theta)}{\exp(-T_{i'} \bar{\varphi}^{-\theta})} \right) \right. \\ &\quad \times \left. \frac{\exp(-T_i (w_i \tau_{ij})^{-\theta} p^\theta)}{\exp(-T_i \bar{\varphi}^{-\theta})} T_i (w_i \tau_{ij})^{-\theta} \theta p^{\theta-1} dp \right] \\ &= T_i (w_i \tau_{ij})^{-\theta} \sum_{n=i}^N \int_{p_{nj}}^{p_{(n+1)j}} \gamma_j^{1:n} \exp(-p^\theta \Phi_j^{1:n}) \theta p^{\theta-1} dp \end{aligned}$$

$$= T_i (w_i \tau_{ij})^{-\theta} \sum_{n=i}^N \frac{\gamma_j^{1:n}}{\Phi_j^{1:n}} \left[\exp \left(-\underline{p}_{nj}^\theta \Phi_j^{1:n} \right) - \exp \left(-\underline{p}_{(n+1)j}^\theta \Phi_j^{1:n} \right) \right], \quad (1.14)$$

where $\underline{p}_{(N+1)j} = \infty$, and exporters are sorted according to inequalities (1.6).⁴⁴ The second line splits the integral at the exporters' lower-bound prices to exploit the piecewise definition of $G_j(p)$ in equation (1.7). The variables $\gamma_j^{1:n}$ and $\Phi_j^{1:n}$ are as defined in equation (1.7). However, note that π_{ij} is not necessarily the probability that country i exports to country j in this model, as households do not necessarily use every variety with the “0-1” assembly function in equation (1.1). Put differently, there is not necessarily positive demand for a variety at every price level.

Using the notation from equation (1.2), only varieties with prices smaller than or equal to $p_j^k(\bar{\omega}_j^k)$ will be used for the composite final good. We thus need to find the mass of intermediate varieties that is exported from country i to country j conditional on prices being at most equal to an upper bound, $\bar{p}_j$. Denote the mass of varieties that j sources from i by M_{ij} . The per capita equivalent is $m_{ij} = M_{ij}/L_j$. Since there is a continuum of varieties, the latter is determined by the joint probability of $p_{ij}(\omega)$ being below $\bar{p}_j$ and country i being the lowest-cost provider, that is,

$$\begin{aligned} m_{ij} &:= \mathbb{1} \left\{ \bar{p}_j \geq \underline{p}_{ij} \right\} \times \int_{\underline{p}_{ij}}^{\bar{p}_j} \prod_{i' \neq i} [1 - G_{i'j}(p)] g_{ij}(p) dp \\ &= T_i (w_i \tau_{ij})^{-\theta} \sum_{n=i}^N \int_{\min\{\underline{p}_{nj}, \bar{p}_j\}}^{\min\{\underline{p}_{(n+1)j}, \bar{p}_j\}} \gamma_j^{1:n} \exp \left(-p^\theta \Phi_j^{1:n} \right) \theta p^{\theta-1} dp \end{aligned}$$

⁴⁴Letting $\bar{\varphi} \rightarrow \infty$ yields the canonical EK expression for trade shares

$$\pi_{ij} \Big|_{\bar{\varphi} \rightarrow \infty} = \frac{T_i (w_i \tau_{ij})^{-\theta}}{\Phi_j^{1:N}} = \frac{T_i (w_i \tau_{ij})^{-\theta}}{\sum_{i'=1}^N T_{i'} (w_{i'} \tau_{i'j})^{-\theta}},$$

since $\lim_{\bar{\varphi} \rightarrow \infty} \gamma_j^{1:n} = 1 \ \forall n \leq N$, and $\lim_{\bar{\varphi} \rightarrow \infty} \underline{p}_{nj} = 0 \ \forall n \leq N$.

$$= T_i (w_i \tau_{ij})^{-\theta} \times \sum_{n=i}^N \frac{\gamma_j^{1:n}}{\Phi_j^{1:n}} \left[\exp \left(-\min\{p_{nj}, \bar{p}_j\}^\theta \Phi_j^{1:n} \right) - \exp \left(-\min\{p_{(n+1)j}, \bar{p}_j\}^\theta \Phi_j^{1:n} \right) \right],$$

where $\mathbb{1}\{\cdot\}$ is an indicator function. The expression for m_{ij} is almost identical to that of π_{ij} in equation (1.14) but contains one crucial difference: the mass of intermediates per capita that country i sells to country j is smaller than π_{ij} whenever $\bar{p}_j$ is finite. For a discussion see Section 1.2.2.

1.A.A.b Trade Flows

To understand the derivation of equation (1.10) (i.e., country j 's per capita expenditure on intermediate varieties from source i), it is instructive to first derive country j 's total industry-level expenditure. This derivation suppresses the superscript k , as it is analogous across industries. Recall that only one unit of a variety is added to the final good, if that variety is used. Therefore, following equation (1.2), we need to integrate over the prices of employed varieties to obtain expenditures. An easy way to calculate total expenditure is to multiply the mass—or, quantity—of intermediates that are being used, M_j , by the average—or, expected—price of these varieties. The average price of employed varieties multiplied by $G_j(\bar{p}_j)$ is given by

$$\begin{aligned} \mathbb{E}[p_j(\omega) | p_j(\omega) \leq \bar{p}_j] \times G_j(\bar{p}_j) &= \int_{\min\{p_{1j}, \bar{p}_j\}}^{\bar{p}_j} p g_j(p) dp \\ &= \sum_{n=1}^N \gamma_j^{1:n} \int_{\min\{p_{nj}, \bar{p}_j\}}^{\min\{p_{(n+1)j}, \bar{p}_j\}} p \Phi_j^{1:n} \exp(-p^\theta \Phi_j^{1:n}) \theta p^{\theta-1} dp \end{aligned}$$

We can use integration by substitution, setting $x = p^\theta \Phi_j^{1:n}$, and thus $dx = \Phi_j^{1:n} \theta p^{\theta-1} dp$, to obtain

$$\begin{aligned}
\mathbb{E}[p_j(\omega)|p_j(\omega) \leq \bar{p}_j] G_j(\bar{p}_j) &= \sum_{n=1}^N \gamma_j^{1:n} (\Phi_j^{1:n})^{-1/\theta} \int_{\min\{\underline{p}_{nj}^\theta, \bar{p}_j^\theta\} \Phi_j^{1:n}}^{\min\{\underline{p}_{(n+1)j}^\theta, \bar{p}_j^\theta\} \Phi_j^{1:n}} x^{1/\theta} \exp(-x) dx \\
&= \sum_{n=1}^N \left[\gamma_j^{1:n} (\Phi_j^{1:n})^{-1/\theta} \right. \\
&\quad \times \tilde{\Gamma} \left(\frac{\theta+1}{\theta}, \min\{\underline{p}_{nj}^\theta, \bar{p}_j^\theta\} \Phi_j^{1:n}, \min\{\underline{p}_{(n+1)j}^\theta, \bar{p}_j^\theta\} \Phi_j^{1:n} \right) \Big],
\end{aligned}$$

where $\tilde{\Gamma}(s, \underline{x}, \bar{x}) := \int_{\underline{x}}^{\bar{x}} x^{s-1} \exp(-x) dx$. The fact that $M_j = G_j(\bar{p}_j) L_j$ gives us $X_j = \mathbb{E}[p_j(\omega)|p_j(\omega) \leq \bar{p}_j] M_j = \mathbb{E}[p_j(\omega)|p_j(\omega) \leq \bar{p}_j] G_j(\bar{p}_j) L_j$. Total expenditures by country j on intermediates (within a industry), X_j , are thus equal to the expression above multiplied by j 's population size, that is,

$$X_j = L_j \sum_{n=1}^N \gamma_j^{1:n} (\Phi_j^{1:n})^{-1/\theta} \tilde{\Gamma} \left(\frac{\theta+1}{\theta}, \min\{\underline{p}_{nj}^\theta, \bar{p}_j^\theta\} \Phi_j^{1:n}, \min\{\underline{p}_{(n+1)j}^\theta, \bar{p}_j^\theta\} \Phi_j^{1:n} \right). \quad (1.15)$$

By noting that $\tilde{\Gamma}(\frac{\theta+1}{\theta}, p, p) = 0$, we can see how equation (1.15) allows for import zeros. The derivation below makes this clearer.⁴⁵

We can calculate expenditure on intermediates from country i , X_{ij} , analogously. It is easier to first derive per capita expenditure, $x_{ij} = X_{ij}/L_j$. I multiply the average price of varieties provided by country i to destination j by m_{ij} . Recall that m_{ij} is not only the mass of intermediates but also the quantity sourced from i . I thus need to calculate the expected price conditional on i being the lowest-cost provider. Using the fact that m_{ij} is equal to the probability of i being the lowest-cost provider while $p \leq \bar{p}_j$, the expected

⁴⁵To see the connection to the standard model, let $\bar{\varphi} \rightarrow \infty$ and $\bar{p}_j \rightarrow \infty$. This yields $m_j = 1$, and

$$x_j = (\Phi_j^{1:N})^{-1/\theta} \tilde{\Gamma} \left(\frac{\theta+1}{\theta}, 0, \infty \right),$$

which is essentially the price index in the benchmark EK framework. The only difference is that there the second element on the right-hand side is $\tilde{\Gamma}(\frac{\theta-(\sigma-1)}{\theta}, 0, \infty)^{-1/(\sigma-1)}$ for $\sigma > 0$, which is a common constant across importers. Note that such a constant shifter is irrelevant to any calibration exercise or welfare calculation.

price can be calculated as

$$\begin{aligned} & \frac{1}{m_{ij} + \mathbb{1}\{m_{ij} = 0\}} \times \int_{\min\{\underline{p}_{ij}, \bar{p}_j\}}^{\bar{p}_j} p \prod_{i' \neq i} [1 - G_{i'j}(p)] g_{ij}(p) dp \\ = & \frac{1}{m_{ij} + \mathbb{1}\{m_{ij} = 0\}} \sum_{n=i}^N \gamma_j^{1:n} \frac{T_i(w_i \tau_{ij})^{-\theta}}{\Phi_j^{1:n}} \\ & \times \int_{\min\{\underline{p}_{nj}, \bar{p}_j\}}^{\min\{\underline{p}_{(n+1)j}, \bar{p}_j\}} p \Phi_j^{1:n} \exp(-p^\theta \Phi_j^{1:n}) \theta p^{\theta-1} dp \end{aligned}$$

As before, we can use integration by substitution to solve the integral. Per capita expenditures on intermediates from source country i , x_{ij} , are equal to m_{ij} times the expression above

$$\begin{aligned} x_{ij} = & T_i(w_i \tau_{ij})^{-\theta} \times \\ & \sum_{n=i}^N \frac{\gamma_j^{1:n}}{\Phi_j^{1:n}} (\Phi_j^{1:n})^{-1/\theta} \tilde{\Gamma} \left(\frac{\theta+1}{\theta}, \min\{p_{nj}^\theta, \bar{p}_j^\theta\} \Phi_j^{1:n}, \min\{p_{(n+1)j}^\theta, \bar{p}_j^\theta\} \Phi_j^{1:n} \right), \end{aligned}$$

which corresponds to equation (1.10). Total expenditures on intermediates from i are given by $X_{ij} = L_j x_{ij}$. For a discussion see Section 1.2.2.

1.A.A.c Equilibrium with General Utility Function

In this part of the appendix, I characterize the equilibrium derived in Section 1.2 for a more general additively separable utility function. The major difference to the explicit function in the main text is that here I allow for finite marginal utility of consumption at zero. With this more general specification households may not consume every product k in equilibrium.

Preferences. Households obtain utility from consuming a finite set of non-tradable final goods

$$U_j = \sum_{k=1}^K u(q_j^k),$$

with $u'(\cdot) > 0$, $u''(\cdot) < 0$, $u'(0) \leq \infty$, and $q_j^k \geq 0$.⁴⁶ The household maximizes its utility subject to a budget constraint $E_j \geq \sum_{k=1}^K \left[\int_0^1 p_j^k(\omega^k) q_j^k(\omega^k) d\omega^k \right]$. As in the main text, q_j^k is assembled according to equation (1.1), and varieties are sorted such that the intermediate's price, $p_j^k(\omega^k)$, is increasing in the index value ω^k . The “marginal variety” that is just used in the assembly process is denoted by $\bar{\omega}_j^k \in [0, 1]$. The household maximizes its utility with respect to $\{\bar{\omega}_j^k\}_k$, which yields the following first-order conditions

$$\begin{aligned} u'(q_j^k) - \lambda_j \bar{p}_j^k &\leq 0 \\ \Rightarrow q_j^k (u'(q_j^k) - \lambda_j \bar{p}_j^k) &= 0, \end{aligned}$$

where λ_j is the Lagrange multiplier, and $\bar{p}_j^k = p_j^k(\bar{\omega}_j^k)$ if $q_j^k > 0$. Note that the utility function allows for binding non-negativity constraints if $u'(0) < \infty$. In other words, it may be optimal for households not to consume the entire set of available goods. The household does not consume product k if

$$p_j^k(0) \geq \frac{u'(0)}{\lambda_j},$$

where $u'(\cdot) > 0$ implies $\lambda_j > 0$. Hence, whenever $\underline{p}_{1j}^k \geq u'(0)/\lambda_j$, we have $q_j^k = 0$.⁴⁷ For $q_j^k > 0$, let average expenditure on intermediates from industry k be given by P_j^k , as defined in equation (1.2).

Equilibrium. To characterize the equilibrium in this economy, we need to determine intermediate choke prices, $\{\bar{p}_j^k\}_{j,k}$, and the Lagrange multipliers, $\{\lambda_j\}_j$. Let final products be indexed such that the household consumes goods $k = \{1, \dots, \bar{k}\}$, whereas it does not consume $k > \bar{k}$ in equilibrium, with $\bar{k} \leq K$. For $k \leq \bar{k}$, and thus $q_j^k > 0$, the demand function is given by

$$q_j^k = u'^{-1}(\lambda_j \bar{p}_j^k).$$

⁴⁶The derivations below would be identical if $u(\cdot)$ were j - and k -specific (with the above properties). A specification with $u'(0) < \infty$ is close in spirit to that of [Hepenstrick and Tarasov \(2015\)](#). For other studies that introduce non-homothetic preferences into an EK-type environment, see, e.g., [Caron et al. \(2014\)](#); [Fieler \(2011\)](#).

⁴⁷Recall that $\underline{p}_{ij}^k := w_i \tau_{ij}^k / \bar{\varphi}$, and that exporters are sorted according to inequalities (1.6).

Multiplying both sides by P_j^k , and summing over k yields

$$\sum_{k=1}^{\bar{k}} P_j^k q_j^k = \sum_{k=1}^{\bar{k}} u'^{-1}(\lambda_j \bar{p}_j^k) P_j^k.$$

Note that the left-hand side is equal to the household's total expenditure, and therefore equal to w_j . Using $P_j^k = X_j^k / (G_j^k(\bar{p}_j^k) L_j)$ when $q_j^k > 0$, we obtain

$$w_j L_j = \sum_{k=1}^{\bar{k}} u'^{-1}(\lambda_j \bar{p}_j^k) \frac{X_j^k}{G_j^k(\bar{p}_j^k)}. \quad (1.16)$$

For a given set of wages and choke prices, equation (1.16) pins down λ_j . $G_j^k(\bar{p}_j^k)$ and X_j^k are still given by equations (1.9) and (1.11), respectively, multiplied by L_j . Choke prices are given by

$$\bar{p}_j^k = \frac{u'(G_j^k(\bar{p}_j^k))}{\lambda_j} \quad \text{for } k \leq \bar{k}. \quad (1.17)$$

For a given set of wages, Lagrange multipliers, and parameters, equation (1.17) implicitly determines $\bar{p}_j^k$ for $k \leq \bar{k}$. For $k > \bar{k}$, and thus $q_j^k = 0$, set $\bar{p}_j^k = \underline{p}_{1j}^k$.

To sum up, for given parameters θ , $\mathbf{L}$, $\bar{\varphi}$, $\mathbf{T}$, $\boldsymbol{\tau}$, the equilibrium is determined by a set of Lagrange multipliers, wages, and choke prices, $\boldsymbol{\lambda}$, $\mathbf{w}$, $\bar{\mathbf{p}}$, that satisfy the equilibrium conditions in equations (1.9)–(1.12), and (1.16)–(1.17).

1.A.A.d Equilibrium with Sectoral Linkages

In this part of the appendix, I outline a simple way to extend the model presented in Section 1.2 to allow for sectoral linkages. Final goods can now be assembled using varieties from all industries. However, intermediate varieties are still produced using labor as the only input. I denote the “input industry” by s , and the “output industry” by k .

Preferences. As in the main text, households obtain utility from consuming

a finite set of non-tradable final goods

$$U_j = \frac{\sigma}{\sigma-1} \sum_{k=1}^K (\xi_j^k)^{\frac{1}{\sigma}} (q_j^k)^{\frac{\sigma-1}{\sigma}}.$$

However, final goods are now assembled according to

$$q_j^k = \sum_{s=1}^K \alpha_j^{s,k} \ln \left[\int_0^1 q_j^{s,k}(\omega^s) d\omega^s \right],$$

where $q_j^{s,k}(\omega^s) \in \{0, 1\}$ is the quantity of variety ω^s used to assemble product k , and $\alpha_j^{s,k} > 0$ are exogenous s - k -specific productivity shifters.⁴⁸ A household first uses the cheapest varieties from s in each industry k . As in Section 1.2.1, let varieties' indexes ω be increasing in prices. Households maximize their utility with respect to $\{\bar{\omega}_j^{s,k}\}_{s,k}$, where the Lagrange function is given by

$$\begin{aligned} \mathcal{L} \left(\{\bar{\omega}_j^{s,k}\}_{s,k}, \lambda_j \right) &= \frac{\sigma}{\sigma-1} \sum_{k=1}^K (\xi_j^k)^{\frac{1}{\sigma}} \left(\sum_{s=1}^K \alpha_j^{s,k} \ln \left[\int_0^{\bar{\omega}_j^{s,k}} d\omega^s \right] \right)^{\frac{\sigma-1}{\sigma}} \\ &\quad + \lambda_j \left[E_j - \sum_{k=1}^K \sum_{s=1}^K \int_0^{\bar{\omega}_j^{s,k}} p_j^s(\omega^s) d\omega^s \right]. \end{aligned}$$

The first-order condition for $\bar{\omega}_j^{s,k}$ yields

$$(\xi_j^k)^{\frac{1}{\sigma}} (q_j^k)^{-\frac{1}{\sigma}} \frac{\alpha_j^{s,k}}{q_j^{s,k}} = \lambda_j p_j^s(\bar{\omega}_j^{s,k}),$$

⁴⁸The derivations would be analogous with a standard Cobb-Douglas function like

$$q_j^k = \prod_{s=1}^K \left[\int_0^1 q_j^{s,k}(\omega^s) d\omega^s \right]^{\alpha_j^{s,k}},$$

although the elasticities in equation (1.19) would differ. The log-specification is chosen solely for illustrative purposes.

where $q_j^{s,k} := \left(\int_0^{\bar{\omega}_j^{s,k}} d\omega^s \right)$ is the quantity of intermediates from industry s used to assemble final product k . Note that we now have an industry-pair-specific choke price, $\bar{p}_j^{s,k} := p_j^s(\bar{\omega}_j^{s,k})$. Rearranging we obtain

$$q_j^k = \xi_j^k \lambda_j^{-\sigma} \left(\alpha_j^{s,k} \right)^\sigma \left(\bar{p}_j^{s,k} q_j^{s,k} \right)^{-\sigma} \quad \forall s, k. \quad (1.18)$$

Analogous to the main text, let P_j^k be average expenditure on q_j^k , that is,

$$P_j^k := \frac{1}{q_j^k} \sum_{s=1}^K \int_0^{\bar{\omega}_j^{s,k}} p_j^s(\omega^s) d\omega^s.$$

Multiplying both sides of equation (1.18) by P_j^k , and summing over k , yields an expression for λ_j to plug back into equation (1.18). Together with the budget constraint, this yields the following expenditure shares

$$P_j^k q_j^k = E_j \frac{\xi_j^k \left(\alpha_j^{s,k} \right)^\sigma \left(\bar{p}_j^{s,k} q_j^{s,k} \right)^{-\sigma} P_j^k}{\sum_{k'=1}^K \xi_j^{k'} \left(\alpha_j^{s,k'} \right)^\sigma \left(\bar{p}_j^{s,k'} q_j^{s,k'} \right)^{-\sigma} P_j^{k'}}.$$

These conditions pin down the choke prices in equilibrium.

Equilibrium. To characterize the equilibrium in this economy, we need to determine intermediate choke prices, $\{\bar{p}_j^{s,k}\}_{j,s,k}$. Note that due to constant returns to scale of intermediate producers, the expenditures on industry s varieties for a final product k do not affect the industry- s -sales used for product k' . Therefore, for a given choke price $\bar{p}_j^{s,k}$, $q_j^{s,k}$ is simply given by $G_j^s(\bar{p}_j^{s,k})$ in equilibrium. Similarly, per capita expenditures on s -varieties to assemble product k , $x_j^{s,k}$, are given by the expression in equation (1.11) evaluated at $\bar{p}_j = \bar{p}_j^{s,k}$, that is,

$$x_j^{s,k} = \sum_{n=1}^N \gamma_j^{s,1:n} \left(\Phi_j^{s,1:n} \right)^{-1/\theta} \\ \times \tilde{\Gamma} \left(\frac{\theta+1}{\theta}, \min\{\underline{p}_{nj}^s, \bar{p}_j^{s,k}\}^\theta \Phi_j^{s,1:n}, \min\{\underline{p}_{(n+1)j}^s, \bar{p}_j^{s,k}\}^\theta \Phi_j^{s,1:n} \right).$$

Moreover, $q_j^k = \sum_{s=1}^K \alpha_j^{s,k} \ln \left[G_j^s(\bar{p}_j^{s,k}) \right]$, while overall per capita expenditures to assemble product k are given by $x_j^k := \sum_{s=1}^K x_j^{s,k}$. Using $P_j^k = x_j^k / q_j^k$, and $E_j = w_j$, we obtain

$$q_j^k = w_j \frac{\xi_j^k \left(\alpha_j^{s,k} \right)^\sigma \left(\bar{p}_j^{s,k} G_j^s(\bar{p}_j^{s,k}) \right)^{-\sigma}}{\sum_{k'=1}^K \xi_j^{k'} \left(\alpha_j^{s,k'} \right)^\sigma \left(\bar{p}_j^{s,k'} G_j^s(\bar{p}_j^{s,k'}) \right)^{-\sigma} \frac{x_j^{k'}}{q_j^{k'}}}. \quad (1.19)$$

For a given set of wages and parameters, this equation implicitly determines $\bar{p}_j^{s,k}$.

Per capita trade flows from source country i in industry s to assemble product k , $x_{ij}^{s,k}$, are given by equation (1.10) evaluated at $\bar{p}_j = \bar{p}_j^{s,k}$. The balanced trade condition that pins down wages becomes

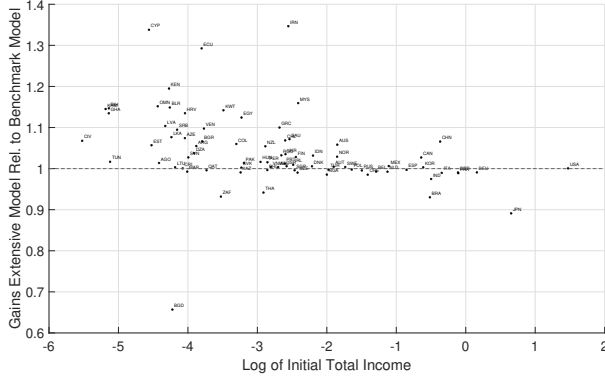
$$\sum_{i \neq j} \sum_{k=1}^K \sum_{s=1}^K L_j x_{ij}^{s,k} - L_i x_{ji}^{s,k} = 0. \quad (1.20)$$

To sum up, equations (1.19) and (1.20) pin down equilibrium choke prices and wages for a given set of parameters.

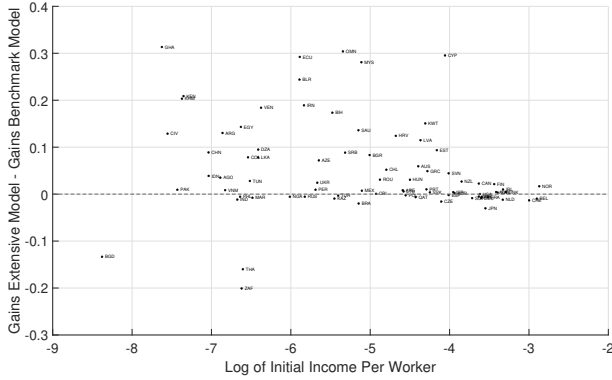
1.A.B Additional Figures

Figure 1.A.B.1: Welfare Changes after 10% Drop in Global Trade Costs – Add-On

(a) Relative Welfare Gains vs. Income



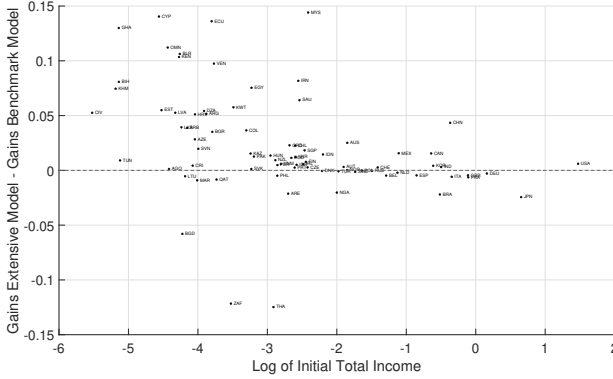
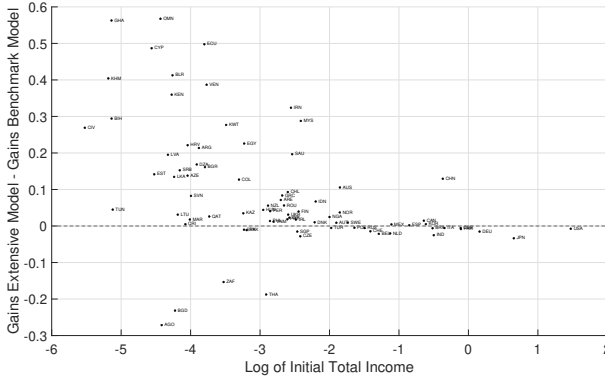
(b) Δ Welfare Gains vs. Income per Worker



Notes. This figure depicts welfare changes after a drop in global trade costs of 10%. The productivity upper bound is flexible in the extensive model, and infinite in the benchmark model. Panel (a) relates the (log) initial total income level to the welfare gains predicted by the extensive model relative to those predicted by the benchmark model. Panel (b) relates the (log) initial income level per worker, $\log(w_i)$, to the difference between the two models' welfare gains. Welfare gains are in percent, and thus the differences in Panel (b) depict percentage points. The results are discussed in Section 1.4.

Data. See Section 1.3.1.

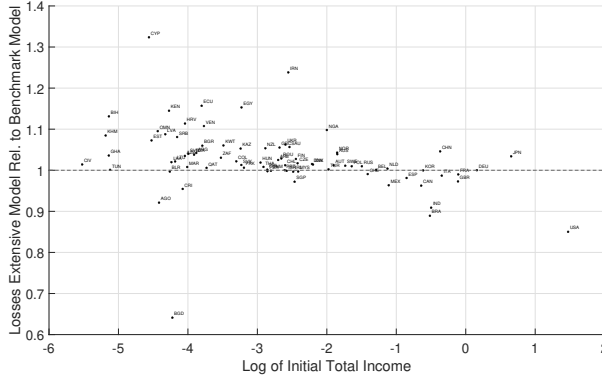
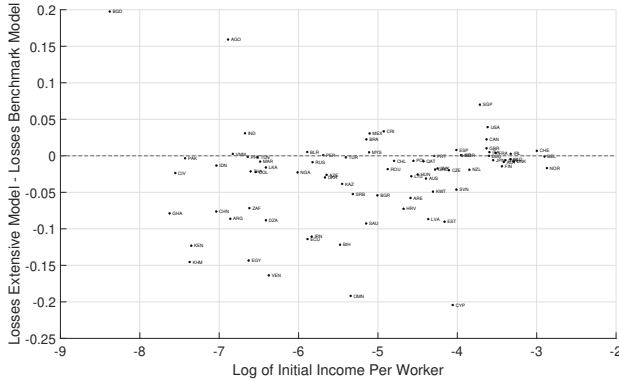
Graph. Author's representation.

Figure 1.A.B.2: Welfare Changes after Drop in Global Trade Costs**(a) 5% Decrease in τ : Δ Welfare Gains vs. Income****(b) 15% Decrease in τ : Δ Welfare Gains vs. Income**

Notes. These figures depict welfare changes after a drop in global trade costs of 5% (Panel (a)) and 15% (Panel (b)). The productivity upper bound is flexible in the extensive model, and infinite in the benchmark model. Both panels relate the (log) initial production level, $\log(X_j)$, to the difference between the two models' welfare gains. Welfare gains are in percent, and thus the differences depict percentage points. The results are discussed in Section 1.4.

Data. See Section 1.3.1.

Graph. Author's representation.

Figure 1.A.B.3: Welfare Changes after 10% Rise in Global Trade Costs – Add-On**(a) Relative Welfare Changes vs. Income****(b) Δ Welfare Changes vs. Income per Worker**

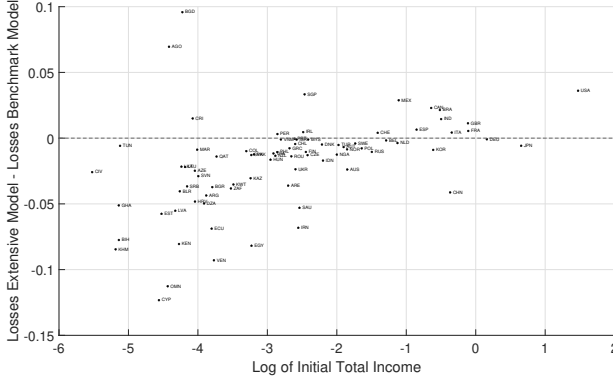
Notes. This figure depicts welfare changes after a rise in global trade costs of 10%. The productivity upper bound is flexible in the extensive model, and infinite in the benchmark model. Panel (a) relates the (log) initial total income level to the welfare changes predicted by the extensive model relative to those predicted by the benchmark model. Panel (b) relates the (log) initial income level per worker, $\log(w_j)$, to the difference between the two models' welfare losses. Welfare changes are in percent, and thus the differences in Panel (b) depict percentage points. The results are discussed in Section 1.4.

Data. See Section 1.3.1.

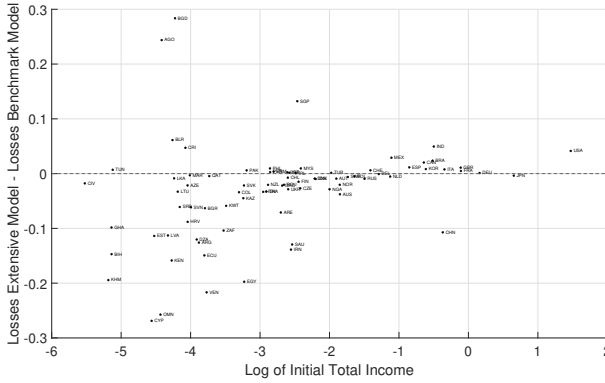
Graph. Author's representation.

Figure 1.A.B.4: Welfare Changes after Increase in Global Trade Costs

(a) 5% Increase in τ : Δ Welfare Changes vs. Income



(b) 15% Increase in τ : Δ Welfare Changes vs. Income



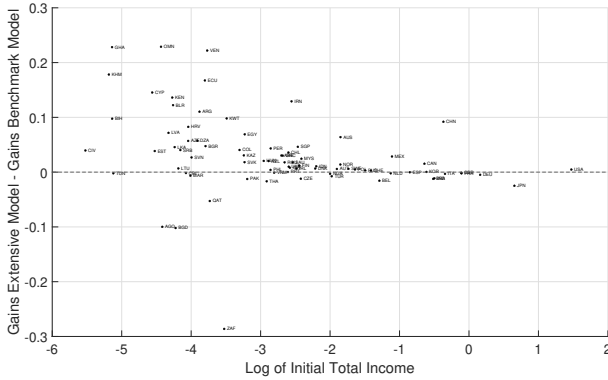
Notes. These figures depict welfare changes after a rise in global trade costs of 5% (Panel (a)) and 15% (Panel (b)). The productivity upper bound is flexible in the extensive model, and infinite in the benchmark model. Both panels relate the (log) initial production level, $\log(X_j)$, to the difference between the two models' welfare losses. Welfare changes are in percent, and thus the differences depict percentage points. The results are discussed in Section 1.4.

Data. See Section 1.3.1.

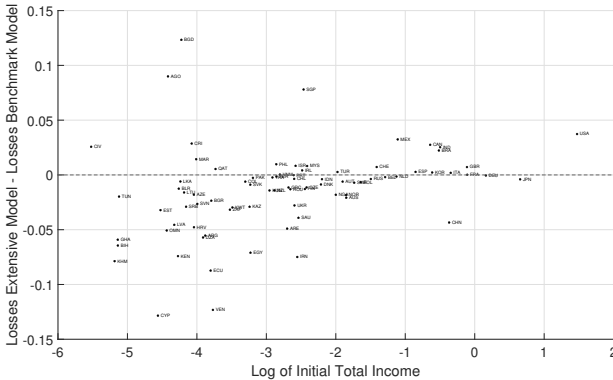
Graph. Author's representation.

Figure 1.A.B.5: Welfare Changes after Global Trade Cost Shocks – Robustness ($\sigma = 1.5$)

(a) 10% Decrease in τ : Δ Welfare Changes vs. Income



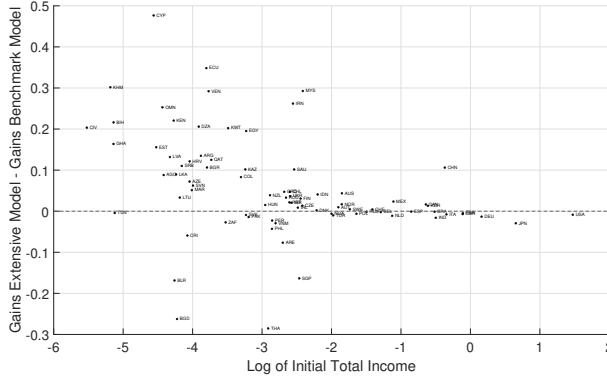
(b) 10% Increase in τ : Δ Welfare Changes vs. Income



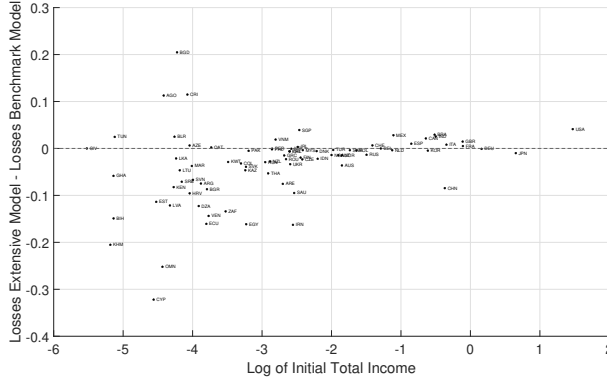
Notes. These plots are the counterparts of Figures 1.1b (Panel (a)) and 1.2b (Panel (b)) when setting $\sigma = 1.5$ instead of $\sigma = 2$ in the calibration. For details see Figures 1.1b and 1.2b.

Figure 1.A.B.6: Welfare Changes after Global Trade Cost Shocks – Robustness ($\sigma = 3$)

(a) 10% Decrease in τ : Δ Welfare Changes vs. Income



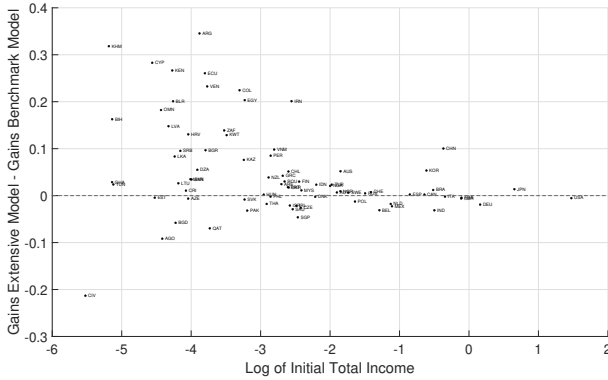
(b) 10% Increase in τ : Δ Welfare Changes vs. Income



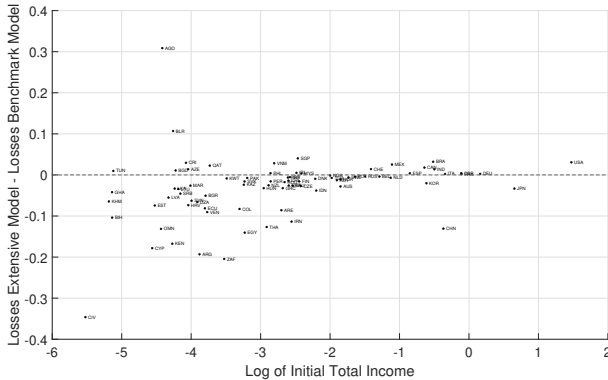
Notes. These plots are the counterparts of Figures 1.1b (Panel (a)) and 1.2b (Panel (b)) when setting $\sigma = 3$ instead of $\sigma = 2$ in the calibration. For details see Figures 1.1b and 1.2b.

Figure 1.A.B.7: Welfare Changes after Global Trade Cost Shocks – Robustness ($\theta = 6$)

(a) 10% Decrease in τ : Δ Welfare Changes vs. Income



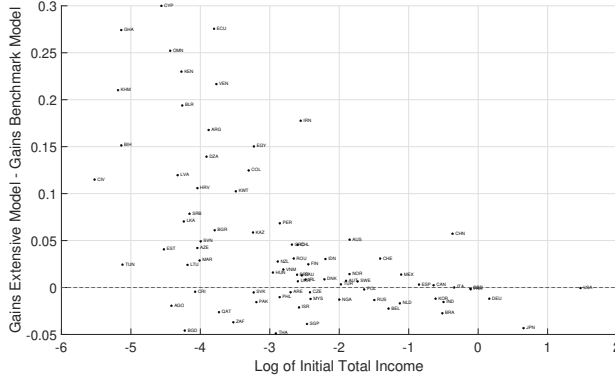
(b) 10% Increase in τ : Δ Welfare Changes vs. Income



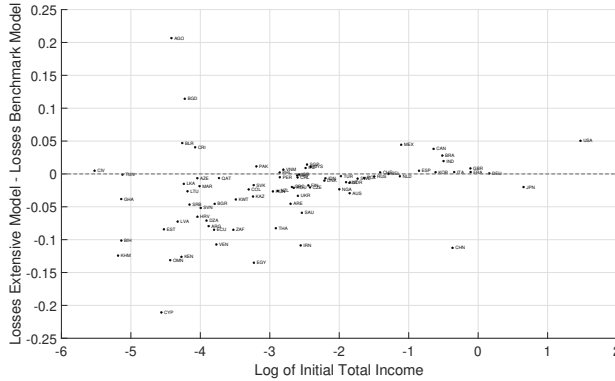
Notes. These plots are the counterparts of Figures 1.1b (Panel (a)) and 1.2b (Panel (b)) when setting $\theta = 6$ instead of $\theta = 4.14$ in the calibration. For details see Figures 1.1b and 1.2b.

Figure 1.A.B.8: Welfare Changes after Global Trade Cost Shocks – Robustness ($\theta = 5$)

(a) 10% Decrease in τ : Δ Welfare Changes vs. Income



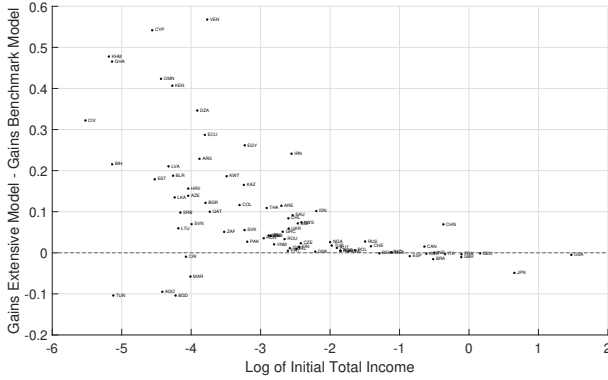
(b) 10% Increase in τ : Δ Welfare Changes vs. Income



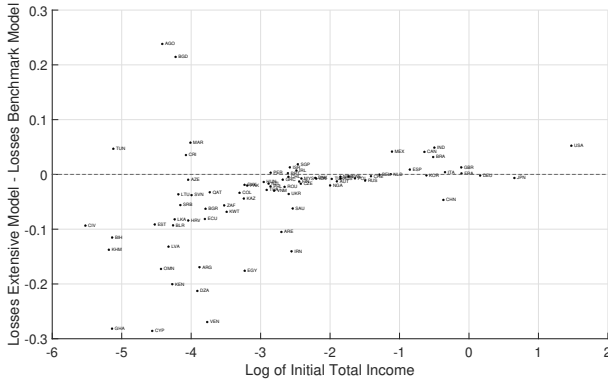
Notes. These plots are the counterparts of Figures 1.1b (Panel (a)) and 1.2b (Panel (b)) when setting $\theta = 5$ instead of $\theta = 4.14$ in the calibration. For details see Figures 1.1b and 1.2b.

Figure 1.A.B.9: Welfare Changes after Global Trade Cost Shocks – Robustness (ξ_j^k from Imports)

(a) 10% Decrease in τ : Δ Welfare Changes vs. Income



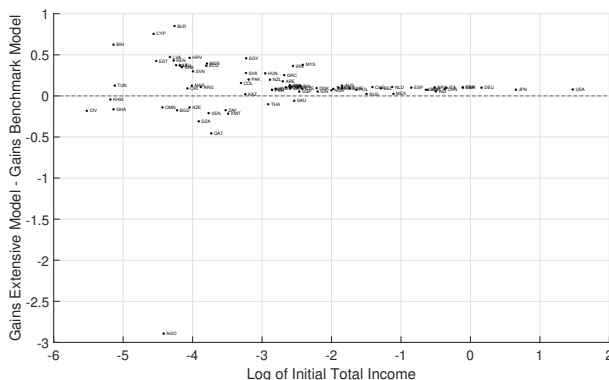
(b) 10% Increase in τ : Δ Welfare Changes vs. Income



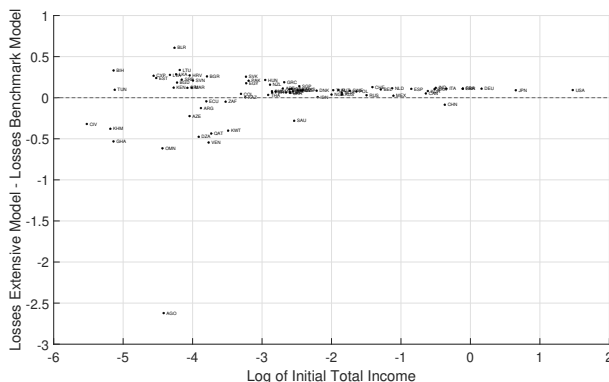
Notes. These plots are the counterparts of Figures 1.1b (Panel (a)) and 1.2b (Panel (b)) when setting ξ_j^k equal to industry k 's import share in country j . For details see Figures 1.1b and 1.2b.

Figure 1.A.B.10: Welfare Changes after Global Trade Cost Shocks – Robustness (Baseline with Trade Imbalances)

(a) 10% Decrease in τ : Δ Welfare Changes vs. Income



(b) 10% Increase in τ : Δ Welfare Changes vs. Income



Notes. These plots are the counterparts of Figures 1.1b (Panel (a)) and 1.2b (Panel (b)) when starting from a baseline with unbalanced trade as opposed to an equilibrium with zero net exports. For details see Figures 1.1b and 1.2b.

1.A.C Additional Tables

Table 1.A.C.1: The Extensive Margin of Trade – Data vs. Benchmark Model (with Cutoff)

Panel A: Data	Min.	Q1	Median	Q3	Max.
Sh. Exported Industries	0.58	0.88	0.97	0.98	1
Sh. Served Destinations ⁽ⁱ⁾	0.01	0.1	0.3	0.66	1
Sh. Imported Industries	0.94	0.97	0.98	0.98	1
Sh. Source Countries ⁽ⁱⁱ⁾	0.01	0.19	0.34	0.52	1
Nr. of Observations	7,836,800				
Panel B: Model	Min.	Q1	Median	Q3	Max.
Sh. Exported Industries	0.84	0.97	0.98	0.99	1
Sh. Served Destinations ⁽ⁱ⁾	0.01	0.28	0.62	0.91	1
Sh. Imported Industries	0.98	0.99	1	1	1
Sh. Source Countries ⁽ⁱⁱ⁾	0.01	0.37	0.62	0.78	1
Nr. of Observations	7,836,800				

Notes. This table reproduces Table 1.2 but using the “benchmark model” for Panel B instead of the “extensive model.” The “benchmark model” sets $\bar{\varphi}_i^k \rightarrow \infty \forall i, k$, thereby removing the extensive margin of trade from the full model. In both panels, trade flows below 1,000 USD are set equal to zero. Further details are given in Section 1.3.2 and the notes to Table 1.2.

⁽ⁱ⁾ Conditional on exporting in a given industry.

⁽ⁱⁱ⁾ Conditional on importing in a given industry.

Data source. [Atlas of Economic Complexity](#).

Results. Author’s computations.

Table 1.A.C.2: The Extensive Margin of Trade – Data vs. Model (with Cutoff)

Panel A: Data	Min.	Q1	Median	Q3	Max.
Sh. Exported Industries	0.58	0.88	0.97	0.98	1
Sh. Served Destinations ⁽ⁱ⁾	0.01	0.1	0.3	0.66	1
Sh. Imported Industries	0.94	0.97	0.98	0.98	1
Sh. Source Countries ⁽ⁱⁱ⁾	0.01	0.19	0.34	0.52	1
Nr. of Observations	7,836,800				
Panel B: Model	Min.	Q1	Median	Q3	Max.
Sh. Exported Industries	0.87	0.97	0.99	1	1
Sh. Served Destinations ⁽ⁱ⁾	0.01	0.08	0.24	0.61	1
Sh. Imported Industries	0.89	0.99	1	1	1
Sh. Source Countries ⁽ⁱⁱ⁾	0.01	0.16	0.33	0.54	1
Nr. of Observations	7,836,800				

Notes. This table reproduces Table 1.2 using a fixed threshold for trade flows, i.e., trade flows below 1,000 USD are set equal to zero (in both panels). Further details are given in Section 1.3.2 and the notes to Table 1.2.

⁽ⁱ⁾ Conditional on exporting in a given industry.

⁽ⁱⁱ⁾ Conditional on importing in a given industry.

Data source. [Atlas of Economic Complexity](#).

Results. Author's computations.

Table 1.A.C.3: The Extensive Margin of Trade – Data vs. PPML Fit (with Cutoff)

Panel A: Data	Min.	Q1	Median	Q3	Max.
Sh. Exported Industries	0.58	0.88	0.97	0.98	1.00
Sh. Served Destinations ⁽ⁱ⁾	0.01	0.10	0.30	0.66	1.00
Sh. Imported Industries	0.94	0.97	0.98	0.98	1.00
Sh. Source Countries ⁽ⁱⁱ⁾	0.01	0.19	0.34	0.52	1.00
Nr. of Observations	7,836,800				
Panel B: PPML	Min.	Q1	Median	Q3	Max.
Sh. Exported Industries	0.95	0.99	1.00	1.00	1.00
Sh. Served Destinations ⁽ⁱ⁾	0.01	0.38	0.76	0.99	1.00
Sh. Imported Industries	1.00	1.00	1.00	1.00	1.00
Sh. Source Countries ⁽ⁱⁱ⁾	0.01	0.51	0.68	0.84	1.00
Nr. of Observations	7,836,800				

Notes. This table reproduces Table 1.2 but using a PPML fit for Panel B instead of the “extensive model.” The PPML fit is obtained from regressing absolute trade flows on exporter-industry- and importer-industry-specific fixed effects, as well as country-pair fixed effects. In both panels, trade flows below 1,000 USD are set equal to zero. Further details are given in Section 1.3.2 and the notes to Table 1.2.

⁽ⁱ⁾ Conditional on exporting in a given industry.

⁽ⁱⁱ⁾ Conditional on importing in a given industry.

Data source. [Atlas of Economic Complexity](#).

Results. Author's computations.

**Table 1.A.C.4: Bilateral Zeros after Changes in Global Trade Costs
– Robustness**

%Change in τ	%Change in Bilateral Zeros Relative to Baseline					
	$\sigma = 1.5$	$\sigma = 3$	$\theta = 5$	$\theta = 6$	ξ_j^k Importer-Specific	Unbalanced
–10%	–5.7%	–5.9%	–6.5%	–7.1%	–6.6%	–6.1%
+10%	5.3%	5.4%	6.1%	6.9%	6.1%	5.2%

Notes. This table reports how bilateral industry-level trade zeros change after a shock to trade barriers. Each column shows the percentage change in the number of zeros relative to the baseline equilibrium. The main specification in Table 1.3 is altered via the elasticity of substitution, σ , the trade elasticity, θ , the preference shifters, ξ , or by using a baseline calibration without balanced trade (“Unbalanced”). Details are given in Section 1.4.2.

Chapter 2

Triangle Inequalities in International Trade: The Neglected Dimension

Joint with Reto Foellmi and Christian Hopenstrick

2.1 Introduction

An important driver of globalization is the reduction in trade barriers of any kind. It is a long-standing question how large these are and how much they impede the flow of goods, services and ideas between countries. To determine the size of trade barriers or trade costs in a world with many countries, we typically assume that trade costs are subject to a variant of the triangle inequality in Euclidean geometry. Trade costs between countries i and j must be lower than or equal to the trade costs between country i , an intermediate country k and country j . The reason for that is simple: a cost-minimizing

producer will send a product through the cheapest transportation channel. A positive trade flow between countries i and j directly is only possible if the triangle inequality of trade costs holds.

However basic this assumption is, the literature typically does not account for potential violations of the triangle inequality (TI) when there are shocks to the trade cost matrix. After changes in trade costs, re-routing opportunities might arise if a shipment through an intermediary becomes cheaper under counterfactual trade costs. Neglecting violations of the TI has serious consequences for the validity of structural trade models, and therefore for their quantitative predictions on the (welfare) impact of trade liberalizations.

To address the TI problem, this paper proceeds in two steps. First, we document that standard estimates of trade costs violate the TI for many bilateral trade pairs and a large part of international trade volume. It would be inconsistent with the theory to use these estimates for counterfactual exercises. Second, as a solution to the aforementioned finding, we provide a model-consistent method to estimate trade costs, satisfying the TI in a standard quantitative trade model. If the inequality were not to hold the implied trade flow between two countries would be zero, since direct sourcing from a country where the TI does not hold is not cost-minimizing. Hence, a basic premise to obtain trade costs that are logically consistent is an estimation of a trade model that is constrained by the different TIs.

We must verify the TI whenever there is a change in trade costs. This holds also in cases where we focus on relative changes without measuring absolute trade costs—an approach labeled “exact hat algebra,” which was first introduced by [Dekle et al. \(2007\)](#) and is now wide-spread in the literature.¹ To see this, consider a globalization step that results in an overall decrease of 10 percent of all bilateral trade costs. When using hat algebra, we do not need to know the initial level of trade costs. This procedure is only valid if the point of departure consists of trade costs that fulfill the TI. But even when the TI holds initially, an overall reduction of 10 percent in trade costs could create violations of the TI that are not taken into account if a standard hat algebra

¹See, e.g., [Caliendo et al. \(2019\)](#); [Caliendo and Parro \(2015\)](#); [Costinot and Rodríguez-Clare \(2014\)](#); [Eaton et al. \(2016\)](#).

approach is applied. For example, let (initial) ad-valorem trade costs to export from country i to country j be equal to τ_{ij} . The triangle inequality for this trade relation reads $\tau_{ij} \leq \tau_{ik}\tau_{kj}$ for any potential intermediary country k . If we now assume that trade costs change by a factor $\gamma < 1$, such that new trade costs equal $\gamma\tau < \tau$, the TI holds only if $\tau_{ij} \leq \gamma\tau_{ik}\tau_{kj}$. If γ is sufficiently smaller than one, this inequality is violated, implying a profitable re-routing opportunity for country pair ij .²

To solve the problem, we propose to estimate trade costs using constrained regressions with constraints satisfying the TIs. We demonstrate our method building on the seminal contribution by [Waugh \(2010\)](#). Similar to his approach we use a gravity equation to estimate trade costs via standard OLS, but in addition we impose constraints to respect the TIs. The resulting bilateral trade costs satisfy the TI and cannot be smaller than one, which is different from standard estimates, as we will show. When we consider changes in trade costs, the system is updated and potential re-routing is endogenously taken into account.

The specific examination of the TI has important quantitative implications for welfare. In one of our counterfactual exercises, we look at a global proportional drop in trade costs of 25%, as done by [Eaton and Kortum \(2012\)](#). Importantly, counterfactual iceberg trade costs are not allowed to fall below one. Our estimates suggest that, after this shock, about 13% of trade relations have at least one arbitrage opportunity, and would thus be better off by involving an intermediary. The trade-weighted share is almost twice as large. In a framework à la [Eaton and Kortum \(2002\)](#) focusing on aggregate trade flows, this has substantial implications for the welfare gains that arise from re-routing. We find that most countries benefit from the option to re-route. The third of the sample with the largest welfare differences gains on average about 5% more real income. To put this into perspective, we scale gains from re-routing by the well-known welfare losses that would arise from a complete trade shutdown. The average return in our sample from re-routing is equal to 17.5% of the net welfare loss from going to autarky. Our results

²Furthermore, we do not know whether trade costs hit their theoretical lower bound (i.e., $\tau_{ij} = 1$) after the trade cost shock if we do not estimate their initial levels.

suggest that accounting for arbitrage opportunities after shocks to trade costs becomes more important the larger the shock size. This indicates a sort of “local stability” of the trade cost matrix.³

We show that our method can be easily applied to more complex models. We repeat our counterfactual exercises using the powerful multi-sector extension of the [Eaton and Kortum \(2002\)](#) model with input-output linkages by [Caliendo and Parro \(2015\)](#). The lesson from these additional analyses is twofold. First, in smaller samples re-routing is less likely to be profitable. This is intuitive, since arbitrage opportunities are more likely the larger our set of potential intermediaries. Second, although we observe less re-routing with fewer countries in our sample, involving intermediaries tends to be highly profitable when sectoral linkages are included in the analysis. After a 25% drop in bilateral sector-level trade costs, the average welfare gain increases by about 20% once we allow for re-routing.

In economic geography, it is argued that an improvement in one segment of a transportation infrastructure may alter the shipping costs along the entire network (e.g., [Allen and Arkolakis \(2020\)](#) analyze changes in the U.S. highway network). The reason lies in the fact that, e.g., lower travel time between two cities may lead to changes of travel paths—re-routing—in the whole network. We argue in the paper that similar mechanisms are (implicitly) present in canonical international trade models when we consider the welfare effects of trade cost changes. A related study in this regard is that of [Ganapati et al. \(2020\)](#), who show that entrepôts create an incentive to ship via indirect routes.

While the TI received little attention in quantitative trade models, it plays a central role in trade policy. An exporter from a third country i could have an incentive to ship a product to country j via country k , taking advantage of a favorable bilateral free trade agreement between countries k and j when there does not exist a similar treaty between i and j . Sophisticated rules-of-

³Qualitatively, the arguments we make also apply to models that feature variable and fixed trade costs, such as the [Melitz \(2003\)](#) model (see also Chapter 3). However, quantitatively results may well differ, as our trade cost estimates are a compound measure of iceberg and fixed costs through the lens of a [Melitz \(2003\)-Chaney \(2008\)](#) framework. The welfare gains from re-routing which we measure after a shock to trade costs (cf. Section 2.4) would therefore not necessarily be the same in those models.

origin clauses in free trade agreements, typically based on value-added shares in each country, should preclude such trade deflection. However, [Felbermayr et al. \(2019\)](#) show that trade deflection to take advantage of tariff differentials is not profitable, as tariff differences do not outweigh increased transportation costs. Their argument underlies that the TI of trade costs should hold when we estimate trade costs, and that tariff differences (even if they could be exploited at all) are not sufficiently large to create violations of the TI.

The paper is organized as follows. In [Section 2.2](#), we introduce the theory. [Section 2.3](#) elaborates the estimation strategy, and [Section 2.4](#) discusses counterfactual exercises. [Section 2.5](#) concludes.

2.2 Theory

Consider the following simple theoretical structure. The world consists of N countries and many goods. Each of the countries purchases its products from the respective cheapest source in the world. This basic feature holds in a broad set of trade models, for instance in a Ricardian trade model à la [Eaton and Kortum \(2002\)](#) (henceforth, EK).

Transportation sector. Trade is costly and takes place within a transportation sector that ships goods from country i to country j . The transportation sector is characterized by perfect competition. This sector uses the particular good as the only input, no labor is employed in transportation. Following the literature, we assume that transportation costs are proportional to the marginal costs of the good: if marginal costs of production in country i equal c_i , the marginal costs of producing a good in country i and selling the item to country j equal $\tau_{ij}c_i \geq c_i$, with $\tau_{ij} \geq 1$. Thus, we make the common assumption that international trade entails constant ad valorem trade costs—known as “iceberg” trade cost ([Samuelson, 1954](#)). Trade costs are a c.i.f. measure, hence they account for trade costs from border i to border j . Therefore it is consistent to set within-country trade costs $\tau_{jj} = 1$. In this setting, the

triangle inequality (TI) of trade costs takes the familiar form⁴

$$\tau_{ij} \leq \tau_{ik} \tau_{kj}.$$

The TI requires that the trade cost of shipping the product via some intermediary k , τ_{ik} times τ_{kj} , may not be lower than the costs of shipping the product directly from country i to j , τ_{ij} . If the TI were violated, a cost-minimizing producer would choose to source the product through the trade cost minimizing path.⁵ Assume there exists at least one intermediary k where $\tau_{ij} > \tau_{ik} \tau_{kj}$. We define

$$\tau_{ij}^1 = \min_k \{\tau_{ik} \tau_{kj}\}, \quad (2.1)$$

where $\tau_{ij}^1 = \tau_{ij}$ when $\arg \min_k \{\tau_{ik} \tau_{kj}\} \in \{i, j\}$. Note that it suffices to look at three-countries triangle violations, and to replace trade costs that violate the TI with their updated value τ^1 iteratively. To see this, assume trade costs are lower if the good is shipped via two intermediaries k and k' , but not if it is shipped via only one of them, i.e., $\tau_{ij} > \tau_{ik} \tau_{kk'} \tau_{k'j}$ but $\tau_{ij} < \tau_{ik} \tau_{kj}$ and $\tau_{ij} < \tau_{ik'} \tau_{k'j}$. Taking the former two inequalities together—obviously we get the identical result when using the latter inequality—, we see that the violation of the TI originates from a three-countries triangle violation. We obtain $\tau_{kj} > \tau_{kk'} \tau_{k'j}$, trade from k to j is cheaper via k' . In a violation where n intermediaries are involved, we apply the same reasoning to reduce the inequality to a violation where $n - 1$ intermediaries are involved. As the number of countries is finite, the number of potential multi-intermediary paths

⁴If within-country trade costs are non-zero (e.g., for retail distribution), total trade costs between i and j would amount to $\tau_{ij} \tau_{jj}$. This leaves the TI unchanged as we would have to compare $\tau_{ij} \tau_{jj} \leq \tau_{ik} \tau_{kj} \tau_{jj}$ in this case, assuming that within-country distribution costs have not to be paid in intermediate country k . For an analysis of internal trade frictions see, e.g., [Sotelo \(2020\)](#).

⁵In the words of [Eaton and Kortum \(2005, p. 172\)](#), “arbitrage would eliminate violations of the TI since k would emerge as an entrepôt, so that [using our notation] $\tau_{ij} = \tau_{ik} \tau_{kj}$.” Note further that we are searching for a baseline trade cost matrix that is theoretically consistent rather than exploring reasons why violations might actually be present in the data. The presence of preferential trade agreements could be such a reason, as argued, e.g., by [Mossay and Tabuchi \(2015\)](#).

for a trade flow between countries i and j is finite as well.⁶

Applying the procedure laid out in equation (2.1) to updated trade costs, we obtain $\tau_{ij}^2 = \min_k \left\{ \tau_{ik}^1 \tau_{kj}^1 \right\}$. We repeat this until all violations of the three-countries triangle inequality are eliminated, and therefore violations with $n > 1$ intermediaries are eliminated too. This leads to the following definition.

Definition 1 *We define $\tilde{\tau}_{ij} = \min_k \{ \tilde{\tau}_{ik} \tilde{\tau}_{kj} \}$, where $\forall \{i, j\}$ it holds that $\arg \min_k \{ \tilde{\tau}_{ik} \tilde{\tau}_{kj} \} \in \{i, j\}$.*

We will achieve the optimal $\tilde{\tau}$ after a finite number of iteration steps as in equation (2.1). The resulting updated trade cost between two countries follows an optimal trade trajectory along a finite set of countries. Formally, the optimal trade cost is a product of initial trade costs, i.e., $\tilde{\tau}_{ij} = \prod_{u=1}^{u=n-1} \tau_{k_u k_{u+1}}$ with $k_1 = i$, $k_n = j$, and $n \geq 2$. We summarize this insight in the following proposition.

Proposition 1 *Consider a trade cost matrix T with elements $\{\tau_{ij}\}$, where $i, j \in \{1, 2, \dots, N\}$. There exists a unique trade cost matrix $\tilde{T}$ where all trade costs satisfy the triangle inequality $\tilde{\tau}_{ij} \leq \tilde{\tau}_{ik} \tilde{\tau}_{kj}$.*

Proof Existence is shown in the text. To see that $\tilde{T}$ is unique, assume there exists a second matrix $\tilde{T}'$ with at least one $\tilde{\tau}'_{ij} \neq \tilde{\tau}_{ij}$. Note that this would violate the definition of $\tilde{\tau}_{ij} = \min_k \{ \tilde{\tau}_{ik} \tilde{\tau}_{kj} \}$. $\square$

Counterfactuals. It is important to note what this implies for the structural estimation of trade models. Since the particular good is the input in the transportation sector, we can simply replace τ_{ij} by $\tilde{\tau}_{ij}$ in the welfare calculations of trade policies. In particular, it does not matter that the physical trade flow between i and j goes through intermediary k .⁷ If $\tilde{\tau}_{ij} = \tilde{\tau}_{ik} \tilde{\tau}_{kj}$, updated trade costs are the same whether trade takes the direct route between i and

⁶Note further that multiple re-shipping via the same intermediary can never be optimal, as trade costs are theoretically bounded to satisfy $\tau_{ij} \geq 1$, and thus $\tau_{ij}^x \geq \tau_{ij}$ for any constant $x \geq 1$.

⁷Recall that the transportation sector is perfectly competitive, and thus there are no rents from being involved as an intermediary.

j , or goes via the intermediary. If the total amount of trade between i and j equals X , all values between 0 and X for direct trade between i and j are consistent with the theory (with the remainder going through k). Hence, if we use the updated trade cost matrix, $\tilde{T}$, we can continue to work with standard gravity equations in simulations, even when the initial matrix exhibits violations of the TI.

Corollary 1 *The gravity equations of trade flows remain functionally unchanged with the trade cost matrix $\tilde{T}$ replacing T .*

What does this result imply for empirical studies of trade cost changes? Many counterfactual experiments make use of the fact that workhorse trade models allow us to estimate relative changes in economic outcomes as a function of relative changes in trade costs and initial trade shares only. There is thus no need to calibrate the level of trade costs (or size-related parameters such as the state of technology and population size). This simple and powerful tool is known as “exact hat algebra,” where the “exact” refers to the fact that no approximations are involved in the estimation process. Consistent with the theoretical setup, one assumes that the TIs are satisfied in the initial equilibrium. However, as the level of trade costs is only implicitly present in standard structural formulae, we do not know whether, after a shock to the trade cost matrix, profitable re-routing opportunities arise.⁸ Consider, e.g., a trade pair where the TI holds initially, i.e., $\tau_{ij} < \tau_{ik}\tau_{kj}$. Let us retake our thought experiment from the introduction and assume there is an overall reduction of bilateral trade costs by a factor $\gamma < 1$. Obviously, there exists a value of γ such that $\gamma\tau_{ij} > \gamma^2\tau_{ik}\tau_{kj}$. This implies that a simple hat-algebra approach, relying solely on changes in trade costs, could lead to logically inconsistent results, since violations of the TI must be taken into account in any simulation. In other words, potential second-order effects of trade cost changes on the transportation sector are neglected. Moreover, we cannot be sure whether γ is large enough for τ_{ij} to hit its lower bound (i.e., $\tau_{ij} = 1$). To ensure

⁸Counterfactuals involving trade costs are unique in this respect. Other shocks such as changes in technology parameters will not encounter the issues discussed in this paper.

a model-consistent counterfactual that respects potential spillover effects, we therefore need to estimate the initial level of τ_{ij} , and after a shock update the trade cost matrix, as outlined in Proposition 1, in case any TI is violated.⁹ Before turning to the quantitative importance of TIs in Section 2.4, the next section describes our approach to estimate τ_{ij} and compute counterfactuals.

2.3 Methodology

According to Section 2.2, we need to calibrate the initial trade cost matrix in order to verify whether violations of the TI arise after a shock to trade costs. In this section, we illustrate how we estimate a trade cost matrix that is consistent with the TIs. The underlying off-the-shelf EK frameworks are standard in the literature, and thus only briefly outlined in Appendix 2.A.A.

TI violations. There are several contributions that estimate iceberg trade costs in variants of the EK framework (e.g., [Eaton and Kortum, 2002](#); [French, 2016](#); [Novy, 2013](#); [Parro, 2013](#); [Simonovska and Waugh, 2014](#); [Waugh, 2010](#)). To our knowledge, however, we are the first to plug these estimates into the TIs to check that they are satisfied. We find that all of the listed methods generate a trade cost matrix that features violations of the TI. The share of trade relations which would be better off shipping via a third country varies between 8% and 50%. Moreover, the calibrated τ 's are not restricted to be greater than or equal to one, and indeed some estimates fall below this theoretical lower bound. Through the lens of the model, this would then imply that value is created—instead of lost—in transit. A trade cost matrix that does not fully satisfy the TIs is inconsistent with the theory, and thus cannot be employed to compute counterfactuals.¹⁰

⁹Note that the widely-used reference point of welfare losses under complete autarky is an exception. With $\tau_{ij} \rightarrow \infty \forall i \neq j$, there cannot be any violation of the TI. See [Arkolakis et al. \(2012\)](#) for a detailed discussion on welfare losses from autarky implied by standard trade models. More generally, whenever all trade costs are raised by the same factor (i.e., $\gamma > 1$ in the example above), we will not encounter violations if the initial trade cost matrix was consistent with the TIs.

¹⁰The TI violations found using [Waugh's \(2010\)](#) method are discussed in Appendix 2.A.C, since our estimation will be based on this procedure.

Trade cost estimates. To overcome the encountered issues, we augment [Vaugh's \(2010\)](#) influential method to ensure that the calibrated τ 's will obey the TI, and satisfy $\tau_{ij} \geq 1 \forall i, j$. The reason we choose [Vaugh \(2010\)](#) is twofold. First, these estimates help the model mimic empirically observed bilateral trade flows and prices. Second, the calibrated trade costs are entirely based on OLS coefficients, making the implementation very simple.

As detailed in Appendix [2.A.A.a](#), the EK model yields a log-linear relationship between trade flows (scaled by home absorption) and trade costs, τ_{ij} , as well as importer- and exporter-specific terms

$$\log \left(\frac{\pi_{ij}}{\pi_{jj}} \right) = \log (T_i w_i^{-\theta}) - \log (T_j w_j^{-\theta}) - \theta \log (\tau_{ij}), \quad (2.2)$$

where π_{ij} is exporter i 's share in importer j 's total expenditure, $T_i w_i^{-\theta}$ is a measure of country i 's productivity, and θ the so-called trade elasticity. Let $S_j^m := -\log (T_j w_j^{-\theta})$, where the superscript m is chosen to indicate that this expression will be captured by the importer fixed effect (FE).

The functional form imposed to estimate trade costs is a key assumption. In line with [Vaugh \(2010\)](#), trade costs are assumed to be a log-linear function of an exporter-specific shifter, ex_i , as well as bilateral dummy variables, that is,

$$-\theta \log (\tau_{ij}) = ex_i + \sum_{l=1}^D \beta^l d_{ij}^l, \quad (2.3)$$

where β^l is the coefficient on a dummy variable d_{ij}^l . Examples for dummy variables are common border, common language, or distance bins. The shifter ex_i captures all exporter-specific aspects affecting trade costs that are not picked up by the dummy variables. Let $S_i^x := ex_i - S_i^m$, where the superscript x is chosen to indicate that this expression will be captured by the exporter FE.

By plugging equation (2.3) into equation (2.2), and applying the definitions

of S_i^x and S_j^m , we obtain our gravity regression

$$\log\left(\frac{\pi_{ij}}{\pi_{jj}}\right) = S_i^x + S_j^m + \sum_{l=1}^D \beta^l d_{ij}^l + \varepsilon_{ij}, \quad (2.4)$$

where ε_{ij} denotes the error term.¹¹ The dependent variable is directly observed in the data. The elements on the right-hand side of equation (2.4) are an exporter FE, an importer FE, and a set of bilateral dummy variables. Note that the exporter-specific component of trade barriers in equation (2.3) is simply the sum of the exporting country's exporter and importer FEs. In the absence of ex_i , the exporter and importer FEs would add up to zero according to the theory. Empirical deviations from this theoretical target will thus be attributed to trade costs. The higher the exporter-specific shifter, the lower trade costs are.

As in [Waugh \(2010\)](#), our baseline regression includes dummy variables indicating whether the country pair shares a common border, and whether the geographical distance between the two countries lies within a specific distance interval (cf. Table 2.A.C.1). In Appendix 2.A.C, we run this regression on a sample of 127 countries in the year 2004. We show that this procedure can yield estimates that violate the TI, and sometimes fall below one. The data are described in Appendix 2.A.B.

What leads to these violations of the theory? Let t_{ij} be the estimate of τ_{ij} , which is calculated as

$$-\hat{\theta} \log(t_{ij}) = ex_i + \sum_{l=1}^D \hat{\beta}^l d_{ij}^l,$$

where we set $\hat{\theta} = 4.14$, which is [Simonovska and Waugh's \(2014\)](#) preferred estimate, and $\hat{\beta}^l$ denotes the estimate of β^l . The TI takes a log-linear form,

¹¹Alternatively, one could add the error term to the trade cost function in equation (2.3). In Appendix 2.A.C, we show that this leads to even more TI violations. We chose the specification in equation (2.3) because this enables us to drastically reduce the number of constraints required in our constrained estimation, as explained below.

and can therefore be rewritten as

$$\log(t_{ij}) \leq \log(t_{ik}) + \log(t_{kj}).$$

Multiplying by $-\hat{\theta}$, and plugging in the expression above, we receive

$$0 \geq ex_k + \sum_{l=1}^D \hat{\beta}^l (d_{ik}^l + d_{kj}^l - d_{ij}^l) \quad \forall i, k, j. \quad (2.5)$$

Inequality (2.5) nicely illustrates how TI violations may emerge.¹² To gain intuition, suppose that the triplet ikj does not share a common border, and all nodes fall within the same geographical distance bin l . Inequality (2.5) reduces to $0 \geq ex_k + \hat{\beta}^l$. The additional distance cost from involving intermediary k is equal to $\hat{\beta}^l$ (distance has a negative impact on trade flows, i.e., $\hat{\beta}^l < 0$). If country k is a relatively strong exporter (i.e., ex_k is high), engaging this intermediary may outweigh the additional distance penalty, and thus yield estimates that suggest an arbitrage opportunity (i.e., $ex_k + \hat{\beta}^l > 0$).

Similarly, to have $t_{ij} \geq 1$, we need the following “lower-bound” condition to hold

$$0 \geq ex_i + \sum_{l=1}^D \hat{\beta}^l d_{ij}^l \quad \forall i, j. \quad (2.6)$$

In the modeling framework outlined above, inequalities (2.5) and (2.6) are necessary and sufficient conditions to make sure that the TI holds and the trade costs are limited from below. Next, we outline how to perform estimations consistent with the TIs.¹³

A model-consistent calibration. Standard gravity regressions that other-

¹²Note that the long debated size of θ is irrelevant to the TI. It will matter, however, in the welfare calculations below.

¹³Recall from Section 2.2 that checking re-routing via multiple intermediaries is not necessary (as long as $t_{ij} \geq 1$), so that inequalities (2.5) and (2.6) suffice to guarantee that there are no arbitrage opportunities in the trade cost matrix. Requiring only one of the two inequalities is not enough.

wise follow the theory closely neglect the additional implicit constraints (2.5) and (2.6). Our approach is a model-consistent calibration that imposes the latter. We minimize the sum of squared residuals subject to a set of linear inequality constraints. Accordingly, the solution to this convex optimization problem is a global minimum. However, the number of constraints is substantial. Let N be the number of countries in the sample. For every potential intermediary k , we have $(N - 1) \times (N - 2)$ TIs, as well as $(N - 1)$ lower-bound conditions (inequality (2.6)). The theory therefore implicitly requires a total of $N \times (N - 1)^2$ constraints—over two million in our sample with 127 countries. A “brute force” solution to the issues above would be to impose all of them in the estimation, which slows down the computation enormously.¹⁴

We can instead use the fact that ex_k does not vary across importers. Moreover, since the trade cost function contains exclusively dummy variables, t_{ij} may draw only a limited number of unique values.¹⁵ For a given ex_k , the right-hand side of inequality (2.6) can take up at most 2^D unique values. Similarly, the “loss term” $d_{ik}^l + d_{kj}^l - d_{ij}^l$ has four potential combinations. Inequality (2.5) thus has at most 4^D unique values for a given ex_k . This yields a total of $N \times (4^D + 2^D)$ unique constraints, which is similar to $N \times (N - 1)^2$ in our baseline regression with $N = 127$ and $D = 7$. However, we can strongly reduce the number of constraints, considering that most combinations are empirically irrelevant. For instance, the distance between countries i and j can obviously only fall within one specific interval. This implies that inequality (2.6) will take up at most 12 ($= 2 \times 6$) rather than 128 ($= 2^7$) unique values for a given ex_k . Out of these remaining 12 combinations several are again redundant, since countries share a border only with geographically close neighbors. In a similar vein, we can trim the 4^D unique values in inequality (2.5). For example, if countries i and j are in the top distance bin while i and k are in the bottom interval, it must be that the distance between k and j falls within one of the two top bins. Following this pruning procedure we heavily reduce

¹⁴The rest of this section is only relevant to understanding our constrained estimation but in no way essential to understand the remainder of the paper.

¹⁵Analogous to distance, we can easily accommodate continuous control variables by dividing them into bins, and including dummy variables that indicate into which bin an observation falls.

Table 2.1: Constrained vs. Unconstrained Coefficients for Trade Cost Estimates

	Coefficients		%-effect on t_{ij} with $\theta = 4.14$	
	Unconstrained	Constrained	Uncons.	Cons.
Distance [0, 375)	-3.565	-4.479	136.6	195
Distance [375, 750)	-4.323	-4.768	184.1	216.4
Distance [750, 1500)	-5.415	-5.641	269.9	290.6
Distance [1500, 3000)	-6.851	-6.804	423.2	417.3
Distance [3000, 6000)	-7.861	-7.77	567.8	553.3
Distance [6000, max.]	-8.532	-8.396	685.3	659.9
Contiguity	1.26	0.579	-26.2	-13.1
Nr. of observations	13,722	13,722		

Notes. This table reports OLS results for the gravity equation (2.4) in Section 2.3 using the unconstrained (first and third columns) and constrained (second and fourth columns) versions. Regressions include importer and exporter fixed effects. The constrained version imposes inequalities (2.5) and (2.6) to ensure that estimates satisfy the TI. The left-most column replicates the first column in Table 2.A.C.1. Distance is weighted by population. Distance intervals are in miles. Trade flows are measured relative to domestic absorption. Dependent variable is log of trade shares. t_{ij} is the estimate of τ_{ij} . The percentage effect on t_{ij} is equal to $100 \times (\exp(\text{coeff.}/(-\theta)) - 1)$. The method and results are discussed in Section 2.3.

Data. See Appendix 2.A.B.

Results. Authors' computations.

the number of constraints from over two million to about ten thousand (or, by a factor of 200). The time required to compute the constrained regression drops from several hours to a few seconds. Importantly, most of these linear inequality constraints will not be binding in the estimation (only twelve in our main specification), and thus it is not an issue that the number of constraints is still close to the number of observations.¹⁶

Comparison. Let $\tilde{t}_{ij}$ be our *constrained* estimate of τ_{ij} , which we obtain by running equation (2.4) while imposing the constraints described above. By construction, none of these estimates violates the TI, and we have $\tilde{t}_{ij} \geq 1 \forall i, j$. To compare the constrained and unconstrained values, we set $t_{ij} = 1$ whenever we get $t_{ij} < 1$.

¹⁶The procedure is the same irrespective of whether we use OLS or PPML. We use OLS for all our main results.

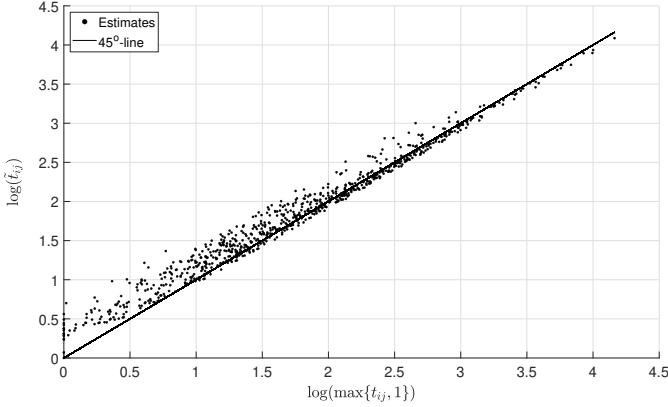


Figure 2.1: Standard Trade Cost Estimates vs. Constrained Trade Cost

Notes. This figure depicts the log of estimated trade costs vs. their constrained counterparts. As detailed in Section 2.3, we estimate trade costs via both unconstrained and constrained OLS. The regression is described by equation (2.4). The constrained version imposes inequalities (2.5) and (2.6) to ensure that estimates satisfy the TI. t_{ij} ($\tilde{t}_{ij}$) are the unconstrained (constrained) estimates of τ_{ij} . If an unconstrained estimate falls below one, we set it equal to one (i.e., $t_{ij} \geq 1$).

Data. See Appendix 2.A.B.

Graph. Authors' representation.

Table 2.1 reports the coefficients from the unconstrained and constrained regressions, as well as their percentage impact on trade costs. The most apparent difference is that of the coefficients on the border dummy. The suggested benefit on trade costs more than halves to ensure that the TI remains satisfied. The impact of distance also flattens, with a notable drop in the gain from being in the lowest distance bin rather than in the second lowest one. By contrast, the values of ex_i stay very similar, although we see a decline for very strong exporters (cf. Figure 2.A.D.2 in Appendix 2.A.D).

These insights are reflected in Figure 2.1, where we plot t_{ij} against $\tilde{t}_{ij}$. The constrained estimation pushes up low values of t_{ij} , whereas it leaves higher values largely unchanged, and, if anything, slightly decreases them. Intuitively, close countries cannot have trade barriers that are so low that involving a

neighbor as an intermediary is beneficial. As a last step, we verify that the ad hoc solution of taking unconstrained estimates and computing post-estimation updated values following Proposition 1 does not yield the same results. In Figure 2.A.D.3 of Appendix 2.A.D, we show that there is a substantial difference between $\tilde{t}_{ij}$ and updated values of t_{ij} . Note that the updated value of t_{ij} is, by construction, always weakly smaller than t_{ij} , whereas this is clearly not the case with $\tilde{t}_{ij}$. In fact, constrained trade costs are generally closer to unconstrained ones than to the updated values.

2.4 Counterfactual Exercises

We want to understand the impact of the TI, working through potential re-routing of trade flows, on welfare gains from trade liberalizations. For that aim, we conduct counterfactual experiments where we lower trade costs. We compute welfare gains (i.e., relative changes in real income) using hat algebra as explained in Appendix 2.A.A, and we do so in two ways.¹⁷ First, we shock the constrained trade cost matrix and do not verify whether TI violations arise (standard approach). Second, we check the TI and update trade costs whenever a new route allows for a cheaper delivery of the good. The relative change in trade costs is then the *effective* drop in trade barriers that materializes after the shock, which takes potential spillover (re-routing) effects into account. We do not allow trade costs to fall below one in either case.

To make our argument precise, we set the initial value of τ_{ij} equal to its constrained estimate, $\tilde{t}_{ij}$, so that the starting point features no arbitrage opportunity. For instance, let the counterfactual exercise be a proportional decrease in trade costs of 25%. The ratio of counterfactual to initial trade costs, denoted by $\hat{\tau}_{ij} = \tau'_{ij}/\tau_{ij}$, equals 75%, as long as the new trade cost τ'_{ij} does not fall below one. Formally, the first step sets $\hat{\tau}_{ij} = \max\{0.75, 1/\tau_{ij}\} \forall i \neq j$,

¹⁷We do not explicitly account for trade imbalances in these experiments, since this dimension is not directly relevant to our discussion. A simple way to do so would be to include an additional parameter which captures net exports in the initial equilibrium (see Dekle et al., 2007).

Table 2.2: TI Violations after 25% Drop in Global Trade Costs

	No restriction	Restrict $t \geq 1$
% of trade pairs with arbitrage opportunity	21%	13%
... trade-weighted	32%	24%
Distribution of potential savings:		
Q1: 1%, Median: 5%, Mean: 8%, Q3: 15%, Max: 36%		
Top 10 hubs (desc. order):		
DEU SGP THA MYS ARE USA SAU FRA ITA CHN		

Notes. This table reports violations of the triangle inequality that emerge after we decrease global bilateral trade costs by 25%. The right-most column sets the new trade cost values equal to one if they drop below one. The second row shows the share in total trade that has a potential arbitrage opportunity. The third row refers to reductions in trade costs from re-routing when there is an arbitrage opportunity (restricting $t \geq 1$). A “hub” refers to an intermediary which allows for a cheaper delivery than the direct shipment between two countries. The results are discussed in Section 2.4.

Data. See Appendix 2.A.B.

Results. Authors’ computations.

and computes welfare gains as described in Appendix 2.A.A.¹⁸ The second step verifies whether counterfactual trade costs, $\tau'_{ij} = \hat{\tau}_{ij}\tau_{ij}$, satisfy the TI. In case there are arbitrage opportunities, we follow Proposition 1 to obtain updated values, $\tilde{\tau}'_{ij}$. The welfare gains are then computed in the standard way but setting $\hat{\tau}_{ij} = \tilde{\tau}'_{ij}/\tau_{ij}$, i.e., we use the *effective* change in trade costs taking the option of re-routing explicitly into account.

Uniform drop in τ . We use the single-sector setup outlined in Appendix 2.A.A.a, and decrease global trade costs proportionally by 25%—an exercise conducted by Eaton and Kortum (2012).

How many arbitrage opportunities arise? Table 2.2 reports the share of trade pairs that faces at least one violation of the TI when trade costs drop by a quarter. Observe that some values would fall below one if we did not

¹⁸This procedure slightly differs from the standard hat algebra approach, which would set $\hat{\tau}_{ij} = 0.75 \forall i \neq j$. However, we need to impose the theoretical lower bound, otherwise we would obtain extreme results when allowing for re-routing.

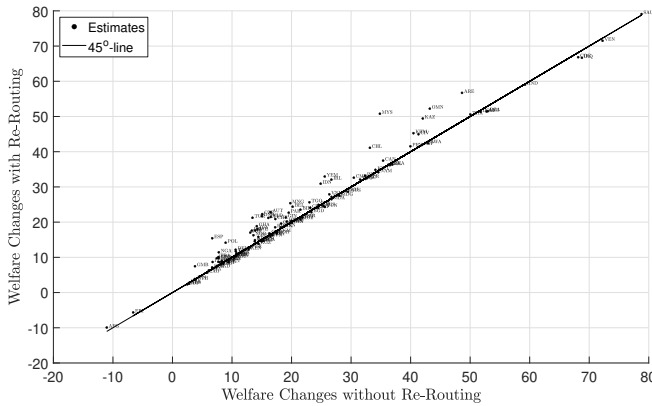


Figure 2.2: Welfare Changes after 25% Drop in Global Trade Costs

Notes. This figure depicts welfare changes (in %) after a drop in global trade costs of 25%. In one case, we do not allow for re-routing (standard approach), and in the other we allow for re-routing when the TI is violated. In both cases, we do not allow for trade costs to fall below one. The results are discussed in Section 2.4.

Data. See Appendix 2.A.B.

Graph. Authors' representation.

impose this lower bound. But even with this restriction, about 13% of trade pairs—or 24% of total trade—can supposedly decrease their delivery costs by involving a third country. The implied savings potential for these relations is meaningful, considering that trade costs can on average drop by a further 8%. Thus, the actual reduction in trade barriers is considerably larger for several trade relations when we account for potential second-order effects.

The welfare implications are depicted by Figure 2.2. Intuitively, countries that are close to the main hubs in Table 2.2, such as Spain or Malaysia, benefit most from re-routing. The results suggest that Spain could more than double its welfare gain by re-routing in response to this uniform globalization experiment. There are also countries which lose through this step, but the losses are usually negligible. As is well-known in the literature, the one-sector model generally yields relatively low gains from trade (see Costinot and Rodríguez-Clare (2014) and Ossa (2015) for a detailed discussion). It thus makes sense to put these numbers into perspective. The welfare loss from going to full

autarky (i.e., $\tau_{ij} \rightarrow \infty \forall i \neq j$) is a widely-used benchmark, not least because calculating it is straightforward (see Appendix 2.A.A.a).

In Table 2.A.E.5 of Appendix 2.A.E, we list the 40 countries with the largest return from re-routing. Column (3) reports the relative real income loss that would result if a country were to completely shut down international trade, which serves as a benchmark. In Column (4), we scale the additional income gain from re-rerouting (i.e., Column (1) minus Column (2)) by the absolute value of Column (3). The extra gains are substantial relative to the benchmark values. For instance, Spain can collect an additional 8% of its initial real income by exploiting arbitrage opportunities. This is double the reference number of a 4% loss in the case of a complete trade collapse. Although the magnitude of the latter strikes as unrealistic, it shows how, through the lens of the model, the implied welfare differences are sizable. These countries can on average extract almost half of their benchmark real income changes by involving intermediaries. This exercise nicely illustrates why it is important to consider potential spillovers even in counterfactuals that involve uniform trade cost changes.

Table 2.A.E.6 in Appendix 2.A.E shows that verifying the existence of arbitrage opportunities becomes more important the larger the shock to trade barriers is, suggesting that the trade cost matrix is “locally stable.” Repeating the exercise from above for a 5% drop in global trade costs yields only few re-routing opportunities. By contrast, involving intermediaries creates a median net welfare change equal to 22.5% when we halve trade costs. On average, the net real income change from re-routing divided by the net loss from going to autarky is huge in the latter case, and non-negligible in the former scenario. Nevertheless, these statistics hide a lot of heterogeneity, as the more detailed results above show.

Asymmetric shock to τ . We also compute a counterfactual with non-uniform trade cost changes. Specifically, following [Vaugh \(2010\)](#), we eliminate bilateral asymmetries in trade barriers by setting the counterfactual trade costs equal to the minimum of the (potentially different) bilateral trade costs;

Table 2.3: TI Violations after Eliminating Asymmetries in Trade Costs

	No restriction	Restrict $t \geq 1$
% of trade pairs with arbitrage opportunity	58%	58%
... trade-weighted	7%	7%
Distribution of potential savings:		
Q1: 24%, Median: 45%, Mean: 50%, Q3: 70%, Max: 197%		
Top 10 hubs (desc. order):		
DEU USA SGP SAU ARE ZAF CAN CHN ARG ITA		

Notes. This table reports violations of the triangle inequality that emerge after we eliminate asymmetries in trade costs (i.e., we set $\tilde{t}_{ij} = \min\{\tilde{t}_{ij}, \tilde{t}_{ji}\}$). The right-most column sets the new trade cost values equal to one if they drop below one. The second row shows the share in total trade that has a potential arbitrage opportunity. The third row refers to reductions in trade costs from re-routing when there is an arbitrage opportunity (restricting $t \geq 1$). A “hub” refers to an intermediary which allows for a cheaper delivery than the direct shipment between two countries. See Section 2.4 for further details.

Data. See Appendix 2.A.B.

Results. Authors’ computations.

formally $\tau'_{ij} = \min\{\tau_{ij}, \tau_{ji}\}$. A policy change that would come close to an equalization of bilateral trade costs would be the abolishment of artificial, policy-related bilateral distortions. Note that trade costs cannot fall below one due to this shock, but nonetheless the TI may not hold anymore.¹⁹ In fact, Table 2.3 shows that more than half of the trade relations now face at least one arbitrage opportunity, with an impressive extra reduction of 50% on average. The winners in this counterfactual scenario are relatively small countries, which is reflected in the low trade share that would benefit from re-

¹⁹How do re-routing opportunities arise? Suppose that for the triplet ikj we have $ex_k > ex_i \geq ex_j$. By eliminating bilateral asymmetries (which stem exclusively from the exporter-specific trade cost components), inequality (2.5) becomes

$$0 \geq \underbrace{(ex_k - ex_i)}_{>0} + ex_k + \sum_{l=1}^D \hat{\beta}^l (d_{ik}^l + d_{kj}^l - d_{ij}^l),$$

which will be violated if inequality (2.5) was (close to) binding to begin with. In words, country k is likely to become a hub if it is a relatively strong exporter.



Graph. Authors' representation.

Multi-sector model. We can easily repeat the two main simulations above within the multi-sector framework with sectoral input-output linkages outlined in Appendix 2.A.A.b. The reasoning above on transportation and potential arbitrage opportunities applies entirely to multi-sector environments as well.

The model yields a sector-level equivalent to the aggregate gravity equation. We can thus apply the method from Section 2.3 to estimate sector-specific iceberg trade costs, τ_{ij}^k , using sector-specific fixed effects and elasticities (i.e., having a sector-specific equation (2.4)). This in turn enables us to compute counterfactuals with and without the option of re-routing.

Our sector-level data cover only 34 countries, and therefore there will be fewer potential intermediaries a country pair may select than in the single-sector counterfactuals above. This should dampen the impact of re-routing opportunities compared to the previous analyses. For completeness, Table 2.A.E.8 in Appendix 2.A.E reproduces the symmetric counterfactual (see Table 2.A.E.5) using the WIOD data instead of the ITPD-E data. Indeed the gains from re-routing after the symmetric shock to trade costs tend to be considerably smaller in the reduced sample.

What happens when the sectoral dimension is added? Trade cost changes affect the equilibrium in a more complex way when input-output linkages are taken into account, and therefore additional reductions in trade costs may have strong effects.²⁰ Table 2.4 replicates Table 2.A.E.5 using the multi-sector setup, i.e., we decrease all bilateral sector-level trade costs by 25%. The share of trade relations which face at least one arbitrage opportunity after the shock varies substantially across sectors, ranging from 0% to 66%, with a median value of 11%. The option to re-route has a sizable impact on welfare, with the average real income gain rising by about 20%. This is considerably larger than the gains implied by the one-sector counterpart (see Table 2.A.E.8).

The suggested benefit from re-routing after eliminating bilateral asymmetries in trade barriers is huge in both the single-sector (Table 2.A.E.9) and the multi-sector version (Table 2.A.E.10), although at a higher level in the latter case. This result is largely driven by the ROW becoming a hub—which ought to be expected, as 40% of bilateral trade comes from or flows to ROW, making it the strongest exporter, and thus an attractive intermediary (footnote 19). The ROW appears to lose substantially when we introduce the re-routing option in the single-sector model. Nevertheless, interpreting ROW as a hub is

²⁰The sectoral dimension tends to generate larger welfare changes in general, which is well-understood in the literature (see, e.g., Costinot and Rodríguez-Clare, 2014).

Table 2.4: Welfare Changes after 25% Drop in Global Trade Costs (Multi-Sector)

Country	(1) With re-routing	(2) No re-routing	(3) Autarky loss	(4) ((1)-(2))/ (3)
AUS	16%	15%	-16%	3%
AUT	30%	24%	-53%	11%
BEL	36%	33%	-56%	6%
BRA	12%	10%	-7%	20%
CAN	18%	15%	-32%	8%
CHN	13%	12%	-8%	22%
CZE	40%	33%	-36%	21%
DEU	18%	16%	-22%	6%
DNK	26%	20%	-41%	14%
ESP	16%	12%	-19%	21%
FIN	29%	25%	-21%	19%
FRA	14%	10%	-18%	17%
GBR	18%	14%	-23%	18%
GRC	16%	16%	-24%	1%
HUN	41%	34%	-56%	13%
IDN	23%	17%	-13%	48%
IND	13%	11%	-10%	26%
IRL	58%	42%	-38%	44%
ITA	15%	12%	-17%	17%
JPN	8%	6%	-5%	40%
KOR	23%	19%	-14%	26%
MEX	17%	14%	-19%	18%
NLD	28%	24%	-43%	8%
POL	26%	21%	-35%	15%
PRT	23%	21%	-39%	7%
ROU	23%	19%	-27%	14%
RUS	20%	18%	-38%	5%
ROW	15%	13%	-30%	4%
SVK	44%	38%	-53%	11%
SVN	36%	29%	-60%	11%
SWE	28%	23%	-26%	19%
TUR	14%	11%	-28%	12%
TWN	28%	28%	-18%	-4%
USA	7%	6%	-8%	12%
Average	23%	19%	-28%	16%

Notes. This table reports welfare changes after a decrease in global bilateral sector-level trade costs of 25%, calculated using the multi-sector model. Column (1) allows for re-routing after the shock. Column (3) reports the real income loss from going to full autarky (i.e., $\tau_{ij}^k \rightarrow \infty \forall i \neq j$, and k). Column (4) shows the difference between Columns (1) and (2) scaled by the absolute value of Column (3). See Section 2.4 for further details.

Data. See Appendix 2.A.B.

Results. Authors' computations.

somewhat abstract, and we therefore tried excluding it from the sample. This reduces the gains from re-routing, but they nonetheless remain substantial (see Tables 2.A.E.12 and 2.A.E.14). In the symmetric counterfactual, ROW plays a way smaller role (see Tables 2.A.E.11 and 2.A.E.13).

In a nutshell, spillovers from trade cost changes appear to be an important factor to consider even in small samples. This seems to hold especially in analyses with multiple sectors.

2.5 Conclusion

We show how the triangle inequality (TI) of trade costs plays a crucial role for counterfactual analyses in general equilibrium trade models. We provide a method to estimate trade costs that satisfy the TI. Standard estimation methods do not take into account potential violations of the TI, and often yield estimates that imply arbitrage opportunities. The latter are inconsistent with the theory, and should thus not be employed in simulations.

We use our estimates to calculate the welfare gains that arise from re-routing after a shock to the trade cost matrix. Using both uniform and asymmetric reductions in trade costs, we show that counterfactual trade cost matrices feature many arbitrage opportunities. That is, several trade pairs could further reduce their delivery costs by involving an intermediary country after the shock. The effective drop in trade costs is therefore larger for these dyads than the “direct” reduction in trade barriers suggests. The TI—and thus re-routing—is quantitatively important, in that welfare gains are substantially higher for many countries when we check for arbitrage opportunities.

2.A Appendices to Chapter 2

2.A.A Model Setup

The Ricardian frameworks we employ are picked off the shelf from well-known contributions in the literature. The only difference being our handling of trade cost (changes). First, we outline the one-sector [Eaton and Kortum \(2002\)](#) model, and then introduce the multi-sector version with input-output linkages.

2.A.A.a Single-Sector EK

We briefly describe the [Eaton and Kortum \(2002\)](#) version with one sector and labor as the only input (see, e.g., [Simonovska and Waugh, 2014](#)). For a detailed discussion of the model consult these papers.

Setup. This static world consists of N countries, indexed by i and j . There is a single tradable final-goods sector, and a continuum of varieties indexed by $\omega \in [0, 1]$. Labor is supplied inelastically by a measure of L_i consumers. The representative agent has CES utility over a bundle of final goods with elasticity of substitution $\sigma > 1$.

Firms operate in perfect competition and employ a constant returns to scale technology, where the cost of producing a unit of good ω in country i is given by $w_i/z_i(\omega)$. w_i is the economy's wage rate, and $z_i(\omega)$ denotes country i 's efficiency in producing good ω , which is independently drawn from a Fréchet distribution

$$F_i(z) = \exp(-T_i z^{-\theta}).$$

For one unit of the good to arrive in destination j , producers in country i need to ship $\tau_{ij} \geq 1$ units, as described in Section 2.2. We normalize $\tau_{jj} = 1 \forall j$, and assume that the triangle inequality holds, i.e., $\tau_{ij} \leq \tau_{ik}\tau_{kj} \forall i, j, k$.

Due to perfect competition, country i charges for good ω its marginal cost

of production and transportation to consumers in j , that is,

$$p_{ij}(\omega) = \frac{\tau_{ij} w_i}{z_i(\omega)},$$

and consumers pick the source which offers the lowest price. The productivity distribution coupled with this pricing behavior yields the CES price index²¹

$$P_j = \left[\sum_{k=1}^N T_k (w_k \tau_{kj})^{-\theta} \right]^{-\frac{1}{\theta}},$$

and the gravity equation

$$\pi_{ij} = \frac{X_{ij}}{X_j} = \frac{T_i (w_i \tau_{ij})^{-\theta}}{\sum_{k=1}^N T_k (w_k \tau_{kj})^{-\theta}},$$

where π_{ij} is the so-called trade share, X_{ij} is the amount spent by country j 's consumers on products from source i , and X_j is country j 's total expenditure on final goods. In Section 2.3, we work with trade shares that are normalized by domestic absorption, that is,

$$\frac{\pi_{ij}}{\pi_{jj}} = \frac{T_i w_i^{-\theta} \tau_{ij}^{-\theta}}{T_j w_j^{-\theta}}.$$

To close the model, we assume that trade is balanced (i.e., every country has net exports equal to zero). As labor is the only source of income in this model, the set of wages satisfies the following condition

$$w_i L_i = \sum_{j=1}^N \pi_{ij} w_j L_j,$$

where we have $X_j = w_j L_j$.

Hat algebra. How does the equilibrium change after a shock to trade costs?

²¹We abstract from constant terms which will be irrelevant to the entire discussion.

Let the value of a variable x in the counterfactual equilibrium be equal to x' , and let $\hat{x} := x'/x$. Using this “hat-notation,” we can depict changes in outcomes as a function of relative trade costs, $\hat{\tau}_{ij}$, and initial expenditure levels, X_{ij} . This way we avoid calibrating the technology shifters, T_i . It is easy to show that the relevant set of equilibrium equations becomes

$$\begin{aligned}\pi'_{ij} &= \frac{\pi_{ij} (\hat{w}_i \hat{\tau}_{ij})^{-\theta}}{\sum_{k=1}^N \pi_{kj} (\hat{w}_k \hat{\tau}_{kj})^{-\theta}} \\ \hat{w}_i X_i &= \sum_{j=1}^N \pi'_{ij} \hat{w}_j X_j.\end{aligned}$$

After solving for the set of $\{\hat{w}_i\}$ that satisfies this equation system, we can calculate welfare changes (i.e., changes in real wages). More precisely, the gravity equation implies

$$\pi_{ii} = T_i \left(\frac{w_i}{P_i} \right)^{-\theta},$$

which means that welfare changes depend only on the trade elasticity, θ , and changes in the home share, π_{ii} ,²²

$$\frac{\hat{w}_i}{\hat{P}_i} = (\hat{\pi}_{ii})^{-\frac{1}{\theta}}.$$

A useful benchmark is the real income loss from going to autarky (i.e., $\tau_{ij} \rightarrow \infty \forall i \neq j$). In that case we obtain

$$\frac{\hat{w}_i^a}{\hat{P}_i^a} = (\pi_{ii})^{\frac{1}{\theta}},$$

because we know that $\pi_{ii}^a = 1$, and thus $\hat{\pi}_{ii}^a = 1/\pi_{ii}$.

²²For more details on this type of counterfactuals in quantitative trade models, see [Costinot and Rodríguez-Clare \(2014\)](#). For a discussion on the equilibrium’s uniqueness, see [Alvarez and Lucas \(2007\)](#).

2.A.A.b Multi-Sector EK

We use the multi-sector extension of the model above from [Caliendo and Parro \(2015\)](#). The only difference to [Caliendo and Parro \(2015\)](#) is that we do not explicitly model tariff revenues and trade imbalances to simplify exposition. Below we briefly sketch the setup.

Setup. We have K sectors, indexed by k and s , with sectoral input-output linkages. Labor is mobile across sectors. Households derive utility from the consumption of final goods, C_i^k . Preferences are given by a Cobb-Douglas aggregator

$$u(C_i) = \prod_{k=1}^K (C_i^k)^{\alpha_i^k}, \quad \text{where} \quad \sum_{k=1}^K \alpha_i^k = 1.$$

Firms produce varieties using two inputs, labor and a composite of intermediate goods. The value added share in sector k is denoted by $\gamma_i^k \geq 0$. The share of materials from sector k used in country i 's sector s is denoted by $\gamma_i^{k,s} \geq 0$, with $\sum_{k=1}^K \gamma_i^{k,s} = 1 - \gamma_i^s$. The cost of an input bundle, c_i^k , is given by

$$c_i^k = w_i^{\gamma_i^k} \prod_{s=1}^K (P_i^s)^{\gamma_i^{s,k}},$$

where we again abstract from constant terms that will be irrelevant to the entire discussion.

Iceberg trade costs, τ_{ij}^k , and the parameters of the Fréchet distribution, T_i^k and θ^k , are sector-specific, which coupled with perfect competition yields the price of the composite intermediate good

$$P_j^k = \left[\sum_{i=1}^N T_i^k (c_i^k \tau_{ij}^k)^{-\theta^k} \right]^{-\frac{1}{\theta^k}},$$

and the gravity equation

$$\pi_{ij}^k = \frac{T_i^k (c_i^k \tau_{ij}^k)^{-\theta^k}}{(P_j^k)^{-\theta^k}}.$$

Finally, the aggregate price index is equal to

$$P_i = \prod_{k=1}^K \left(\frac{P_i^k}{\alpha_i^k} \right)^{\alpha_i^k},$$

and total household income, I_i , is given by

$$I_i = w_i L_i = \sum_{k=1}^K \gamma_i^k \sum_{j=1}^N \pi_{ij}^k X_j^k,$$

where X_j^k is total expenditure on goods from sector k , including households' final consumption as well as expenditure by firms on composite intermediate goods. Note that this model reduces to the one in Appendix 2.A.A.a if we set $K = 1$ and $\gamma_i = 1 \forall i$.

Hat algebra. Analogous to the procedure above, we solve the model in relative changes after a shock to trade costs. The relevant system of equations is given by

$$\begin{aligned} \hat{c}_i^k &= \hat{w}_i^{\gamma_i^k} \prod_{s=1}^K (\hat{P}_i^s)^{\gamma_i^{s,k}} \\ \hat{P}_j^k &= \left[\sum_{i=1}^N \pi_{ij}^k (\hat{c}_i^k \hat{\tau}_{ij}^k)^{-\theta^k} \right]^{-\frac{1}{\theta^k}} \\ (\pi_{ij}^k)' &= \pi_{ij}^k \frac{(\hat{c}_i^k \hat{\tau}_{ij}^k)^{-\theta^k}}{(\hat{P}_j^k)^{-\theta^k}} \\ (X_i^k)' &= \sum_{s=1}^K \gamma_i^{k,s} \sum_{j=1}^N (\pi_{ij}^s)' (X_j^s)' + \alpha_i^k \hat{w}_i I_i \end{aligned}$$

$$\sum_{k=1}^K \sum_{j=1}^N (\pi_{ij}^k)' (X_j^k)' - (\pi_{ji}^k)' (X_i^k)' = 0.$$

For a given guess of wage changes $\{\hat{w}_i\}$, we solve for (changes in) price indices and total expenditure, and then use the balanced trade condition to update the vector of relative wages, thereby adopting the tâtonnement algorithm from [Alvarez and Lucas \(2007\)](#). Relative changes in real wages, and thus welfare changes, are calculated as follows

$$\begin{aligned} \log \left(\frac{\hat{w}_i}{\hat{P}_i} \right) = & - \sum_{k=1}^K \frac{\alpha_i^k}{\theta^k} \log (\hat{\pi}_{ii}^k) \\ & - \sum_{k=1}^K \frac{\alpha_i^k}{\theta^k} \frac{1 - \gamma_i^k}{\gamma_i^k} \log (\hat{\pi}_{ii}^k) - \sum_{k=1}^K \frac{\alpha_i^k}{\gamma_i^k} \log \left(\prod_{s=1}^K \left[\frac{\hat{P}_i^s}{\hat{P}_i^k} \right]^{\gamma_i^{s,k}} \right). \end{aligned}$$

For the multi-sector version, we require more data compared to the single-sector model above. Specifically, we need sector-specific trade shares, π_{ij}^k , input shares, $\gamma_i^{k,s}$, value added shares, γ_i^k , expenditure shares in final demand, α_i^k , and value added, $w_i L_i$. Following [Costinot and Rodríguez-Clare \(2014\)](#), we obtain all these values from the data described in Appendix 2.A.B. The computation also requires sectoral trade elasticities, θ^k , which we take from [Caliendo and Parro \(2015\)](#) (see Appendix 2.A.B).

2.A.B Data

Aggregate data. For the one-sector model, we use the International Trade and Production Database for Estimation ([ITPD-E](#)). The ITPD-E provides bilateral trade flows and production data retrieved from several administrative data sources (see [Borchert et al. \(2021\)](#) for details). Production data is used to calculate domestic trade as well as trade shares. The dataset covers the years 2000–2016, and up to 243 countries and 170 industries. For each country pair, we sum over all industries to obtain aggregate bilateral trade flows. We include only countries that appear as both importers and exporters in a given year, and we drop countries with a total population below one million. We merge

China and Hong Kong, as well as Belgium and Luxembourg, which is often done in the literature due to large shares of re-exports. Our main results use the year 2004.

The gravity covariates such as distance and contiguity are taken from the widely-used [CEPII Gravity database](#) ([Head et al., 2010](#)).

Sector-level data. Despite its richness, the ITPD-E does not cover input-output linkages. Thus, our sector-level analysis uses the well-known 2013 release of the WIOD [World Input-Output Tables](#) ([Timmer et al., 2015](#)). This dataset covers 40 countries and a model for the “rest of the world” (ROW). It includes 35 sectors and their input-output links over the period 1995–2011. We add Luxembourg to Belgium, and following [Costinot and Rodríguez-Clare \(2014\)](#), we merge several smaller countries to the ROW: Bulgaria, Cyprus, Estonia, Latvia, Lithuania, and Malta. According to [Costinot and Rodríguez-Clare \(2014\)](#), because of classification issues across countries, there are sectors for which countries have zero consumption and production. To avoid issues in this regard, we use the year 2008 and apply the same sector aggregation as [Costinot and Rodríguez-Clare \(2014\)](#), called “basic aggregation” in Table 3 of their online appendix ([Costinot and Rodríguez-Clare, 2013](#)). These steps leave us with 34 countries and 31 sectors. We use the sector-level trade elasticities that are listed in Table 3 of their online appendix ([Costinot and Rodríguez-Clare, 2013](#)), which were taken from [Caliendo and Parro \(2015, Table 1\)](#).

For the regressions, we use the median distance between a specific country in the sample and ROW-countries. The dummy variable indicating contiguity is set equal to zero between ROW and the other countries in the sample.

2.A.C Violations of the Triangle Inequality

In this section, we present violations of the TI found using [Vaugh’s \(2010\)](#) procedure to estimate τ , as discussed in Section 2.3. Let t_{ij} be the estimate of τ_{ij} . We calculate trade costs for the year 2004 using the data described in Appendix 2.A.B, which gives us a sample of 127 countries. The regression

results are reported in Table 2.A.C.1.²³

TI violations. For every pair, we plug t_{ij} into the TI, and verify whether this no-arbitrage condition is violated for at least one intermediary country. The results are presented in Table 2.A.C.2. We find that 36% of trade relations could reduce their delivery costs by shipping via a third country, where most pairs have several potential intermediaries that are worth involving. Intuitively, the importers which would most benefit from an intermediary are geographically close to the top hubs (e.g., Switzerland and Vietnam). There are 34 relations where we obtain $t_{ij} < 1$, which would imply that trade costs could be pushed toward zero by re-shipping multiple times. But even when we set these t 's equal to one, we see that 20% of relations would still have an arbitrage opportunity. This does not only concern small relations, as 28% of total trade is affected (when $t_{ij} \geq 1$). In Figure 2.A.D.1 of Appendix 2.A.D, we plot the trade cost estimates against their “updated” values, where the latter are computed following Proposition 1. The potential savings for pairs with an arbitrage opportunity are substantial, with the average log-difference amounting to 23%.²⁴

Are the violations driven by only a handful of hubs? In Panel B of Table 2.A.C.2 we exclude the three intermediaries that would minimize trade costs for the most relations (Germany, Singapore, and Thailand), and rerun the regressions without them. The resulting estimates still feature many violations, despite a notable decline. Even when the ten top hubs are excluded—the U.S. being among them—about 6% of trade pairs would be better off involving an intermediary (cf. Panel C). It is important to note that these hubs are not solely the intuitive suspects (such as the Netherlands or Singapore).

²³The coefficients look very similar to those of Simonovska and Waugh (2014), who run the same regression on a sample with 123 countries in the year 2004.

²⁴Despite this relatively large average difference, one might still worry that potential savings are not statistically significantly different from zero. We can easily construct confidence bands around t_{ij} using the delta method. For pairs with an arbitrage opportunity, we find that the difference between indirect costs, $t_{ik}t_{kj}$, and direct costs, t_{ij} , is significantly smaller than zero at the five percent level in roughly 80% of the cases. Given that the literature focuses on point estimates and rarely works with confidence bands, we did not investigate this further.

Table 2.A.C.1: Standard Coefficients for Trade Cost Estimates

	OLS	PPML
Distance [0, 375)	-3.565*** (0.352)	-2.978*** (0.621)
Distance [375, 750)	-4.323*** (0.321)	-3.099*** (0.580)
Distance [750, 1500)	-5.415*** (0.316)	-3.964*** (0.566)
Distance [1500, 3000)	-6.851*** (0.311)	-5.538*** (0.582)
Distance [3000, 6000)	-7.861*** (0.312)	-6.246*** (0.590)
Distance [6000, max.]	-8.532*** (0.314)	-6.455*** (0.585)
Contiguity	1.260*** (0.118)	1.122*** (0.215)
Nr. of observations	13,722	16,002

Notes. This table reports OLS and PPML results for the gravity equation (2.4) in Section 2.3. Robust standard errors are reported in parentheses. Regressions include importer and exporter fixed effects. Distance is weighted by population. Distance intervals are in miles. Trade flows are measured relative to domestic absorption. Dependent variable is absolute (log of) trade shares for PPML (OLS). The results are discussed in Appendix 2.A.C.

*** significant at 1%-level.

Data. See Appendix 2.A.B.

Results. Authors' computations.

Robustness. We test different specifications to verify that the pattern we found is persistent. Table 2.A.E.1 of Appendix 2.A.E replicates the analysis in Table 2.A.C.2, the only difference being that we add four dummy variables to the trade cost function which indicate whether a country pair (i) shares a common official language, (ii) a common currency, (iii) ever had a colonial link, or (iv) has a regional trade agreement. The results are very similar, with the share of TI violations slightly increasing.

In Table 2.A.E.2, we add the error term in equation (2.4) to the trade cost

Table 2.A.C.2: TI Violations with Standard Trade Cost Estimates

	Baseline	Restrict $t \geq 1$
Panel A: Full sample		
% of trade pairs with arbitrage opportunity	36%	20%
... trade-weighted	51%	28%
Top 10 hubs (desc. order): DEU SGP THA ARG CHN ARE USA MYS GBR PAN		
Panel B: Excl. top 3 hubs		
% of trade pairs with arbitrage opportunity	23%	13%
Panel C: Excl. top 10 hubs		
% of trade pairs with arbitrage opportunity	10%	6%

Notes. This table reports violations of the triangle inequality implied by trade cost estimates, t , that we computed using [Vaugh's \(2010\)](#) method. In the right-most column, we set $t = 1$ whenever we find $t < 1$. Panel A uses the full sample, the second row showing the share in total trade that has a potential arbitrage opportunity. Panels B and C exclude the three and ten main hubs, respectively, found in the full sample. A “hub” refers to an intermediary which allows for a cheaper delivery than the direct shipment between two countries. Details are given in [Appendix 2.A.C](#).

Data. See [Appendix 2.A.B](#).

Results. Authors' computations.

function. This table uses only trade relations with positive trade flows, as we cannot calculate trade costs for the remaining pairs in this specification.²⁵ We obtain even more TI violations when perfectly matching the observations with our trade cost estimates. In this specification, we may interpret the error term as a shock to trade costs rather than measurement error ([Simonovska and Vaugh, 2014](#)).

In [Table 2.A.E.3](#), we reproduce [Table 2.A.C.2](#) using PPML instead of OLS, as recommended by [Santos Silva and Tenreiro \(2006\)](#). The results are similar, although PPML appears to alleviate the issue. Especially the share of trade that is affected drops substantially. However, there are still implied arbitrage

²⁵Alternatively, we could set $t_{ij} \rightarrow \infty$ for $\pi_{ij} = 0$. This would, however, lead to even more alleged arbitrage opportunities.

opportunities left even when we exclude the ten major hubs. The correlation between the OLS and PPML trade cost estimates is 0.89. Note that four of the ten top hubs in Table 2.A.C.2 are confirmed by this exercise. In line with OLS, results look similar (with slightly more violations) when we include additional covariates.

In Table 2.A.E.4, we see that a similar picture arises in other years. The share of violations appears to increase with the number of countries in the sample. This is intuitive: as more potential intermediaries become available, the TI is more likely to be violated (cf. discussion in Section 2.4).

2.A.D Additional Figures

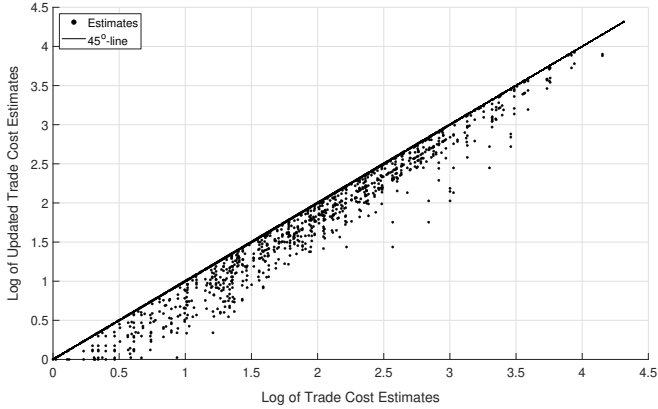


Figure 2.A.D.1: Standard Trade Cost Estimates vs. Updated Trade Cost

Notes. This figure depicts the log of estimated trade costs vs. their updated counterparts. As detailed in Appendix 2.A.C, we estimate trade costs using [Vaugh's \(2010\)](#) method. Whenever there is a violation of the triangle inequality, we obtain updated trade costs following Proposition 1. If an estimate falls below one, we set it equal to one (i.e., $t_{ij} \geq 1$).

Data. See Appendix 2.A.B.

Graph. Authors' representation.

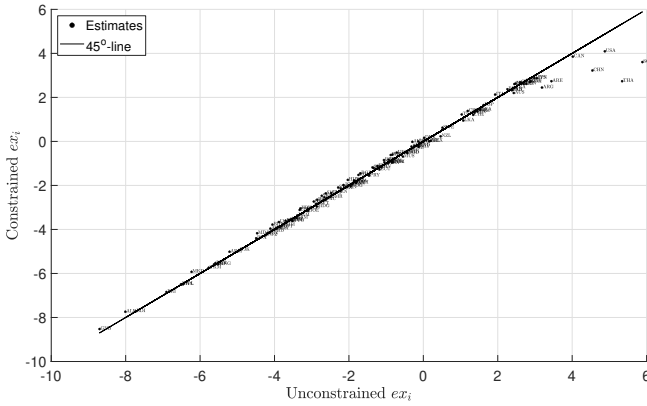


Figure 2.A.D.2: Unconstrained vs. Constrained Values of ex_i

Notes. This figure depicts the unconstrained vs. constrained estimates of ex_i . As detailed in Section 2.3, we estimate trade costs via both unconstrained and constrained OLS. The regression is described by equation (2.4). The constrained version imposes inequalities (2.5) and (2.6) to ensure that estimates satisfy the TI. ex_i is the sum of country i 's exporter and importer fixed effects.

Data. See Appendix 2.A.B.

Graph. Authors' representation.

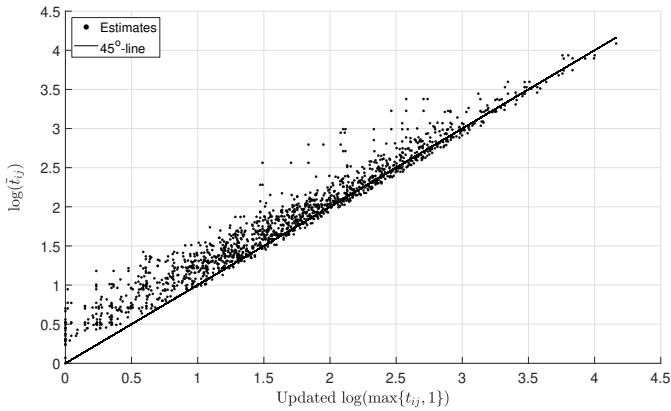


Figure 2.A.D.3: Updated Trade Cost Estimates vs. Constrained Trade Cost

Notes. This figure depicts the log of “updated” unconstrained trade costs vs. their constrained counterparts. As detailed in Section 2.3, we estimate trade costs via both unconstrained and constrained OLS. The regression is depicted by equation (2.4). The constrained version imposes inequalities (2.5) and (2.6) to ensure that estimates satisfy the TI. t_{ij} ($\tilde{t}_{ij}$) are the unconstrained (constrained) estimates of τ_{ij} . If an unconstrained estimate falls below one, we set it equal to one (i.e., $t_{ij} \geq 1$). Whenever there is a violation of the triangle inequality, we obtain updated trade costs following Proposition 1.

Data. See Appendix 2.A.B.

Graph. Authors’ representation.

2.A.E Additional Tables

Table 2.A.E.1: TI Violations with Standard Trade Cost Estimates (Additional Covariates)

	Baseline	Restrict $t \geq 1$
Panel A: Full sample		
% of trade pairs with arbitrage opportunity	41%	27%
... trade-weighted	62%	40%
Top 10 hubs (desc. order): DEU SGP THA ARG BEL ARE SAU USA GBR MYS		
Panel B: Excl. top 3 hubs		
% of trade pairs with arbitrage opportunity	28%	18%
Panel C: Excl. top 10 hubs		
% of trade pairs with arbitrage opportunity	15%	10%

Notes. This table reproduces Table 2.A.C.2 adding four variables to the regressions that indicate whether a country pair (i) shares a common official language, (ii) a common currency, (iii) ever had a colonial link, or (iv) has a regional trade agreement. Details are given in the notes to Table 2.A.C.2.

Table 2.A.E.2: TI Violations with Standard Trade Cost Estimates (Error Term in τ)

	Baseline	Restrict $t \geq 1$
Panel A: Full sample		
% of trade pairs with arbitrage opportunity	50%	48%
... trade-weighted	11%	9%
Top 10 hubs (desc. order):		
SGP ARE SAU THA RUS ZAF PAN ARG VEN USA		
Panel B: Excl. top 3 hubs		
% of trade pairs with arbitrage opportunity	47%	46%
Panel C: Excl. top 10 hubs		
% of trade pairs with arbitrage opportunity	39%	39%

Notes. This table reproduces Table 2.A.C.2 but adding the error term in the regression equation (2.4) to the trade cost function (equation (2.3)). The table includes only trade pairs with positive trade flows, while the remaining trade relations are omitted. Details are given in the notes to Table 2.A.C.2.

Table 2.A.E.3: TI Violations with Standard Trade Cost Estimates (PPML)

	Baseline	Restrict $t \geq 1$
Panel A: Full sample		
% of trade pairs with arbitrage opportunity	26%	16%
... trade-weighted	7%	5%
Top 10 hubs (desc. order): SGP ZAF SAU PAN THA ARE RUS KAZ KEN VEN		
Panel B: Excl. top 3 hubs		
% of trade pairs with arbitrage opportunity	10%	6%
Panel C: Excl. top 10 hubs		
% of trade pairs with arbitrage opportunity	2%	1%

Notes. This table reproduces Table 2.A.C.2 using PPML instead of OLS. Details are given in the notes to Table 2.A.C.2.

Table 2.A.E.4: TI Violations with Standard Trade Cost Estimates (By Year)

Year	% of trade pairs with arbitrage opportunity		Nr. of countries
	Baseline	Restrict $t \geq 1$	
2000	22%	13%	122
2001	28%	17%	124
2002	26%	14%	125
2003	27%	16%	126
2004	36%	20%	127
2005	37%	21%	127
2006	39%	22%	129
2007	37%	22%	131
2008	43%	27%	132
2009	43%	26%	133
2010	38%	23%	135
2011	47%	28%	136
2012	57%	36%	137
2013	61%	39%	137
2014	59%	37%	133

Notes. This table reproduces the top row in Table 2.A.C.2 for several years. Details are given in the notes to Table 2.A.C.2.

Table 2.A.E.5: Welfare Changes after 25% Drop in Global Trade Costs (Selection)

Country	(1) With re-routing	(2) No re-routing	(3) Autarky loss	(4) ((1)-(2))/ (3)
ARE	57%	49%	-45%	18%
AUT	23%	16%	-9%	68%
BEL	24%	20%	-13%	32%
BIH	24%	21%	-34%	6%
CAN	37%	35%	-22%	10%
CHE	21%	17%	-10%	35%
CHL	41%	33%	-18%	44%
CIV	45%	41%	-26%	14%
CMR	33%	30%	-15%	14%
CZE	22%	15%	-10%	72%
DNK	22%	15%	-8%	80%
ESP	15%	7%	-4%	208%
FRA	10%	8%	-5%	51%
GBR	10%	8%	-6%	34%
GHA	19%	14%	-15%	32%
GIN	21%	19%	-10%	22%
GMB	7%	4%	-29%	13%
IDN	31%	25%	-11%	56%
IRL	32%	27%	-12%	45%
ISR	18%	13%	-7%	59%
ITA	9%	7%	-4%	49%
JOR	17%	13%	-21%	19%
KAZ	49%	42%	-25%	29%
KHM	45%	40%	-21%	23%
LBN	10%	7%	-45%	5%
LTU	21%	16%	-12%	43%
LVA	18%	14%	-10%	36%
MNG	25%	20%	-15%	37%
MYS	51%	35%	-19%	84%
NGA	11%	8%	-3%	120%
NLD	21%	16%	-10%	51%
NOR	16%	14%	-6%	46%
OMN	52%	43%	-25%	36%
PAK	23%	20%	-18%	17%
POL	14%	9%	-6%	84%
PRY	42%	40%	-27%	6%
SWE	18%	14%	-7%	51%
TGO	26%	23%	-16%	17%
TUR	21%	13%	-9%	88%
YEM	33%	26%	-15%	49%
Average	26%	21%	-16%	45%

Notes. This table reports welfare changes after a decrease in global bilateral trade costs of 25%. The table is restricted to the 40 countries with the highest difference between Columns (1) and (2). Column (1) allows for re-routing after the shock. Column (3) reports the real income loss from going to full autarky (i.e., $\tau_{ij} \rightarrow \infty \forall i \neq j$). Column (4) shows the difference between Columns (1) and (2) scaled by the absolute value of Column (3). See Section 2.4 for further details.

Data. See Appendix 2.A.B.

Table 2.A.E.6: Welfare Changes after Drop in Global Trade Costs (Averages)

Drop in trade costs	Distribution of net welfare changes from re-routing					
	Min.	Q1	Q2	Q3	Max.	Mean (scaled by net autarky losses)
5%	0%	0%	0%	0%	0.4%	0.2%
10%	0%	0%	0%	0.1%	1.9%	0.8%
15%	0%	0%	0.1%	0.3%	5.7%	2.6%
25%	0%	0.4%	0.7%	2.4%	15.9%	17.5%
40%	0%	1.8%	9.3%	18.3%	80.3%	108%
50%	0.4%	9.8%	22.5%	51%	249.9%	268.7%

Notes. This table reports the distribution of net welfare changes from re-routing after a decrease in global bilateral trade costs of 5%, 10%, 15%, 25%, 40%, or 50%. Q1, Q2, Q3 refer, respectively, to the first, second and third quartiles. The right-most column reports the average net real income change relative to the net loss from going to full autarky (i.e., $\tau_{ij} \rightarrow \infty \forall i \neq j$). See Section 2.4 for further details.

Data. See Appendix 2.A.B.

Results. Authors' computations.

Table 2.A.E.7: Welfare Changes after Eliminating Asymmetries in Trade Costs (Selection)

Country	(1) With re-routing	(2) No re-routing	(3) Autarky loss	(4) ((1)-(2))/ 3
ARM	154%	144%	-17%	61%
AZE	126%	106%	-14%	145%
BIH	140%	91%	-34%	142%
CRI	97%	88%	-12%	68%
CZE	32%	19%	-10%	128%
DNK	21%	13%	-8%	95%
DZA	124%	115%	-29%	30%
ERI	140%	108%	-15%	207%
EST	63%	54%	-12%	71%
ETH	72%	63%	-13%	67%
GEO	88%	70%	-25%	71%
GMB	111%	99%	-29%	42%
HND	101%	92%	-41%	23%
HRV	93%	72%	-12%	175%
HUN	50%	42%	-11%	69%
IRQ	187%	162%	-39%	65%
JAM	189%	170%	-24%	78%
JOR	60%	51%	-21%	42%
KAZ	57%	45%	-25%	48%
KHM	236%	221%	-21%	68%
KWT	142%	124%	-10%	178%
LAO	198%	176%	-17%	127%
LTU	82%	62%	-12%	171%
LVA	93%	65%	-10%	267%
MDA	130%	119%	-21%	53%
MKD	117%	101%	-7%	238%
MNG	262%	246%	-15%	104%
MWI	83%	74%	-7%	111%
MYS	20%	12%	-19%	44%
NIC	145%	135%	-36%	29%
OMN	139%	128%	-25%	44%
PRY	82%	71%	-27%	40%
SLV	99%	91%	-44%	18%
SVK	60%	47%	-12%	113%
SVN	60%	46%	-11%	136%
TJK	121%	87%	-10%	331%
TKM	152%	127%	-7%	368%
TTO	194%	186%	-14%	63%
URY	31%	22%	-15%	61%
YEM	169%	136%	-15%	218%
Average	113%	97%	-19%	110%

Notes. This table reports welfare changes after eliminating asymmetries in trade costs (i.e., we set $\tilde{t}_{ij} = \min\{\tilde{t}_{ij}, \tilde{t}_{ji}\}$). The table is restricted to the 40 countries with the highest difference between Columns (1) and (2). Column (1) allows for re-routing after the shock. Column (3) reports the real income loss from going to full autarky (i.e., $\tau_{ij} \rightarrow \infty \forall i \neq j$). Column (4) shows the difference between Columns (1) and (2) scaled by the absolute value of Column (3). See Section 2.4 for further details.

Table 2.A.E.8: Welfare Changes after 25% Drop in Global Trade Costs (Single-Sector – WIOD Sample)

Country	(1) With re-routing	(2) No re-routing	(3) Autarky loss	(4) $((1)-(2))/ (3) $
AUS	6%	6%	-3%	0%
AUT	14%	13%	-7%	19%
BEL	19%	18%	-9%	12%
BRA	4%	4%	-2%	14%
CAN	10%	10%	-5%	8%
CHN	5%	5%	-2%	4%
CZE	15%	14%	-7%	23%
DEU	11%	11%	-5%	-1%
DNK	14%	13%	-7%	18%
ESP	6%	6%	-4%	0%
FIN	11%	11%	-5%	1%
FRA	6%	6%	-4%	0%
GBR	7%	7%	-4%	-1%
GRC	6%	6%	-5%	0%
HUN	17%	17%	-10%	-1%
IDN	7%	7%	-4%	0%
IND	5%	5%	-3%	2%
IRL	20%	20%	-10%	0%
ITA	6%	6%	-3%	0%
JPN	4%	4%	-2%	0%
KOR	10%	10%	-5%	0%
MEX	8%	8%	-4%	0%
NLD	16%	15%	-7%	8%
POL	9%	9%	-5%	1%
PRT	8%	8%	-5%	1%
ROU	7%	7%	-5%	0%
RUS	6%	6%	-3%	0%
ROW	10%	10%	-6%	0%
SVK	16%	16%	-9%	-1%
SVN	14%	14%	-8%	-1%
SWE	12%	12%	-6%	1%
TUR	5%	5%	-4%	1%
TWN	16%	16%	-8%	0%
USA	3%	3%	-2%	0%
Average	10%	10%	-5%	3%

Notes. This table reproduces Table 2.A.E.5 using the WIOD data instead of the ITPD-E data, and thus reducing the sample to 34 countries. Details are given in the notes to Table 2.A.E.5.

Table 2.A.E.9: Welfare Changes after Eliminating Asymmetries in Trade Costs (Single-Sector – WIOD Sample)

Country	(1) With re-routing	(2) No re-routing	(3) Autarky loss	(4) $((1)-(2))/ (3) $
AUS	29%	4%	-3%	915%
AUT	45%	13%	-7%	467%
BEL	34%	7%	-9%	288%
BRA	29%	3%	-2%	1437%
CAN	31%	4%	-5%	609%
CHN	13%	2%	-2%	498%
CZE	53%	14%	-7%	539%
DEU	16%	8%	-5%	154%
DNK	58%	16%	-7%	617%
ESP	22%	3%	-4%	501%
FIN	57%	9%	-5%	897%
FRA	18%	3%	-4%	408%
GBR	19%	3%	-4%	409%
GRC	68%	20%	-5%	953%
HUN	63%	14%	-10%	500%
IDN	36%	6%	-4%	844%
IND	35%	5%	-3%	1036%
IRL	63%	15%	-10%	508%
ITA	20%	4%	-3%	448%
JPN	12%	3%	-2%	457%
KOR	20%	5%	-5%	287%
MEX	28%	6%	-4%	545%
NLD	35%	7%	-7%	372%
POL	37%	8%	-5%	544%
PRT	67%	14%	-5%	998%
ROU	61%	13%	-5%	892%
RUS	30%	6%	-3%	853%
ROW	44%	60%	-6%	-290%
SVK	83%	23%	-9%	655%
SVN	108%	45%	-8%	757%
SWE	44%	9%	-6%	581%
TUR	42%	9%	-4%	951%
TWN	35%	8%	-8%	347%
USA	6%	2%	-2%	194%
Average	40%	11%	-5%	593%

Notes. This table reproduces Table 2.A.E.7 using the WIOD data instead of the ITPD-E data, and thus reducing the sample to 34 countries. Details are given in the notes to Table 2.A.E.7.

Table 2.A.E.10: Welfare Changes after Eliminating Asymmetries in Trade Costs (Multi-Sector)

Country	(1) With re-routing	(2) No re-routing	(3) Autarky loss	(4) ((1)-(2))/ 3
AUS	106%	21%	-16%	529%
AUT	126%	51%	-53%	143%
BEL	143%	51%	-56%	164%
BRA	143%	22%	-7%	1738%
CAN	137%	25%	-32%	354%
CHN	127%	18%	-8%	1387%
CZE	237%	47%	-36%	525%
DEU	109%	46%	-22%	282%
DNK	150%	40%	-41%	269%
ESP	88%	20%	-19%	347%
FIN	157%	34%	-21%	577%
FRA	74%	21%	-18%	290%
GBR	87%	38%	-23%	217%
GRC	178%	38%	-24%	583%
HUN	280%	59%	-56%	397%
IDN	195%	47%	-13%	1114%
IND	128%	29%	-10%	994%
IRL	199%	76%	-38%	320%
ITA	87%	20%	-17%	386%
JPN	56%	13%	-5%	920%
KOR	127%	26%	-14%	730%
MEX	176%	43%	-19%	701%
NLD	120%	49%	-43%	165%
POL	163%	42%	-35%	349%
PRT	185%	45%	-39%	363%
ROU	313%	70%	-27%	882%
RUS	297%	58%	-38%	624%
ROW	374%	353%	-30%	69%
SVK	327%	83%	-53%	456%
SVN	367%	133%	-60%	388%
SWE	134%	34%	-26%	379%
TUR	175%	56%	-28%	432%
TWN	120%	29%	-18%	500%
USA	56%	27%	-8%	352%
Average	169%	52%	-28%	527%

Notes. This table reports welfare changes after eliminating asymmetries in sector-level trade costs (i.e., we set $\tilde{t}_{ij}^k = \min\{\tilde{t}_{ij}^k, \tilde{t}_{ji}^k\}$), calculated using the multi-sector model. Column (1) allows for re-routing after the shock. Column (3) reports the real income loss from going to full autarky (i.e., $\tau_{ij}^k \rightarrow \infty \forall i \neq j$, and k). Column (4) shows the difference between Columns (1) and (2) scaled by the absolute value of Column (3). See Section 2.4 for further details.

Data. See Appendix 2.A.B.

Results. Authors' computations.

Table 2.A.E.11: Welfare Changes after 25% Drop in Global Trade Costs (Single-Sector – WIOD Sample excl. ROW)

Country	(1) With re-routing	(2) No re-routing	(3) Autarky loss	(4) $((1)-(2))/ (3) $
AUS	5%	5%	-2%	0%
AUT	13%	11%	-6%	23%
BEL	17%	16%	-9%	13%
BRA	2%	2%	-1%	0%
CAN	9%	9%	-4%	0%
CHN	4%	4%	-1%	0%
CZE	14%	13%	-7%	26%
DEU	10%	10%	-5%	-1%
DNK	11%	10%	-6%	21%
ESP	5%	5%	-3%	0%
FIN	9%	9%	-5%	0%
FRA	5%	5%	-3%	0%
GBR	6%	6%	-3%	-1%
GRC	3%	3%	-4%	0%
HUN	15%	15%	-9%	-1%
IDN	5%	5%	-2%	0%
IND	3%	3%	-1%	0%
IRL	18%	18%	-9%	0%
ITA	5%	5%	-3%	-1%
JPN	3%	3%	-1%	0%
KOR	8%	8%	-4%	0%
MEX	7%	7%	-4%	0%
NLD	15%	14%	-6%	10%
POL	8%	8%	-5%	0%
PRT	7%	7%	-5%	0%
ROU	6%	6%	-5%	-1%
RUS	5%	5%	-2%	0%
SVK	15%	15%	-8%	-2%
SVN	12%	12%	-7%	-1%
SWE	9%	10%	-5%	-1%
TUR	4%	4%	-2%	0%
TWN	14%	14%	-5%	0%
USA	2%	2%	-2%	0%
Average	8%	8%	-4%	2%

Notes. This table reproduces Table 2.A.E.8 excluding “rest-of-the-world” (ROW) from the sample. Details are given in the notes to Table 2.A.E.8.

Table 2.A.E.12: Welfare Changes after Eliminating Asymmetries in Trade Costs (Single-Sector – WIOD Sample excl. ROW)

Country	(1) With re-routing	(2) No re-routing	(3) Autarky loss	(4) $((1)-(2))/ (3) $
AUS	1%	1%	-2%	0%
AUT	9%	7%	-6%	38%
BEL	4%	3%	-9%	11%
BRA	1%	1%	-1%	0%
CAN	3%	3%	-4%	0%
CHN	2%	2%	-1%	0%
CZE	11%	8%	-7%	46%
DEU	8%	8%	-5%	-3%
DNK	8%	6%	-6%	47%
ESP	1%	1%	-3%	0%
FIN	4%	4%	-5%	2%
FRA	1%	1%	-3%	2%
GBR	2%	2%	-3%	-1%
GRC	5%	4%	-4%	5%
HUN	7%	7%	-9%	0%
IDN	2%	2%	-2%	0%
IND	1%	1%	-1%	0%
IRL	8%	8%	-9%	3%
ITA	1%	1%	-3%	0%
JPN	1%	1%	-1%	0%
KOR	2%	2%	-4%	0%
MEX	4%	4%	-4%	0%
NLD	4%	4%	-6%	10%
POL	4%	4%	-5%	8%
PRT	7%	7%	-5%	6%
ROU	5%	5%	-5%	6%
RUS	2%	2%	-2%	0%
SVK	14%	14%	-8%	1%
SVN	22%	21%	-7%	4%
SWE	3%	3%	-5%	1%
TUR	2%	2%	-2%	8%
TWN	4%	4%	-5%	0%
USA	1%	1%	-2%	0%
Average	5%	4%	-4%	6%

Notes. This table reproduces Table 2.A.E.9 excluding “rest-of-the-world” (ROW) from the sample. Details are given in the notes to Table 2.A.E.9.

Table 2.A.E.13: Welfare Changes after 25% Drop in Global Trade Costs (Multi-Sector – excl. ROW)

Country	(1) With re-routing	(2) No re-routing	(3) Autarky loss	(4) $((1)-(2))/ (3) $
AUS	15%	15%	-13%	1%
AUT	27%	22%	-52%	10%
BEL	32%	29%	-55%	6%
BRA	11%	10%	-5%	15%
CAN	17%	14%	-31%	9%
CHN	12%	11%	-7%	27%
CZE	36%	30%	-35%	18%
DEU	16%	15%	-21%	7%
DNK	24%	18%	-40%	15%
ESP	16%	12%	-18%	21%
FIN	25%	21%	-20%	19%
FRA	13%	10%	-17%	16%
GBR	17%	13%	-22%	17%
GRC	13%	12%	-23%	1%
HUN	39%	32%	-55%	13%
IDN	22%	16%	-9%	73%
IND	13%	10%	-8%	39%
IRL	56%	40%	-37%	44%
ITA	14%	11%	-16%	18%
JPN	7%	5%	-3%	59%
KOR	21%	17%	-12%	34%
MEX	17%	14%	-18%	14%
NLD	26%	22%	-42%	8%
POL	24%	19%	-34%	14%
PRT	23%	21%	-38%	4%
ROU	21%	17%	-26%	13%
RUS	18%	16%	-36%	5%
SVK	40%	35%	-51%	11%
SVN	31%	26%	-58%	9%
SWE	24%	19%	-24%	18%
TUR	13%	10%	-26%	11%
TWN	26%	28%	-16%	-8%
USA	6%	5%	-8%	13%
Average	22%	18%	-26%	17%

Notes. This table reproduces Table 2.4 excluding “rest-of-the-world” (ROW) from the sample. Details are given in the notes to Table 2.4.

Table 2.A.E.14: Welfare Changes after Eliminating Asymmetries in Trade Costs (Multi-Sector – excl. ROW)

Country	(1) With re-routing	(2) No re-routing	(3) Autarky loss	(4) $((1)-(2))/ 3 $
AUS	12%	8%	-13%	28%
AUT	50%	37%	-52%	24%
BEL	52%	47%	-55%	9%
BRA	14%	7%	-5%	131%
CAN	34%	20%	-31%	45%
CHN	11%	8%	-7%	44%
CZE	72%	29%	-35%	123%
DEU	51%	44%	-21%	30%
DNK	44%	26%	-40%	44%
ESP	19%	11%	-18%	42%
FIN	32%	15%	-20%	82%
FRA	21%	15%	-17%	38%
GBR	35%	30%	-22%	23%
GRC	37%	16%	-23%	94%
HUN	70%	32%	-55%	69%
IDN	21%	18%	-9%	39%
IND	16%	8%	-8%	107%
IRL	95%	70%	-37%	67%
ITA	21%	13%	-16%	50%
JPN	7%	5%	-3%	54%
KOR	16%	13%	-12%	28%
MEX	49%	33%	-18%	89%
NLD	67%	57%	-42%	25%
POL	40%	23%	-34%	50%
PRT	48%	25%	-38%	61%
ROU	65%	30%	-26%	134%
RUS	39%	17%	-36%	61%
SVK	98%	39%	-51%	116%
SVN	89%	43%	-58%	80%
SWE	31%	17%	-24%	61%
TUR	46%	33%	-26%	48%
TWN	19%	16%	-16%	16%
USA	25%	24%	-8%	19%
Average	41%	25%	-26%	59%

Notes. This table reproduces Table 2.A.E.10 excluding “rest-of-the-world” (ROW) from the sample. Details are given in the notes to Table 2.A.E.10.

Chapter 3

Gravity with History: On the Aggregate Implications of Incumbency Effects in International Trade

Joint with Reto Foellmi and Ulrich Schetter

3.1 Introduction

Countries open up to trade gradually over time. When firms start serving foreign markets, they therefore need to compete against domestic and potentially third-country incumbents. The latter have a first-mover advantage, as they, e.g., already built up a distribution network, got to know market structures, and established their brand. An extensive literature on industrial organization

shows how incumbency effects can deter entry (see, e.g., [Arping and Diaw \(2008\)](#); [Goolsbee and Syverson \(2008\)](#); [Noh and Moschini \(2006\)](#); [Schivardi and Viviano \(2011\)](#)), and it is well-known that sunk market access costs can generate hysteresis effects in international trade (see, e.g. [Alessandria et al. \(2020\)](#); [Baldwin \(1988\)](#); [Baldwin and Krugman \(1989\)](#); [Dixit \(1989\)](#)). It is less well understood what this implies for aggregate trade and “gravity” models of international trade. In this paper, we seek to contribute to closing this gap by introducing incumbency effects into a tractable dynamic trade model with heterogeneous firms.

The key novelty in our framework is that a firm’s market access costs are (weakly) decreasing with its tenure in the market. When confronted with market entry by foreign competitors, incumbent firms therefore weigh discounted future profits against lower market access costs. This leads to less exit by (low-productive) incumbents and, hence, less entry vis-à-vis a version of the model without tenure effects.¹ A major benefit of our model is that it reduces to the canonical (static) [Melitz \(2003\)](#) model when market access costs are constant over time. We derive a gravity equation that explicitly accounts for potential incumbency effects. It implies that, *ceteris paribus*, countries trade less if they liberalized their trade relation later. Accordingly, trade flows are not only shaped by trade frictions today—as implicitly assumed in standard gravity equations—but they also depend on the history of trade liberalizations. In other words, we argue that there is strong path dependence in international trade.² Incumbency effects may thus help to explain the “home bias” in international trade, and, more generally, why historic events such as colonial ties or the “Iron Curtain” have long-lasting consequences for trade flows. The model’s

¹This is in line with e.g. [Schivardi and Viviano \(2011\)](#), who show that entry barriers are empirically associated with lower productivity of incumbent firms.

²The term “path dependence” is sometimes used to refer to different things, and it has become a “*trendy way to say that history matters*” ([Page, 2006](#), p. 87). In economics, it is often used to refer to situations where random shocks can have long-lasting effects. A prominent example are agglomeration effects, where small random differences in the initial allocation can have long-lasting and large consequences. In this paper, we apply a more literal use of the term “path dependence.” In particular, we use it to refer to a situation where today’s outcomes are not only shaped by where we are today (in terms of trade costs), but also by which route we took to get to this point.

implications are also in line with novel indirect evidence on the response of fixed market access costs to a shock on variable trade costs.

Our key assumption of decreasing fixed market access costs and, hence, sunk upfront investments is in line with persistence in firms' export status (Alessandria et al., 2020; Eaton et al., 2007; Roberts and Tybout, 1997; Ruhl and Willis, 2017)). It is also consistent with Krishna et al. (2021), who show that the fixed (documentation) costs of using preferential trade agreements are decreasing with firms' experience of exploiting the agreement. To further motivate our analysis, we consider market exit of Colombian firms in response to a large negative trade shock with Venezuela in Section 3.2. We show that firms are less likely to stop exporting a product in response to the shock the longer their tenure, controlling for firm and product fixed effects and very flexibly for a firm's sales of a product. This suggests that fixed market access costs are indeed decreasing with tenure.

We start our theoretical analysis of tenure-dependent market access costs with a simple example in Section 3.3. In this example, firms are myopic and there are two perfectly symmetric countries with free trade—i.e. no variable trade costs and symmetric market access costs. We use this example to show that incumbency effects can give rise to selection into exporting and a “home bias” of international trade flows even with free trade between perfectly symmetric countries. In the remainder of the paper, we generalize this basic insight to a set-up with forward-looking firms, many asymmetric countries, arbitrary trade frictions, and an arbitrary sequence of trade liberalizations.

In Section 3.4, we present our model. Analogous to the canonical static version of the Melitz model (see, e.g., Melitz and Redding (2014)) there is a fixed cost of entry. Upon entry firms receive a random productivity draw from a Pareto distribution. Serving a market involves a fixed cost of market access, and a variable iceberg trade cost. The key difference to the canonical set-up is that our model is dynamic, with fixed market access costs that decrease with tenure in a market. Firms are forward-looking: in their (market) entry decision, they weigh the present value of future profits against the present value of fixed costs. They do not anticipate any changes to the exogenous

trade costs, but have otherwise perfect foresight.

These dynamics notwithstanding, our model is tractable enough to analytically characterize the equilibrium. We show this in Section 3.5, where we derive a gravity equation for international trade. First, we solve the equilibrium for the case where all trade liberalizations happen simultaneously in the very first period. In this scenario, tenure effects trivially play no role, and the gravity equation is identical to that in the canonical Melitz model. Second, we allow for an arbitrary sequence of trade liberalizations. In this case, our theory predicts that, *ceteris paribus*, countries trade more if they liberalized their trade relationship earlier. Our work reveals a novel, multiplicatively separable “incumbency term” in the nominator of the gravity equation, which captures the impact of “history” on current trade flows.

This gravity equation can be estimated following standard steps from the literature. It implies, however, that in addition to standard proxies for fixed and variable trade costs, we need to control for the history of trade liberalizations. To illustrate the potential relevance of this novel term, we include indicators for major historic events—colonial ties and the “Iron Curtain”—in a standard gravity regression in Section 3.6.1. Colonial ties between country pairs are associated with larger trade flows, *ceteris paribus*. This control variable is typically included in the “gravity” literature. As a new insight, we document that countries which were separated by the Iron Curtain in the past trade significantly less today. The differences compared to non-separated relations are large and persistent. These outcomes are stable across different specifications for bilateral trade costs. We also obtain the same conclusions when considering sea and air trade only, where trade should be less affected by pair-specific infrastructure investments. While an indicator for colonial ties between country pairs is typically included in gravity regressions, our theory provides a mechanism that can help explain these empirical patterns: incumbency effects that may prevent firm entry and exit upon trade liberalizations.

Yet, there may be other channels through which the Iron Curtain and colonial ties impact trade today. In Section 3.6.2, we therefore turn to our underlying mechanism. We derive a structural estimator that maps changes

in the importance of tenure for survival in a market to average firm sales (i.e., total exports divided by the number of firms serving a market). With that, we are able to empirically test the model's predictions using data on average firm exports for a large set of countries. Specifically, our theory predicts that in response to a trade liberalization between an exporter i and a destination j , incumbency effects become less important for firms from i relative to firms from other countries that serve j . Intuitively, if i liberalizes its trade relationship with j , the profit potential of firms from i in j increases, which facilitates survival in the market with or without incumbency benefits. Considering relatively large tariff cuts, we find evidence in support of this theoretical prediction. In turn, this implies that the gain from a decrease in variable trade costs is partly offset by the strength of incumbents from other source countries.

Relation to the Literature. Our paper is related to several strands of literature.

We develop a simple dynamic variant of the canonical (static) version of a Melitz model. This model has been extended in many ways, but mostly focusing on static environments (see, e.g. [Chaney \(2008\)](#), [Helpman et al. \(2008\)](#), [Arkolakis \(2010\)](#), [Eaton et al. \(2011\)](#), [Sampson \(2014\)](#), [Arkolakis et al. \(2018\)](#), and the survey articles by [Redding \(2011\)](#) and [Melitz and Redding \(2014\)](#).) A few exceptions consider dynamic effects of trade liberalizations. [Melitz and Ottaviano \(2008\)](#) distinguish between short- and long-run effects of trade liberalizations in a model with variable mark-ups, where in the former case there is a fixed number of incumbent firms. [Chaney \(2005\)](#) shows that sluggish exit of firms may result in productivity overshooting after a trade liberalization. Similarly, [Burstein and Melitz \(2013\)](#) find slow transitional dynamics in response to a trade liberalization in a model with endogenous innovation by firms. [Costantini and Melitz \(2007\)](#) also consider firm innovation and exporting, documenting interesting interactions between the two. They show that innovation may increase prior to anticipated trade liberalizations. Our focus is different, since we are interested in understanding how the order of trade liberalizations shapes international trade flows today.

Our paper is closest related to the extensive “gravity” literature in international trade. The gravity equation was originally motivated by an analogy with gravity between objects (Tinbergen, 1962), but we now have a thorough understanding of its foundations and implications thanks to seminal contributions by Eaton and Kortum (2002), Anderson and van Wincoop (2003), Melitz (2003), Chaney (2008), Arkolakis et al. (2012), and others—see Head and Mayer (2014) for a review of the literature. While the majority of the literature in this area is static, a series of recent papers introduces dynamics into gravity models of trade and migration (see, e.g., Allen and Donaldson, 2020; Atkin et al., 2021; Buera and Oberfield, 2020; Caliendo et al., 2021). In these papers, history can impact outcomes today through technological diffusion, agglomeration, or the evolution of country capability. These mechanisms have in common that they impact an exporter (or city / region) “fixed effect,” i.e., the impact of history is not pair-specific. As opposed to that, we present a simple dynamic framework and derive a gravity equation where trade flows are not only shaped by trade frictions today, but also by the history of trade liberalizations. In this regard, our paper is closer to Morales et al. (2019), who show that the probability of an exporter entering a new market hinges on how similar this new destination is to destinations which the firm already served in past periods. In our model, countries who liberalized their trade relationship earlier trade more today. Interestingly, with countries opening up to trade gradually over time, our framework thus also provides a novel mechanism that can help explain a home bias in international trade even with free trade and between perfectly symmetric countries.³

Our analysis is centered on fixed costs of market access which are decreasing

³A “border puzzle” was originally documented in McCallum (1995). While the exact size of the border effect is heavily debated, there appears to be a consensus that it tends to be relatively large. In other words, the empirically observed low shares of trade relative to production are hard to rationalize using standard trade theories. Several explanations have been proposed such as multilateral resistance terms (Anderson and van Wincoop, 2003), technical barriers to trade (e.g., Chen, 2004; Manchin and Pinna, 2009), multi-stage production (Yi, 2010), or non-homothetic preferences (Caron et al., 2014). We complement this literature by showing how tenure effects can naturally induce a home bias, which can help explain a home bias even among U.S. states (Wolf, 2000). Importantly, large home shares and low gains from trade are intimately linked in standard trade theory (Arkolakis et al., 2012).

with tenure. It is well-known at least since [Baldwin \(1988\)](#); [Baldwin and Krugman \(1989\)](#); [Dixit \(1989\)](#) that sunk entry costs can give rise to hysteresis in international trade. Motivated by empirical evidence on firm entry and exit in export markets—see also [Section 3.2](#)—, a growing literature introduces sunk costs of market access into theoretical models to analyze exporter dynamics (see, e.g., [Das et al. \(2007\)](#); [Eaton et al. \(2014\)](#); [Fitzgerald et al. \(2016\)](#); [Ruhl and Willis \(2017\)](#), or [Alessandria et al. \(2020\)](#) for a survey article).⁴ This literature largely focuses on partial equilibrium models and firm-level transitional dynamics. Notable exceptions are, e.g., [Alessandria and Choi \(2014\)](#); [Alessandria et al. \(2014\)](#); [Impullitti et al. \(2013\)](#), who consider general equilibrium environments, but can only allow for two symmetric countries. Compared to this literature, our dynamics are much simpler, and after a shock our economy immediately jumps to a new steady state. In turn, this allows introducing incumbency effects into a general equilibrium trade model with numerous asymmetric countries, arbitrary bilateral fixed and variable trade costs, and an arbitrary sequence of trade liberalizations, while maintaining tractability.

3.2 Motivating Evidence

Before turning to the model, we motivate our setup by looking at the exiting behavior of Colombian exporters in response to a large negative trade shock. Our goal is to measure the exit propensity of firms depending on their market-specific tenure levels while controlling for size. To do so, we use transaction-level data in USD at the 10-digit HS level from ADUANAS for the years 1994 to 2013 provided by the Colombian tax authority (DIAN). We aggregate this data to yearly firm-level exports at the 2-digit HS level, and then consider

⁴An alternative mechanism to create incumbency effects would be to introduce increasing sales over time, which have been found to be important for explaining some of the micro-patterns in the data (see, e.g., [Fitzgerald et al., 2016](#); [Ruhl and Willis, 2017](#)). In our model, this would have qualitatively similar implications as tenure-dependent fixed costs for aggregate trade and its dependence on the history of trade liberalizations. We therefore focus on tenure-dependent fixed costs, which are analytically more tractable and the predominant assumption in the literature. In [Section 3.2](#), we provide evidence to corroborate our setup.

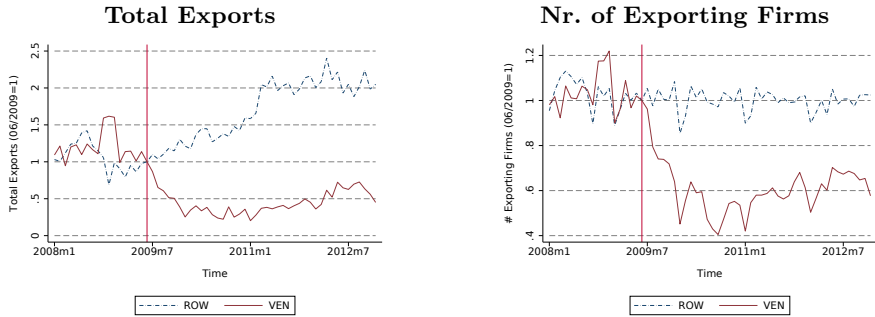
firm exit by product-destination. This or similar datasets have previously been used to document exporter dynamics (see, e.g. [Alessandria et al. \(2020\)](#); [Eaton et al. \(2007\)](#); [Roberts and Tybout \(1997\)](#); [Ruhl and Willis \(2017\)](#)). We add to prior findings (i) by flexibly controlling for log-sales at the product-destination level to emphasize a potential role of tenure-dependent fixed costs, and (ii) by focusing on exit in response to a large negative shock to Colombia's trade with Venezuela rather than on general exporter dynamics.

On July 28, 2009, in response to a dispute over American anti-drug campaigns in Colombia and the supply of weapons to Colombian rebels, the by then president of Venezuela Hugo Chavez declared a freeze on diplomatic ties with Colombia, and expressed the wish to decrease bilateral trade.⁵ While this shock did not entirely close the border, it had a major impact on bilateral trade, as illustrated in Figure 3.1. From July 2009 to December 2009 Colombian exports to Venezuela fell by over 50%, whereas exports to the rest of the world slightly increased. The number of firms that exported to Venezuela fell by over 40%. By contrast, the number of firms that served the rest of the world remained fairly constant. These are sizable shocks, especially considering the pre-crisis importance of Venezuela as a trading partner. In 2008, the last full year before the crisis, more than 15% of Colombian exports went to Venezuela, and about one third of all exporters served that market.

To analyze the exit response of Colombian firms to this (exogenous) shock, we restrict our sample to all firm-product pairs with positive exports to Venezuela in the period July 2008 to June 2009. In our baseline specification, we then treat a firm-product pair as an exit if the firm does no longer export that product to Venezuela in the period from January 2010 to December 2010. This allows comparing pre-crisis levels to post-crisis levels, allowing for the shock to manifest itself gradually over time, as suggested by Figure 3.1. We run the following linear probability model

$$y_{fp} = \lambda'_{fp}\alpha + x'_{fp}\beta + d'_{fp}\gamma + \epsilon_{fp} , \quad (3.1)$$

⁵See <https://www.economist.com/the-americas/2009/09/10/politics-versus-trade>, accessed on March 26, 2021.

Figure 3.1: The Venezuela Shock on Colombian Firm Exports

Notes: These figures illustrate monthly total exports from Colombia to Venezuela vs. the rest of the world (left panel), and the number of firms exporting to Venezuela vs. the rest of the world (right panel). Both graphs depict indexed values such that June 2009 is equal to one. See Section 3.2 for further details.

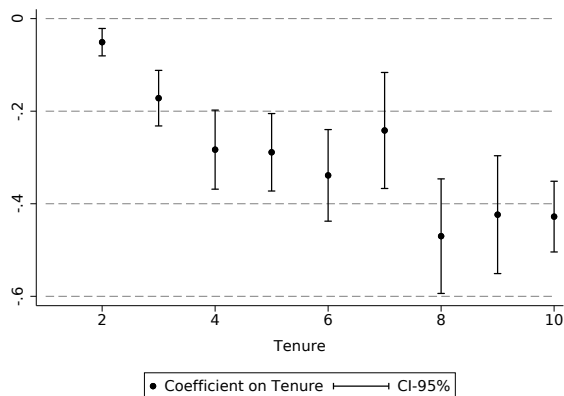
Data source: ADUANAS (DIAN)

Graph: Authors' representation.

where y_{fp} is our exit indicator that takes on a value of one if firm f stops exporting product p in response to the shock, and zero otherwise. λ is a set of tenure dummies $[\lambda^1, \dots, \lambda^{10}]$, where λ_{fp}^k equals one if at the time of the shock firm f exports product p to Venezuela for the k^{th} consecutive period.⁶ $\mathbf{x}$ is a (set of) size controls and $\mathbf{d}$ a set of fixed effects as specified momentarily. The error term is denoted by ϵ_{fp} , and we apply two-way clustering at the product and firm levels.

In our main—most conservative—specification, we order all firms that export a product to Venezuela by their exports, group them into quintiles, and then control for (i) a full set of product-quintile fixed effects, and (ii) the firm's log-sales of that product with a product-quintile-specific coefficient. This allows to flexibly control for a firm's size in a given product. We add a full set of firm fixed effects to control for general firm-level trends. In Table 3.A.F.1 of

⁶To measure tenure in July 2008–June 2009, we aggregate trade flows in preceding years over twelve-month periods from July in a year $t-1$ to June in year t . Table 3.A.F.2 of Appendix 3.A.F shows that the basic pattern is the same when aggregating data from August in a year $t-1$ to July in year t and then looking at exit immediately after the shock (i.e., no exports in the period from August 2009 to July 2010).

Figure 3.2: Firm Exit in Response to Venezuela Shock

Notes: These figures summarize the tenure coefficients from estimating equation (3.1) using OLS. Standard errors are clustered at both, the product and the firm level. All tenure levels of 10 years and above have been summarized in one category “tenure 10.” Omitted category is tenure=1. Further details are provided in the main text.

Data source: ADUANAS (DIAN)

Graph: Authors’ representation.

Appendix 3.A.F, we show that less restrictive specifications with fewer fixed effects yield the same basic insights (cf. Appendix 3.A.B for a discussion).

Figure 3.2 summarizes the coefficients on the tenure dummies, which measure tenure effects relative to the omitted category of firms with tenure 1 (i.e., of firms who had started exporting to Venezuela in the year before the shock). This figure clearly reveals sizable effects of tenure on exit: the exit probability decreases by more than 40 percentage points when a firm is highly experienced (tenure ≥ 8), flexibly controlling for sales prior to the shock.

In what follows we analyze what these tenure effects imply for aggregate trade, and how it is shaped by the history of trade liberalizations. To that end, we embed tenure-dependent fixed market access costs into a Melitz model, which gives rise to size-independent tenure effects in line with our findings here.

3.3 A Simple Example

To illustrate our main point of interest, it is instructive to start off with a stylized example, before introducing our entire set-up in the next section.

Consider a Melitz model with two perfectly symmetric countries and free trade. Hence, for the purpose of this example, let us assume that upon trade opening there are no variable trade costs and market access costs f are the same at home and abroad. In the textbook case of the Melitz model, we then trivially have that the cutoff productivity is the same for every destination market and that

$$\frac{X_{ij}}{X_j} = \frac{1}{2},$$

where X_{ij} denotes trade flows from country i to destination country j , and X_j denotes total expenditure in country j .

Now, suppose that both countries start out under autarky at time $t = 0$, and open up to free trade at some time $t > 0$. Economically, this implies that when firms start to export, they need to compete against incumbent local firms. How is that going to affect international trade? The answer is, of course, not at all if we are willing to assume that fixed costs of market access are constant over time. This is, however, no longer true if these costs are decreasing with tenure in a market. In particular, to keep things simple for now, suppose that firms already serving any given market for t periods have to pay a fraction $\alpha < 1$ of fixed costs only, and that they are myopic in their entry decision. Then, in both countries domestic firms will continue serving their home market whenever

$$\pi_{ii}^v(\varphi) \geq \alpha f,$$

where $\pi_{ij}^v(\varphi)$ denotes the variable profits that a firm in country i with productivity φ makes from serving country j . By contrast, firms from country i will enter market j whenever

$$\pi_{ij}^v(\varphi) \geq f.$$

It follows that the sequential opening to trade in itself gives rise to selection

into exporting and a home bias. In particular, it is straightforward to show that

$$\left(\frac{\varphi_d}{\varphi_x}\right)^{\sigma-1} = \alpha,$$

where φ_d is the cutoff productivity for serving the home market, φ_x the cutoff productivity for serving the foreign market, and $\sigma > 1$ the constant elasticity of substitution.

In what follows, we will develop this argument more carefully and show how it impacts aggregate trade flows between a large set of heterogeneous countries with forward-looking firms, arbitrary variable trade costs, and weakly decreasing with tenure (but otherwise arbitrary) fixed market access cost.

3.4 Model

To analyze how the history of trade liberalizations affects international trade flows today, we now develop a dynamic general equilibrium model of international trade. We build on the canonical static single-sector Melitz model (as in, e.g., [Melitz and Redding \(2014, Section 6\)](#), or [Chaney \(2008\)](#) when there is free entry). Analogous to a standard Melitz model, there is free entry, subject to a fixed cost of entry, and serving a market involves a fixed cost. Forward-looking firms decide which markets to serve. The key novelty in our setup is that fixed cost of serving a market are decreasing with tenure in that market. That is, serving a market involves higher fixed cost upfront (e.g., in order to establish a brand or a distribution network), and lower fixed cost in the future (e.g., in order to sustain the brand or the distribution network).⁷ In the case where these fixed costs are constant over time, our model reduces to the canonical baseline framework.

Households. There are I countries, indexed by $i, j \in \mathcal{J}$. Country i is pop-

⁷The decreasing over-time fixed costs resemble models in the exporter dynamics literature that feature a separate upfront investment and per-period market access costs (see [Alessandria et al. \(2020\)](#) for a review of this literature). As opposed to this literature, however, we consider a general equilibrium framework with a large set of heterogeneous countries, and focus on long-run outcomes rather than transitional dynamics.

ulated by a continuum of measure L_i of infinitely lived households. In every period, each household is endowed with one unit of labor that it inelastically supplies to the labor market. Households receive utility from consuming a continuum of horizontally differentiated varieties with preferences given by

$$U_i = \sum_{t=0}^{\infty} (\beta)^t C_i^t \quad (3.2)$$

$$C_i^t := \left[\int_{\omega \in \Omega_i^t} q_i^t(\omega)^{(\sigma-1)/\sigma} d\omega \right]^{\sigma/(\sigma-1)}, \quad \sigma > 1,$$

where a superscript ^{t} indicates period t , $\beta < 1$ is the discount factor, Ω_i^t denotes the set of varieties available in country i at time t , and $q_i^t(\omega)$ denotes the quantity of variety ω consumed by a representative household in country i at time t . Households can invest in an aggregate asset of domestic firms, as we will detail later on. For now, it suffices to note that this asset yields a rate of return which we denote by r_i^t . The representative household in country i then maximizes (3.2) subject to the flow budget constraint

$$\mathbb{E}_t \left[\frac{a_i^{t+1}}{1 + r_i^{t+1}} \right] = a_i^t + w_i^t - x_i^t, \quad (3.3)$$

where a_i^t denotes (pre-determined) asset holdings at the beginning of the period, w_i^t the wage rate, and x_i^t total spending of the representative household in period t . $\mathbb{E}_t[\cdot]$ denotes the expectation operator.

Following standard steps to solve the household's intra-temporal optimization problem yields its demand for variety ω in country i at time t

$$q_i^t(\omega) p_i^t(\omega) = p_i^t(\omega)^{1-\sigma} (P_i^t)^{\sigma-1} x_i^t,$$

where $p_i^t(\omega)$ denotes the price of variety ω and where

$$P_i^t := \left[\int_{\omega \in \Omega_i^t} p_i^t(\omega)^{1-\sigma} d\omega \right]^{\frac{1}{1-\sigma}}$$

is the ideal price index. Intertemporal optimization further requires that in equilibrium we must have

$$1 = \beta \mathbb{E}_t \left[(1 + r_i^{t+1}) \frac{P_i^t}{P_i^{t+1}} \right] \quad (3.4)$$

at all times, i.e., the interest rate adjusts such that households are exactly indifferent between consuming the CES aggregator either today or in the next period.

Firms and production. The competitive environment and production are analogous to the baseline Melitz model, with the only changes in assumptions that fixed costs of serving a market vary with tenure in that market, and that firms anticipate these changes when deciding on whether or not to enter a market. In particular, there is free entry into the market for final consumption goods. Upon entry, a firm is equipped with a new variety ω , and it receives a productivity draw φ from a random distribution with CDF $G(\varphi)$, PDF $g(\varphi)$, and positive and unbounded support Φ . To derive a gravity equation, we will later consider the case of a Pareto distribution of productivities, i.e.,

$$G(\varphi) = 1 - \left(\frac{\varphi}{\underline{\varphi}} \right)^{-\theta}, \quad \text{for } \varphi \geq \underline{\varphi}$$

$$\text{and } g(\varphi) = \theta \underline{\varphi}^\theta \varphi^{-\theta-1}$$

where $\underline{\varphi} > 0$ is the lowest productivity level, and the shape parameter θ is assumed to satisfy $\theta > \sigma - 1$.⁸

Entry in country i entails a fixed cost. For simplicity, we assume that this fixed cost is paid in form of a perpetual annuity of f_{ei} units of domestic labor.

A firm with productivity φ needs $1/\varphi$ units of domestic labor in order to produce one unit of its variety. In addition, firms face fixed costs of serving a market: A firm based in country i needs to employ f_{ij} units of domestic labor in every period it is serving market j . These labor units need to be

⁸We assume that $\underline{\varphi}$ is the same across countries. This is for expositional convenience only.

equipped with tenure-dependent units of the CES aggregator.⁹ In Section 3.5, we will show that—in the absence of trade shocks—the CES price index is constant over time. In anticipation of this result, we simplify the exposition by assuming that the resources needed for market access simply scale the market access costs. In particular, suppose that a firm from country i has *continuously* served market j for $\lambda \geq 0$ periods, where we use λ to denote tenure in a market. Then its fixed costs of serving that market in period t are

$$f_{ij}^t(\lambda) := f_{ij} w_i^t f(\lambda), \quad (3.5)$$

where $f(\lambda)$ is a tenure-dependent scaling factor satisfying $f(\lambda) \geq 1$ and $f'(\cdot) \leq 0$. Out of these fixed costs, $f_{ij} w_i^t$ are spent on labor in country i , while $[f(\lambda) - 1] f_{ij} w_i^t$ are spent on the CES aggregator.¹⁰ Note that equation (3.5) nests the common assumption of constant-over-time market access costs in terms of labor as a special case (with $f(\lambda) = 1 \forall \lambda$).

Trade is further subject to an iceberg trade cost, that is, $\tau_{ij}^t \geq 1$ units of a variety have to be shipped from country i for one unit to arrive at destination country j . Variable trade costs satisfy the triangle inequality, and $\tau_{ii}^t = 1$ for all i and t . We note that the labor requirements for accessing a market, f_{ij} , and for entering, f_{ei} , are constant over time, while the variable trade costs τ_{ij}^t may potentially change, with any such change being unexpected. Hence, to study the effects of sequential trade openings on trade flows today, we focus

⁹We introduce this hybrid market access cost for two reasons. First, economically, it is reasonable to assume that accessing a market requires both personnel and financial resources for, e.g., a marketing campaign, and that the marketing budget needed to maintain a brand is smaller than the budget needed to establish a brand. Second, this choice largely improves the tractability of our model. For one thing, with parts of the fixed market access costs in terms of labor, we can leverage the well-known result that—with a Pareto distribution of firm productivities—the mass of firms is independent of the trade environment. For another, expressing the time-varying part of fixed cost in terms of the consumption aggregator implies that—in the absence of trade shocks—all prices and aggregate trade flows are constant over time. In turn, this allows us to study long-run implications of incumbency effects for aggregate trade while otherwise staying as close as possible to static microfoundations of the gravity equation.

¹⁰Strictly speaking, in order to interpret the scaling factor directly as a requirement in terms of the CES aggregator, $f(\cdot)$ should be exporter-time dependent, as $f(\lambda)$ should be i - t -dependent to capture fully flexible costs in terms of CES aggregator. However, our main interest is on how they change with tenure, and we therefore ignore this here.

on shocks to variable trade costs. This will allow disentangling tenure effects from the direct effects of changes in the trade environment.

In summary, a firm in country i with productivity φ that started serving households in country j λ periods ago, and that wants to sell q units of its variety, faces a total cost of

$$\mathbb{C}_{ij}^t(q; \varphi, \lambda) = f_{ij}^t(\lambda) + \frac{\tau_{ij}^t w_i^t}{\varphi} q.$$

Firm behavior. Firms in country i can borrow or lend at the prevailing market interest rate r_i^t as given by equation (3.4). Firms are forward-looking. In their (market) entry decision, they weigh discounted future costs against benefits, perfectly anticipating the tenure-dependent evolution of market access cost and any potential endogenous evolution of equilibrium outcomes. They do not, however, anticipate any potential changes in the exogenous trade environment. That is, firms expect all τ_{ij} and f_{ij} to remain constant over time, analogous to static gravity equations. We will later on examine how the sequence of past shocks to these exogenous trade costs impact trade flows today.

Given the forward-looking behavior of firms, it will come in handy to introduce the following notation

$$\begin{aligned} \tilde{f}_{ei}^t &:= \left(f_{ei} \sum_{\varsigma=t}^{\infty} \mathbb{E}_t [w_i^{\varsigma} R_i^{t,\varsigma}] \right) \left(\sum_{\varsigma=t}^{\infty} \mathbb{E}_t [R_i^{t,\varsigma}] \right)^{-1} \\ \tilde{f}_{ij}^t(\lambda) &:= \left(f_{ij} \sum_{\varsigma=t}^{\infty} f(\lambda + (\varsigma - t)) \mathbb{E}_t [w_i^{\varsigma} R_i^{t,\varsigma}] \right) \left(\sum_{\varsigma=t}^{\infty} \mathbb{E}_t [R_i^{t,\varsigma}] \right)^{-1}, \end{aligned}$$

where

$$R_i^{t,\varsigma} := \prod_{s=t}^{\varsigma} \frac{1 + r_i^t}{1 + r_i^s}$$

is the discount factor from period $\varsigma \geq t$ to the beginning of period t . In words, $\tilde{f}_{ei}^t$ is an annuity with time t present value equivalent to that of the upcoming stream of fixed entry costs payable by a firm that entered at time t or before.

Similarly, $\tilde{f}_{ij}^t(\lambda)$ is the annuity for the stream of market access costs payable by a firm that accessed market j for the first time in period $t - \lambda$, and continuously served that market ever since.

With CES preferences, firms charge the well-known constant mark-up over marginal cost. A firm from country i selling to market j offers its product at a price

$$p_{ij}^t(\varphi) = \frac{\tau_{ij}^t w_i^t}{\varphi} \frac{\sigma}{\sigma - 1}, \quad (3.6)$$

and it earns variable profits equal to

$$\pi_{ij}^{v,t}(\varphi) = \frac{1}{\sigma} \left(\frac{\sigma}{\sigma - 1} \right)^{1-\sigma} (\tau_{ij}^t w_i^t)^{1-\sigma} (P_j^t)^{\sigma-1} X_j^t \varphi^{\sigma-1}, \quad (3.7)$$

where X_j^t denotes aggregate expenditure on the CES aggregator in country j as further characterized in Lemma 1 below. Analogous to the above, let $\tilde{\pi}_{ij}^{v,t}(\varphi)$ denote an annuity that has the same time t present value as the infinite stream of variable profits associated with serving market j from period t onward.¹¹ Then, a firm in country i with productivity φ and tenure λ in market j will find it profitable to continue (or start in case of $\lambda = 0$) serving j whenever

$$\tilde{\pi}_{ij}^{v,t}(\varphi) \geq \tilde{f}_{ij}^t(\lambda). \quad (3.8)$$

For a given λ , this allows defining a cutoff productivity similar to a canonical Melitz model, but with a subtle difference that will turn out to be important for our understanding of tenure effects in international trade. In particular, in any period t , $\pi_{ij}^{v,t}(\varphi)$ is increasing in φ , and hence so is $\tilde{\pi}_{ij}^{v,t}(\varphi)$. By contrast, for a given λ , the right-hand-side of condition (3.8) is independent of the firm. Accordingly, we can define a cutoff productivity $\underline{\varphi}_{ij}^{\lambda,t}$ for firms in country i that started serving market j λ periods ago. This cutoff is implicitly defined by

$$\tilde{\pi}_{ij}^{v,t}(\underline{\varphi}_{ij}^{\lambda,t}) = \tilde{f}_{ij}^t(\lambda), \quad (3.9)$$

¹¹In principle, it is possible for a firm to anticipate its own market exit in the future. However, as we will show in Section 3.5, this will not be the case in equilibrium. Firms either exit at time $t = 0$ or they expect to be operating forever.

i.e., firms with $\varphi \geq \underline{\varphi}_{ij}^{\lambda,t}$ find it optimal to serve market j . In case $\lambda = 0$, they start serving the market, and equation (3.9) is analogous to the counterpart of the Melitz model. In case $\lambda > 0$, however, it is typically the case that the least productive firm with tenure λ has productivity $\varphi > \underline{\varphi}_{ij}^{\lambda,t}$. As a consequence, condition (3.8) is strictly non-binding for all such firms. This is because the right-hand side of equation (3.9) is (weakly) decreasing in λ . Therefore, all else equal, tenure in a market introduces a buffer between future profits and future market access costs. This buffer implies that, all else equal, if a firm in country i finds it optimal to serve market j today, it will also find it optimal to keep serving this destination in future periods. As a direct consequence, there is less firm exit in response to a negative trade shock, which will give rise to tenure effects in aggregate trade.

As long as firms can make positive profits from entering, they will do so. Free entry then requires that in every period the expected profits from entering in country i must be non-positive, and they must be exactly equal to zero in every period with strictly positive entry. Using the fact that $\left(\tilde{\pi}_{ij}^{v,t}(\varphi)/\tilde{\pi}_{ij}^{v,t}(\underline{\varphi}_{ij}^{0,t})\right) = \left(\varphi/\underline{\varphi}_{ij}^{0,t}\right)^{\sigma-1}$ and equation (3.9), we obtain the following free entry condition

$$\sum_{j \in \mathcal{J}} \int_{\underline{\varphi}_{ij}^{0,t}}^{\infty} \left[\left(\frac{\varphi}{\underline{\varphi}_{ij}^{0,t}} \right)^{\sigma-1} - 1 \right] \tilde{f}_{ij}^t(0) g(\varphi) d\varphi \begin{cases} \leq \tilde{f}_{ei}^t & \text{if } M_{ei}^t = 0 \\ = \tilde{f}_{ei}^t & \text{if } M_{ei}^t > 0 \end{cases}, \quad (3.10)$$

where M_{ei}^t denotes the mass of entrants in period t . Note that with a positive mass of incumbent firms we can have $M_{ei}^t = 0$ in equilibrium. In fact, as we will detail below, this is the case in equilibrium for every $t > 0$.

Sequence of events. Before analyzing the equilibrium in Section 3.5, it is instructive to clarify the sequence of events in our economy. In any period t , the government first announces changes to the (exogenous) trade environment, if there are any. As all such changes are unanticipated, this will result in an instantaneous updating of all asset values in the economy and of future expectations. We use $\mathbb{E}_t$ to denote time- t expectations *after* this updating

has occurred. Analogously, a_i^t , r_i^t (and firm values detailed in the appendix) capture updated values *after* announcement. Finally, firms simultaneously decide on their production and exporting, households on their consumption and savings, prices are determined, and all markets clear.

3.5 Equilibrium

In this section, we derive the equilibrium in our economy. We start with some preliminary considerations on aggregate demand, adjustments to trade cost shocks, and firm entry. We then derive gravity equations, first for the case of simultaneous trade liberalizations, and then for the general case of sequential liberalizations.

3.5.1 Preliminary Considerations

In our model, aggregate demand on the CES aggregator comprises household expenditure on consumption plus a share of firm expenditure on fixed market access costs, which are both time varying. Yet, in the absence of future entry or changes in firms' market access strategies, aggregate demand in country i is a constant-over-time multiple of its wage rate. In particular, firms in our economy are identified by a triple (i, s, φ) . Accordingly, let $\mathcal{J}^t(i, s, \varphi)$ denote the set of destinations that the firm with productivity φ that was born in period s in country i serves in period $t \geq s$. With this notation, we can show the following:

Lemma 1 *Aggregate—firm plus household—expenditure in country i on the CES aggregator is given by*

$$X_i^t = \frac{\sigma}{\sigma - 1} w_i^t \left[L_i - \sum_{s=0}^t \left[M_{ei}^s \left[f_{ei} + \int_{\varphi \in \Phi} \sum_{j \in \mathcal{J}^t(i, s, \varphi)} f_{ij} g(\varphi) d\varphi \right] \right] \right]. \quad (3.11)$$

Proof See Appendix [3.A.D.a](#). □

The intuition behind Lemma 1 is as follows. The value of the firms, and with it the aggregate asset, changes over time for two reasons. First, due to any unanticipated shocks to the trade environment which may trigger entry, exit, and in general will impact the net present value of domestic firms.¹² Second, due to the time-varying component of fixed market access cost. Potential shocks to expected future earnings of firms simply represent a valuation gain (or loss) on the aggregate asset, and they are therefore one-for-one reflected in households' asset holdings. As opposed to that, time-varying fixed costs are associated with net investments by the representative household. In equilibrium, however, any variations over time in households' net investments are offset by variations over time in the fixed cost of the firms. As a consequence, they do not have a direct effect on aggregate demand.

Lemma 1 implies that—with a constant set of operating firms—aggregate expenditure in country i is a constant-over-time multiple of the wage rate, analogous to the canonical Melitz model. In our model, this implies that in response to a shock all prices and trade flows adjust in the period of the shock and are constant thereafter.

Lemma 2 *Let there be a shock to trade costs at time t . Then firm sales, prices, wages, and interest rates adjust at time t , and remain constant thereafter.*

Proof See Appendix 3.A.D.b. □

Firms are forward-looking in their entry and market access decisions. Nevertheless, the equilibrium is highly tractable thanks to Lemmata 1 and 2. These lemmata imply that entry and market-access costs are exogenously given multiples of the wage rate

$$\begin{aligned}\tilde{f}_{ei}^t &= f_{ei} w_i^t \\ \tilde{f}_{ij}^t(0) &= f_{ij} w_i^t \tilde{f}(0),\end{aligned}$$

¹²Note that, due to free entry, firms always have zero value on average in the entry period, i.e., entry will not affect the value of the aggregate asset in the same period.

where

$$\tilde{f}(0) := \left[\sum_{\lambda=0}^{\infty} f(\lambda) \left(\frac{1}{1+r} \right)^{\lambda} \right] \frac{r}{1+r}.$$

is an annuity for the tenure-dependent multiplier of market access costs, evaluated at the equilibrium interest rate $r := \frac{1}{\beta} - 1$. In turn, this allows solving for the mass of entering firms.

Lemma 3 *All firms are born at $t = 0$. The mass of entrants in country i is*

$$M_{ei}^0 = \frac{L_i}{f_{ei}} \frac{\tilde{f}(0)(\sigma - 1)}{(\sigma - 1) [(\theta + 1)\tilde{f}(0) - 1] + \theta}, \quad (3.12)$$

Proof See Appendix 3.A.D.c. □

According to Lemma 3, all firms enter at $t = 0$ (the very first period). Intuitively, this is because the mass of entering firms is independent of the trade environment, analogous to the canonical Melitz model with fixed market access cost in terms of labor and a Pareto distribution of firm productivities. In fact, with $f(\lambda) = 1 \forall \lambda$, and therefore $\tilde{f}(0) = 1$, equation (3.12) reduces to the well-known expression for the canonical Melitz model with free entry (Melitz and Redding, 2014). With $\tilde{f}(0) > 1$, there are more entrants, reflecting the fact that with higher costs of market access, fewer firms will find it profitable to start operating.¹³

With these insights at hand, we will now characterize equilibrium trade flows. We begin with considering the case of trade opening to all countries at time $t = 0$, which is the counterpart of the canonical Melitz model and will form our benchmark for the analysis of sequential liberalizations below.

¹³Fixed cost of market access feed back into the mass of entrants in our model because they are partly paid in terms of the CES aggregator. In equilibrium, this implies that labor demand of the average operating firm is smaller when compared to the labor demand involved in entering, which translates into a larger labor-market-clearing mass of entering firms. See Appendix 3.A.D.c for further details.

3.5.2 Gravity Equation

Our theory gives rise to a gravity equation for aggregate trade flows, as we now explain.

Simultaneous trade liberalization. To highlight the importance of incumbency effects for aggregate trade, it is instructive to first consider the special case where all trade costs are fixed at time $t = 0$ and then held constant forever. In particular, from the previous section we know that all firms enter at $t = 0$ (Lemma 3), and that the equilibrium is constant over time (Lemma 2). As a consequence, all market entry is simultaneous and incumbency effects—which are our main focus here—trivially have no impact on trade flows. Indeed, the following proposition shows that in this special case, the gravity equation is identical to the one in the canonical Melitz model.

Proposition 2 *Let $\tau_{ij}^\varsigma = \tau_{ij}$ for all $0 \leq \varsigma \leq t$. Then time- t bilateral trade shares are given by*

$$\frac{X_{ij}^t}{X_j^t} = \frac{\frac{L_i}{f_{ei}} (\tau_{ij})^{-\theta} (w_i^t)^{\frac{\sigma-1-\sigma\theta}{\sigma-1}} (f_{ij})^{\frac{\sigma-1-\theta}{\sigma-1}}}{\sum_{k \in \mathcal{J}} \frac{L_k}{f_{ek}} (\tau_{kj})^{-\theta} (w_k^t)^{\frac{\sigma-1-\sigma\theta}{\sigma-1}} (f_{kj})^{\frac{\sigma-1-\theta}{\sigma-1}}} . \quad (3.13)$$

Proof See Appendix 3.A.D.d. □

According to Lemma 2, in the absence of further shocks, all prices and firms' market access strategies are constant over time and, hence, so are bilateral trade shares. With Proposition 2 as our benchmark, we now turn to characterizing the general case of sequential trade liberalizations.

Gravity with history. In an economy with sequential trade liberalizations, the order of trade openings matters for trade flows today. This is because pioneering firms have sunk investments in market access, and they are therefore less prone to exit a market upon trade liberalizations. In aggregate, this implies that—*ceteris paribus*—countries who started trading earlier trade more today, as we now explain.

Suppose there is a trade shock at time $t \geq 0$. Then, there is a unique productivity cutoff $\underline{\varphi}_{ij}^t$ such that all firms in i with productivity $\varphi \geq \underline{\varphi}_{ij}^t$ serve j at t , while all firms with $\varphi < \underline{\varphi}_{ij}^t$ do not.¹⁴ Incumbency effects may, however, imply that the cutoff firm is *not* indifferent between serving market j or not. To analyze equilibrium trade flows in our economy, it will therefore be convenient to introduce a (hypothetical) tenure level $\hat{\lambda}_{ij}^t$ such that

$$\pi_{ij}^{v,t}(\underline{\varphi}_{ij}^t) = w_i^t f_{ij} \tilde{f}(\hat{\lambda}_{ij}^t). \quad (3.14)$$

In words, $\hat{\lambda}_{ij}^t$ is the tenure level with which the least productive firm that serves j from i would just be indifferent to exiting. In response to a sufficiently large positive shock (e.g., a drop in trade barriers), there is market entry, implying that $\hat{\lambda}_{ij}^t = 0$. Incumbency effects have no direct impact on i 's exports to j in this case, because all firms serving j from i would also have done so without market tenure. By contrast, in response to a sufficiently large negative trade shock, some firms from i stop serving market j . In that case, $\hat{\lambda}_{ij}^t$ is exactly equal to the time- t tenure of the least productive firm that continues to serve j , $\lambda_j^t(i, 0, \underline{\varphi}_{ij}^t)$. In general, we have $\hat{\lambda}_{ij}^t \in [0, \lambda_j^t(i, 0, \underline{\varphi}_{ij}^t)]$, which indicates the minimum tenure level needed to sustain the level of aggregate trade from i to j after the shock. We will henceforth simply refer to $\hat{\lambda}_{ij}^t$ as the *tenure effect* in trade from i to j .

Proposition 3 *For any history of trade liberalizations, suppose that at time t there is a shock to the trade environment. Then bilateral trade shares are*

$$\frac{X_{ij}^t}{X_j^t} = \frac{\frac{L_i}{f_{ei}} (\tau_{ij}^t)^{-\theta} (w_i^t)^{\frac{\sigma-1-\sigma\theta}{\sigma-1}} (f_{ij} \tilde{f}(\hat{\lambda}_{ij}^t))^{\frac{\sigma-1-\theta}{\sigma-1}}}{\sum_{k \in \mathcal{J}} \frac{L_k}{f_{ek}} (\tau_{kj}^t)^{-\theta} (w_k^t)^{\frac{\sigma-1-\sigma\theta}{\sigma-1}} (f_{kj} \tilde{f}(\hat{\lambda}_{kj}^t))^{\frac{\sigma-1-\theta}{\sigma-1}}} \quad \forall i, j, \quad (3.15)$$

where $\hat{\lambda}_{ij}^t$ is as defined in equation (3.14) and measures the tenure effect in trade from i to j . In the absence of further shocks, trade shares remain unchanged in future periods.

¹⁴This is true for an arbitrary sequence of trade liberalizations because at all times firms with a higher productivity earn higher variable profits and, hence, they must have a weakly longer tenure in that market. Recall that there is only one firm cohort (Lemma 3).

Proof See Appendix 3.A.D.e. □

Proposition 3 carries the main message of our paper. It generalizes the basic insight from our simple example in Section 3.3 to a world with many, asymmetric countries and forward-looking firms. Specifically, it shows how aggregate trade flows are not only influenced by trade costs today, as captured by τ_{ij}^t and f_{ij} , but also by the history of trade liberalizations that resulted in today's trade costs. The latter effect is summarized by the terms $\tilde{f}(\hat{\lambda}_{kj}^t)$ in equation (3.15). This term implies that, *ceteris paribus*, countries trade more with each other if they liberalized earlier, as $\tilde{f}(\hat{\lambda}_{ij}^t)$ is smaller for these relations. This is because firms with longer tenure in a market are less prone to exit in response to a negative trade shock.¹⁵

3.6 Empirical Application

In the previous sections, we have shown how tenure effects imply that the history of trade liberalizations can have lasting effects on international trade. In this section, we provide suggestive evidence that is in line with this theoretical prediction. A key benefit of our analysis is that it gives rise to a gravity equation for trade flows. In turn, this makes the tenure effects very transparent, and allows contrasting our results with the large theoretical and empirical literature that is centered on gravity equations.

We begin our empirical analysis with running standard gravity regressions in Section 3.6.1, but where we exploit two historical events: colonial ties and the fall of the Iron Curtain. Across a series of robustness checks, we find lasting effects of these historic events on trade flows today. While this is in line with our theory, there are channels other than tenure effects through which these events may impact trade today. In Section 3.6.2, we therefore corroborate our findings by inspecting the underlying mechanisms. Specifically, we derive a structural expression for the response of average firm sales to trade liberalizations, and how they depend on tenure. The data for both exercises

¹⁵Note that if $\hat{\lambda}_{ij}^t$ is the same across all exporters i , equation (3.15) reduces to the standard gravity equation (3.13).

are described in Appendix 3.A.A.

3.6.1 Tenure Effects and Gravity

Theory and estimation. As shown in Proposition 3, our theory gives rise to a gravity equation for bilateral trade. This equation can be estimated following standard steps from the literature. In particular, taking logs of equation (3.15), rearranging terms, and omitting time superscripts for simplicity, we obtain

$$\begin{aligned}
 \log(X_{ij}) = & \overbrace{\log\left(\frac{L_i}{f_{ei}}(w_i)^{\frac{\sigma-1-\sigma\theta}{\sigma-1}}\right)}^{\equiv E_i} \\
 & + \overbrace{\log\left(X_j \left[\sum_{k \in \mathcal{J}} \frac{L_k}{f_{ek}} (\tau_{kj})^{-\theta} (w_k)^{\frac{\sigma-1-\sigma\theta}{\sigma-1}} (f_{kj} \tilde{f}(\hat{\lambda}_{kj}))^{\frac{\sigma-1-\theta}{\sigma-1}}\right]^{-1}\right)}^{\equiv M_j} \\
 & - \left[\overbrace{\theta \log(\tau_{ij})}^{TC_{var}} + \overbrace{\frac{\theta - (\sigma - 1)}{\sigma - 1} \log(f_{ij})}^{TC_{fix}} + \overbrace{\frac{\theta - (\sigma - 1)}{\sigma - 1} \log(\tilde{f}(\hat{\lambda}_{ij}))}^{history} \right].
 \end{aligned} \tag{3.16}$$

Equation (3.16) additively separates bilateral trade into exporter-specific terms, importer-specific terms, bilateral (fixed and variable) trade costs, and a bilateral tenure term. We estimate this equation using data on bilateral trade flows (cf. Appendix 3.A.A for details on the data). To control for exporter- and importer-specific terms, we include exporter (E_i) and importer fixed effects (M_j), respectively. To control for variable and fixed trade costs, we use standard proxies: the log of geographic distance, the absolute log-difference in GDP per capita, and dummies for contiguity, common language, common currency, a free trade agreement, and common legal origin, respectively.

The key novelty of our paper—and the main focus of our analysis—is the last summand in equation (3.16). This term implies that it is not sufficient to control for barriers to trade today, but that we need to take into account the

history of trade liberalizations as well. Our theory suggests that—due to incumbency effects—trade relations which liberalize later tend to have relatively lower trade flows today. To analyze whether this theoretical prediction is in line with the data, we would ideally have measures of pair-specific aggregate tenure levels. We are not aware of convincing such measures, and, in either case, they may suffer from endogeneity, as we would expect countries to liberalize earlier with more attractive trading partners. We therefore exploit two sets of historical events with arguably major effects on the history of trade: Colonial ties and the fall of the Iron Curtain. Colonial ties are a typical control variable in the literature, and our theory may help interpreting corresponding long-lasting consequences, which we also find in our estimation below.

The focus on the Iron Curtain is—to our knowledge—new. Specifically, we consider trade between pairs of countries that have been on different sides of the Iron Curtain during the Cold War.¹⁶ This provides a relevant set-up for our purposes for two reasons. First, the Iron Curtain was a major impediment to international trade, and its fall was associated with a dramatic shift in the trade environment (cf. Figure 3.A.E.2 in Appendix 3.A.E). Second, the original trade barriers—as well as the subsequent liberalization—between members and non-members of the Comecon were mostly a matter of geopolitics rather than a matter of strategic trade policy.

To analyze the effects of these historic events on trade today, we include two dummies as additional controls in our gravity regression: a dummy that indicates whether (i, j) had colonial ties, $\mathbb{1}[\text{Colonial Ties}]_{ij}$, and a dummy that indicates whether (i, j) had been separated by the Iron Curtain, $\mathbb{1}[\text{Sep. by Iron Curtain}]_{ij}$, where $\mathbb{1}[\cdot]$ is the indicator function that takes on value of 1 if the term in brackets is true and 0 otherwise. We consider countries to have colonial ties if they had a colonial relationship or a common colonizer after 1945 according to the [CEPII Gravity database](#). In summary, we run the

¹⁶We classify countries as having been behind the Iron Curtain if they were members of the Council for Mutual Economic Assistance (Comecon). The (full) members of this economic association were Albania, Bulgaria, Cuba, Czechoslovakia, East Germany, Hungary, Mongolia, Poland, Romania, the Soviet Union, and Vietnam (cf. [Encyclopædia Britannica](#)).

following regression

$$\begin{aligned} \log(X_{ij}) = & E_i + M_j + \zeta'_{ij}\beta + \phi_1 \times \mathbb{1}[\text{Sep. by Iron Curtain}]_{ij} \\ & + \phi_2 \times \mathbb{1}[\text{Colonial Ties}]_{ij} + v_{ij}, \end{aligned} \quad (3.17)$$

where ζ_{ij} is a (column) vector of the controls for bilateral trade frictions, and v_{ij} is an error term. The remaining variables are as previously discussed. Our main coefficients of interest are ϕ_1 and ϕ_2 , which measure the semi-elasticity for the Iron-Curtain and Colonial-Ties indicator, respectively. Our theory predicts $\phi_1 < 0$ and $\phi_2 > 0$. We run this regression for 2015 using OLS, and cluster standard errors at both the importer and the exporter level.¹⁷

Results. The results are summarized in Table 3.1.¹⁸ The full set of coefficients is reported in Table 3.A.F.4 of Appendix 3.A.F. The left-most column shows the coefficient on our two indicators of interest, controlling only for distance, contiguity, and the fixed effects. The second column depicts the outcome from including our full set of controls. The main insight from these results is twofold. First, country pairs that were separated by the Iron Curtain trade substantially less even 25 years after the fall of the Iron Curtain, while country pairs connected through colonial ties trade substantially more. The point estimate for the Iron Curtain indicator in column (2), for example, implies that countries that were separated by the Iron Curtain trade about 47 percent ($\approx 1 - \exp(-.63)$) less today, controlling for standard trade frictions. Second, most of the suggested discrepancies are not picked up by standard gravity controls, suggesting that these historic events have effects on trade that are orthogonal to conventional proxies for bilateral frictions.

Our theory provides a mechanism that can explain these patterns, as both the Iron Curtain and colonial ties have had a major impact on past trade flows. There may, however, likely be other effects at play. In particular, both colonial

¹⁷We exclude former Yugoslavia countries, since these countries were (economically) close to both Iron Curtain and other non-Iron Curtain countries. The overall share of relations where the separation indicator equals one is around 23%.

¹⁸A robustness check using PPML is provided in Table 3.A.F.6 of Appendix 3.A.F.

Table 3.1: Gravity Regressions, Year 2015, OLS

	Main Specification		Robustness			
	(1)	(2)	(3)	(4)	(5)	(6)
Sep. by Iron Curtain	-0.572*** (0.096)	-0.628*** (0.099)	-0.572*** (0.090)	-0.553*** (0.087)	-0.930*** (0.142)	-0.418*** (0.124)
Colonial Ties	0.781*** (0.095)	0.534*** (0.096)	0.507*** (0.096)	0.497*** (0.098)	0.949*** (0.232)	0.935*** (0.240)
Baseline Controls		✓	✓	✓	✓	✓
Tariffs & NTMs			✓	✓		
CIF/FOB Ratio				✓		
Comext Data					✓	✓
Mean of Dep. Var.	2.469	2.682	2.742	2.918	2.347	2.024
Adj. R-squared	0.737	0.765	0.782	0.796	0.802	0.799
Nr. of Observations	19,847	17,153	16,195	14,511	6,954	6,849

Notes. This table reports OLS results for regression (3.17). Standard errors clustered at both the importer and exporter levels are reported in parentheses. All regressions include importer and exporter fixed effects. “Sep. by Iron Curtain” refers to a dummy variable equal to one if the country pair was historically separated by the Iron Curtain and zero otherwise (see footnote 16). “Colonial Ties” refers to a dummy variable equal to one if the country pair either had a colonial relationship or a common colonizer after 1945. Dependent variable is the log of aggregate trade flows, measured in million USD (columns (1) to (4)) or Euros (columns (5) and (6)). Additional controls have been added as indicated in the respective column, but are not shown for clarity. A long table with coefficients on all controls is provided in Table 3.A.F.4 of Appendix 3.A.F. In column (1) only the log-distance, weighted by population, and a dummy for contiguity are included as additional controls. The “baseline controls” in columns (2) to (6) further include the absolute log-difference in GDP per capita, and dummies for common language, common currency, a free trade agreement, and common legal origin, respectively. “Tariffs” is the log of the simple average of bilateral gross tariffs. “NTMs” refers to non-tariff measures. “CIF/FOB Ratio” is the ratio of CIF to FOB prices. Columns (5) and (6) use data from [Comext](#). Column (6) includes only trade via air or sea transport.

* significant at 10%-level;** significant at 5%-level;*** significant at 1%-level.

Data source. [Atlas of Economic Complexity](#) (trade data), [CEPII/Head et al. \(2010\)](#) (covariates), [WITS](#) (import tariffs), [GTA](#) (NTMs), [ITIC](#) (CIF/FOB), [Comext](#) (trade data in (5)–(6)).

Results. Authors’ computations.

ties and the Iron Curtain may plausibly impact pair-specific trade policy and transport costs even today. In columns (3) to (6) we seek to control for these. In column (3), we control for trade policy, by including average tariffs and non-tariff measures (in addition to the RTA dummy). In column (4) we further proxy for trade costs by adding the CIF/FOB ratio from [ITIC](#). Adding these controls has little effect on our main coefficients of interest. Finally, in column

(6), we take a different approach: We consider trade via sea or air transport only, using data from [Comext](#). The idea being that such trade should not be affected by pair-specific investments in transportation infrastructure. The [Comext](#) dataset reports trade between EU members and non-members only. In column (5) we therefore repeat our baseline regression using this dataset as a reference point. Comparing columns (5) and (6) reveals that the coefficient on colonial ties is virtually unaffected, while that on the Iron Curtain is smaller when looking at sea and air trade only. Still, the point estimate suggests that trade between pairs of countries that were separated by the Iron Curtain during the Cold War is about 35% lower today, even when looking at sea and air trade only.¹⁹

In summary, these regressions point to sizable, long-lasting, and pair-specific implications of history for trade that seem not to be captured by conventional proxies for trade frictions or be entirely attributable to differences in pair-specific transportation infrastructure. These differences call for a better understanding. Our theory provides one channel that can help explain such patterns: incumbency effects that may prevent firm entry and exit upon trade liberalizations. Yet, our Iron Curtain and colonial ties dummies likely capture other effects as well. We therefore turn to our underlying mechanism next.

3.6.2 Average Firm Sales

In our model, history matters for trade flows today because market access costs decline with tenure, implying that there is less exit and entry in response to a trade shock. We cannot directly observe the fixed market access costs. In this section, we therefore exploit the structure of our model and use average firm sales to provide indirect evidence in support of these effects.

Theory and estimation. In our model, there is a tight connection between

¹⁹In Figure 3.A.E.3 of Appendix 3.A.B we reproduce the coefficients of interest from our main specification (column (2) in Table 3.1) for each of the years 1995 to 2015 separately. While both coefficients attenuated at the beginning of the period, they remained remarkably stable since the early 2000s, pointing to persistent differences, in line with our theory.

the fixed costs of serving a market and the average firm sales in that market, analogous to the canonical Melitz model. In particular, it is straightforward to show—see Appendix 3.A.C—that average firm sales from country i to j in period t , $\bar{r}_{ij}^t$, are equal to

$$\bar{r}_{ij}^t = \frac{\sigma\theta}{\theta - (\sigma - 1)} f_{ij} w_i^t \tilde{f}(\hat{\lambda}_{ij}^t), \quad (3.18)$$

where, recall, $\hat{\lambda}_{ij}^t$ captures the *tenure effect* in trade from i to j . Taking log-differences yields

$$\Delta \log [\bar{r}_{ij}^t] = \Delta \log [w_i^t] + \Delta \log [\tilde{f}(\hat{\lambda}_{ij}^t)] \quad \forall (i, j) \in \mathcal{I} \times \mathcal{J} \quad (3.19)$$

where $\Delta \log(x^t) := \log(x^t) - \log(x^{t-1})$. Now, suppose that at time t there is a drop in τ_{ij}^t for some country $i \in \mathcal{I}$. Then, controlling for exporter wages (and market size in the destination), $\tilde{f}(\hat{\lambda}_{ij}^t)$ weakly increases. This is because in the wave of a positive trade shock, the profit potential of firms from i in j increases. Hence, incumbency effects are less important for the decision of firms from i to serve j . In fact, in response to a sufficiently large positive shock there is market entry by new firms (who do not have any incumbency advantages), and tenure has no direct effect on trade from i to j . In either case, incumbency effects become less important for i 's exports to j , and average firm sales increase. Interestingly, this is in contrast to the canonical Melitz model, where controlling for general equilibrium effects on wages, a shock to τ_{ij} would not impact average firm sales.

The key advantage of equation (3.19) is that it allows studying these mechanisms using data on average firm sales. We combine data on average firm sales with average tariffs at the HS2-digit level—the “sector” level (cf. Appendix 3.A.A for details on the data). We study the implications of tariff cuts on average firm sales at the exporter-importer-sector level. We consider tariff changes, as they are arguably an observed shock to τ_{ij} without an (obvious)

effect on f_{ij} .²⁰ In Appendix 3.A.B, we consider a more general event instead: the EU Eastern enlargement. This analysis confirms our main insights presented here.

To test whether average firm sales from i to j increase in response to a tariff cut, we consider country pairs with sizable tariff cuts and compare these to country pairs with non-negative tariff changes. The latter allows controlling for general exporter and importer trends as suggested by our theory. To include only relatively large tariff changes, we consider tariff cuts in the top quartile of all tariff reductions in our baseline specification. This yields a still fairly modest cutoff of -2.3% , with a mean decline of gross tariffs of 6.7% . In order to mitigate concerns regarding mid-year reporting, we further consider changes in average firm sales from t to $t + 2$ in response to a tariff cut at t .²¹ We then estimate the following empirical counterpart of equation (3.19)

$$\begin{aligned} \Delta^2 \log [\bar{r}_{ij,g}^{t+2}] &= \phi \times \mathbb{1}[\text{Tariff Cut in the Top Quartile from } t \text{ to } t + 1]_{ij,g} \\ &\quad + E_{i,g}^{t+2} + M_{i,g}^{t+2} + \varepsilon_{ij,g}^{t+2}, \end{aligned} \quad (3.20)$$

where a subscript $_g$ denotes sector g , $\Delta^2 \log [\bar{r}_{ij,g}^{t+2}] := \log [\bar{r}_{ij,g}^{t+2}] - \log [\bar{r}_{ij,g}^t]$ is the log-change in average firm sales from t to $t + 2$, and $E_{i,g}^{t+2}$ and $M_{i,g}^{t+2}$ are exporter-sector-year and importer-sector-year fixed effects to capture exporter-sector trends in production costs and importer-sector trends in market size. $\varepsilon_{ij,g}^{t+2}$ is an error term. The coefficient of interest is ϕ . The previous discussions

²⁰One potential concern could be the introduction of a free trade agreement between two parties, which often requires additional effort for firms to benefit from preferential tariffs. To alleviate this concern we keep for each exporter only those destinations where the exporter's preferential status did not change within our period of analysis. For instance, when we look at changes of average firm sales from year t to $t + 2$, the exporter either has preferential tariffs throughout t to $t + 2$, or it faces MFN tariffs for this entire time span.

²¹This implies that we need to define how tariffs can evolve from $t + 1$ to $t + 2$ for the “treatment” (with tariff cut of at least 2.3% at time t) and the control groups (with non-negative tariff changes at t). We allow for a partial rebound of up to one third of the tariff reduction between $t + 1$ and $t + 2$ for the “treatment group,” and omit observations where average tariffs recover by more than that. Moreover, we verify that the observations in the “control group” did not experience any tariff cuts between $t - 1$ and $t + 2$, where $t - 1$ is also chosen because of potential mid-year reporting effects. All importer-exporter-sector-year observations that did not meet either of these criteria were excluded from the sample. We use “treatment” and “control group” for ease of reference, which should not indicate that we claim to estimate a causal effect of tariff reductions.

imply that in response to the tariff cut incumbency effects should become less important and, hence, our theory predicts $\phi > 0$. We estimate (3.20), and cluster the standard errors at the exporter-sector, importer-sector, and exporter-importer levels.

Results. Table 3.2 presents the results. The first three columns in the top panel depict the coefficients from our main specification with sector-year FEs, exporter-sector-year FEs, and both exporter-sector-year and importer-sector-year FEs, respectively. The coefficient of interest is stable and statistically significantly larger than zero throughout, as our model predicts. Through the lens of our model, the point estimates suggest that (relatively large) tariff cuts are associated with an eight percentage points relative decline in the importance of incumbency effects for trade from i to j . This empirical exercise therefore reveals that changes in firm sales move in accordance with our predictions. According to our theory, this implies that the benefit from the drop in τ_{ij} was partly offset by a relative increase in $\tilde{f}(\hat{\lambda}_{ij})$.

To conclude, we present several robustness tests. In the top panel, fourth column of Table 3.2 we apply 90% winsorizing to our dependent variable to test whether our results are driven by few large changes in average firm sales. This somewhat reduces our estimate but leaves it within one standard error deviation from our main coefficient in the third column. We also control for exporter-importer-year-specific trends, the results of which are reported in the fifth column of the top panel. The coefficient is still in the same ballpark, although the standard error increases substantially. This is due to the fact that there is little variation in our tariff indicator within a country pair in a given year across sectors.²² In the bottom panel we document the sensitivity of our results with respect to (i) the time span we choose to measure the growth in average firm sales, and (ii) how long the reference group without tariff cuts should maintain at least the same tariff level. The results look similar in all

²²Only about 1% of the observations in the regression sample have variation in the treatment dummy within all three sets of fixed effects.

Table 3.2: Impact of Tariff Changes on Average Firm Sales

	Dep. Var.: Log-Change in Av. Firm Sales from t to $t+2$ Control Group: No Tariff Cuts from $t-1$ to $t+2$				
	Main Specification			Robustness	
				90% Wins.	Pair-Year FE
$\mathbb{1}[\text{Tariff Cut}]$	0.098*** (0.016)	0.089*** (0.016)	0.075** (0.033)	0.059** (0.026)	0.063 (0.072)
Sector-Year FE	✓				
Exp-Sector-Year FE		✓	✓	✓	✓
Imp-Sector-Year FE			✓	✓	✓
Mean of Dep. Var.	0.094	0.093	0.091	0.090	0.089
Adj. R-squared	0.023	0.081	0.091	0.100	0.106
Nr. of Observations	191,979	187,429	151,694	151,694	145,762
	Dep. Var.: Log-Change in Av. Firm Sales from t to $t+\bar{t}$ Control Group: No Tariff Cuts from $t-\underline{t}$ to $t+\bar{t}$				
	Robustness				
	$\bar{t}=2, \underline{t}=0$	$\bar{t}=2, \underline{t}=2$	$\bar{t}=3, \underline{t}=1$	$\bar{t}=3, \underline{t}=2$	$\bar{t}=3, \underline{t}=3$
$\mathbb{1}[\text{Tariff Cut}]$	0.058** (0.029)	0.101*** (0.039)	0.108** (0.045)	0.131** (0.052)	0.141** (0.057)
Exp-Sector-Year FE	✓	✓	✓	✓	✓
Imp-Sector-Year FE	✓	✓	✓	✓	✓
Mean of Dep. Var.	0.100	0.082	0.121	0.103	0.093
Adj. R-squared	0.091	0.089	0.099	0.098	0.100
Nr. of Observations	187,631	124,313	108,518	88,235	73,128

Notes. This table reports OLS results from regressing log-changes in average firm sales on an indicator for relatively large tariff reductions, as described in Section 3.6.2. The regression is outlined in equation (3.20). Standard errors clustered at the exporter-sector, importer-sector, and exporter-importer levels are reported in parentheses. The full sample spans the years 1997–2014. The “treatment group” consists of observations with tariff reductions belonging to the top quartile of overall tariff cuts. The “control group” had non-negative changes in average tariffs between the years $t - \underline{t}$ and $t + \bar{t}$. The analysis is conducted at the sector level (HS2-digit).

* significant at 10%-level; ** significant at 5%-level; *** significant at 1%-level.

Data source. EDD (average firm exports), WITS (import tariffs).

Results. Authors’ computations.

specifications, with the coefficient slightly increasing for longer time horizons.²³ Additional robustness checks with respect to our data cleaning and variable definitions are presented in Appendix 3.A.B.

²³Boehm et al. (2020) document that tariff changes at the HS6-digit level have a ramp-up phase of up to seven years. Unfortunately, average tariffs at the sector level fluctuate substantially more often than those at the product level, and thus our aggregation level impedes investigating tariff shocks over such a long time span.

3.7 Conclusion

In this paper, we have introduced incumbency effects into a Melitz model of international trade that nests the canonical (static) version of this model as a special case. We derive a gravity equation for international trade that shows how trade is not only shaped by trade frictions today—as typically assumed in the literature—, but also by the history of trade liberalizations. Our theory can thus help explain a home bias and, more generally, why historic events can have long-lasting implications for international trade. Our theoretical predictions are in line with the relevance of colonial ties and the Iron Curtain for trade flows today, and with novel, indirect evidence on the response of fixed cost of exporting to trade liberalizations.

Our paper provides a first attempt to integrating sunk market access costs into a dynamic general equilibrium model of international trade with many asymmetric countries. In our model, the economy immediately jumps to a new steady state upon a trade liberalization. This allows introducing incumbency effects into the canonical (static) version of the Melitz model, while maintaining its tractability. Future work may set out to allow for richer dynamics, thereby providing additional insights on the persistence of incumbency effects in international trade.

3.A Appendices to Chapter 3

3.A.A Details on Data

In this appendix, we provide details on our data.

Gravity regressions. To run our gravity regressions, we use data on bilateral goods trade flows for the years 1995 to 2015 from the [Atlas of Economic Complexity](#).²⁴ Data on distance, GDP per capita, colonial ties, contiguity, common language, common currency, free trade agreements, and common legal origin are obtained from the widely-used [CEPII Gravity database](#).

The tariff data in column (3) are from the World Integrated Trade Solution ([WITS](#)), which provides us with most-favored-nation (MFN) and preferential tariffs at the HS6-digit level. Whenever preferential tariffs are available, we set the tariff between two countries to that value, and we take the MFN import tariff for all remaining exporters that are WTO members in a given year. We restrict the sample to WTO members to make sure that we set the correct non-preferential tariff rates. However, most countries do not report tariffs in every year. Following [Felbermayr et al. \(2019\)](#), we fill in missing tariff observations by using the closest preceding year we observe. This accounts for the fact that countries tend to report only upon policy changes. To aggregate tariffs to the country-pair-year level, we use simple averages.

Data on non-tariff measures (NTMs) in column (3) are from the Global Trade Alert ([GTA](#)). The GTA data documents NTMs at the country-pair-product level. We use all 29 measures that are categorized according to the [MAST chapters](#), and include those barriers that were in force in 2015. For the exporter-related measures, we attribute the non-tariff barrier to exports from i to j if i is the implementing jurisdiction while j is the affected one. For all other measures, we attribute the NTM to exports from i to j if j is the

²⁴Since trade flows are mostly reported twice (i.e., as imports by the importer and as exports by the exporter), researchers have to decide which information to use, given that these statistics rarely coincide (not only due to discrepancies between reporting imports including costs of insurance and freight (CIF), and exports free on board (FOB)). The [Atlas of Economic Complexity](#) has developed a unique methodology to extract arguably more consistent flows out of the raw COMTRADE data.

implementing jurisdiction and i the affected one. We then add the log of one plus the number of affected products as a covariate. This yields 29 different control variables for NTMs that are included in column (3).

CIF/FOB ratios in column (4) of Table 3.1 are from the OECD International Transport and Insurance Costs (ITIC) database.²⁵

Finally, the data on trade by mode of transportation used in columns (5) and (6) is from Comext, which reports trade between EU members and non-members.

Average firm sales. Data on average firm exports at the exporter-importer-HS2-digit-year level are from the Exporter Dynamics Database (EDD). This dataset provides us with an unbalanced panel of average firm exports for 62 exporting countries and 95 sectors for the years 1997–2014. Since we are interested in the aggregate implications of trade liberalizations, we verify that these data are in line with aggregate trends. In particular, we compare total exports from country i to j in the EDD to those reported in the *Atlas of Economic Complexity*, and require the two series to be positively correlated over time within an exporter-importer-sector cell. For our baseline results, we chose a minimum correlation of 0.6, and dropped observations with a lower value (roughly 20% of the observations). Moreover, for our main results, we omit observations where total exports belong to the bottom 5% within a sector. We do this because country pairs with very small exports (the median cutoff value is around 2,000 USD) add a lot of noise when we compute log-changes in average firm sales (cf. Figure 3.A.E.1). In Appendix 3.A.B, we show that our results are robust to altering both of these steps.

The tariff data are as detailed above. We take the simple average of import tariffs within HS2-digit sectors for each country pair to have the same aggregation level as the average firm sales from the EDD.

²⁵Many of these values were imputed using a gravity model (Miao and Fortanier, 2017), and we thus prefer omitting this variable for our main results, but reassuringly, adding it has little effect on our main coefficients of interest.

3.A.B Additional Empirical Results

In this part of the Appendix, we provide additional results for our empirical analyses.

Motivating evidence in Section 3.2. Table 3.A.F.1 shows robustness of our insights from Figure 3.2 to using fewer fixed effects and controls. Column (1) shows the raw correlations in the data, Columns (2)–(4) and (5)–(7) results from regressions without and with firm fixed effects, respectively. In Columns (2) and (5), we add product fixed effects and control for log-sales, allowing for the coefficient to differ at the product level. In Columns (3) and (6), we control for product-quintile fixed effects. Lastly, in Columns (4) and (7) we control for product-quintile fixed effects and log-sales, allowing the coefficient to differ at the product-quintile level. Column (7) replicates the regression summarized in Figure 3.2. In all regressions, we find a systematic negative relation between tenure levels and exit propensities in response to the shock.

Tables 3.A.F.2 and 3.A.F.3 provide further robustness checks for our motivating regressions in Figure 3.2 and Table 3.A.F.1. In Table 3.A.F.2, we replicate our analysis of the Venezuela shock, but looking at exit immediately after the shock (i.e., no exports in the period from August 2009 to July 2010). In Table 3.A.F.3, we run regressions analogous to Table 3.A.F.1, but where we exploit our full dataset, looking at firm exit from all country-product pairs, and including all years in our data. The advantage is that it allows us to include firm-product-year fixed effects, thereby controlling for firm-product-specific shocks. The disadvantage is that we are not considering exit in response to a major aggregate shock—our main interest. These robustness checks confirm our insights from Figure 3.2 and Table 3.A.F.1.

Gravity regressions in Section 3.6.1. Figure 3.A.E.2 shows the evolution over time of trade between (former) Council for Mutual Economic Assistance (Comecon) members and (former) OECD members relative to trade between OECD members, normalized by GDP. The group of OECD countries in Figure 3.A.E.2 includes only members that joined before 1991. For the graph, we

lump former Soviet Union countries back together after 1991, and merge the Czech Republic and the Slovak Republic. This figure not only reveals much less Comecon-OECD trade in the 1970s and 80s, but also a dramatic rise after the fall of the Iron Curtain in the course of the 1990s.

Table 3.A.F.4 replicates Table 3.1, showing the coefficients on all controls but for the 29 NTMs. Coefficients on NTMs in column (3) of Table 3.1 are provided in Table 3.A.F.5. Table 3.A.F.6 replicates Table 3.A.F.4, using PPML instead of OLS. This robustness check confirms our conclusions from Section 3.6.1.

In Figure 3.A.E.3 we reproduce the coefficients of interest from our main specification (second column in Table 3.1) for each of the years 1995 to 2015 separately. Both coefficients attenuated at the beginning of the sample period but remained remarkably stable since the early 2000s, pointing to persistent differences.

Robustness of tariff regressions in Section 3.6.2. Tables 3.A.F.7 and 3.A.F.8 provide additional robustness checks for Table 3.2.

The simple average of tariffs within a sector might not reflect the true sector-level tariff burden for a trade relationship in that a country pair may not (want to) trade the underlying goods that cause mean tariffs to fluctuate. To alleviate this concern, Table 3.A.F.7 reproduces Table 3.2 using the weighted average of tariffs within a sector rather than the simple average. The weights are the bilateral trade shares within a sector. The results are reassuring, as the insights from above are entirely confirmed by these estimates. Intuitively, the suggested impact of tariff changes is generally larger in this specification.

In Table 3.A.F.8 we recompute the main specification in the third column of the top panel in Table 3.2 to further test the sensitivity of our main results with respect to our data cleaning steps and the definition of the treatment. In the top panel, we show that the results look very similar if we allocate observations with above-median tariff cuts to the “treatment group,” allow for no rebound or a rebound of up to 50% of the tariff cut, or include the tariff level in year t and an indicator for preferential tariffs as additional control variables.

Moreover, for the main results, we omitted observations with very small total exports (bottom 5% within a sector), or with a correlation between total exports in the aggregate data and the firm-level statistics below 0.6. In the bottom panel of Table 3.A.F.8, we tried using no minimum total exports level (first column) or no minimum correlation value (third column). We also tried a more restrictive setting with total exports required to be either larger than the bottom 10% within a sector (second column) or have a correlation value of at least 0.8 (fourth column). All these adjustments lead to a coefficient close to the main one in Table 3.2.

EU Eastern enlargement and average firm sales. A potential concern about our tariff regressions is that—even though we include only relatively large tariff cuts in our sample—these represent relatively modest shocks. We therefore corroborate our tariff regressions by looking at the 2004 EU East enlargement. The advantage of exploiting this episode is that this presumably represented a large shock to variable trade barriers. The downturn, however, is that it likely had an additional direct effect on fixed market access costs f_{ij} . Nevertheless, to the extent that fixed market access costs decreased in the wave of the trade liberalizations associated with EU accession (due to, e.g., the abolishment of rules of origin), these changes should work against the effect predicted by our theory. We will get back to this point momentarily.

The regression equation follows naturally from Equation (3.19). We want to measure whether the (cutoff) tenure-dependent component of fixed costs has increased for newly created EU member pairs—where at least one new member is involved—relative to established EU pairs. The treaty of accession was signed in 2003, and we therefore use 2002 as base period. We then consider changes in average firm sales from the base period to 2005, 2006, 2007, and 2008, in order to verify that the conclusions do not hinge on the time frame considered. Analogous to Section 3.6.2, we then run the following regressions

$$\begin{aligned} \Delta^{\bar{t}} \log \left[\bar{r}_{ij}^{2002+\bar{t}} \right] &= E_i^{2002+\bar{t}} + M_j^{2002+\bar{t}} + \mu_0 \times \mathbb{1}[\text{EU Pair}]_{ij}^{2002+\bar{t}} \\ &\quad + \mu_1 \times \mathbb{1}[\text{New EU Pair}]_{ij}^{2002+\bar{t}} + \nu_{ij}^{2002+\bar{t}}, \end{aligned} \quad (3.21)$$

where $\Delta^{\bar{t}} \log [\bar{r}_{ij}^{2002+\bar{t}}] \equiv \log [\bar{r}_{ij}^{2002+\bar{t}}] - \log [\bar{r}_{ij}^{2002}]$ is the log-change in average firm sales from 2002 to $2002 + \bar{t}$ with $\bar{t} \in \{3, 4, 5, 6\}$, and $E_i^{2002+\bar{t}}$ and $M_j^{2002+\bar{t}}$ are exporter and importer FEs, respectively. $\nu_{ij}^{2002+\bar{t}}$ is an error term.

To run this regression, we use average firm sales at the country-pair level reported in the OECD Trade by Enterprise Characteristics (TEC) database, as the EDD covers only 4 EU countries. We add country-pair level data from the EDD for country pairs that are not covered by the OECD TEC database, thereby following [Adão et al. \(2020\)](#). To be able to identify $E_i^{2002+\bar{t}}$, we include all trade relations in our dataset. The coefficient μ_0 measures whether, in general, there is a different trend in $\bar{r}_{ij}$ among established EU pairs and non-EU trade relations. Our coefficient of interest is μ_1 . This coefficient measures a semi-elasticity that determines how the tenure-dependent component of fixed costs varies, on average, for new EU pairs relative to old ones. The difference between new EU pairs and non-EU relations is given by the sum of these two estimates. In our regressions, we apply two-way clustering at the importer and exporter levels, and perform data cleaning exercises that are analogous to those for the EDD as detailed in [Appendix 3.A.A](#) above, i.e., we use a correlation cutoff of 0.6 and omit very small country pairs. We verified that the main results remain largely unchanged when we adjust these steps in the same ways as the robustness tests for the tariff regressions above.

Since the EU Enlargement is a negative shock to trade barriers between newly created EU pairs, we expect a rise in $\tilde{f}(\hat{\lambda})$ for these relations relative to pre-existing within-EU relations (i.e., $\mu_1 > 0$). To the extent to which the EU accession has a negative effect on f_{ij} , this provides a lower bound on the incumbency effect. The results are presented in [Table 3.A.F.9](#). The estimates suggest that, on average, the fixed costs of market access increased by about 30 percentage points more among new EU pairs relative to old ones, with the coefficient's magnitude somewhat increasing when we inspect a longer time horizon. The results therefore point to the same direction as our theory and the insights from the general tariff exercise above. Including an importer fixed effect, and hence controlling for overall market size trends, has a minor impact on these estimates.

Bulgaria and Romania entered the EU only in 2007. For our main results in Table 3.A.F.9, we therefore excluded these two countries (as exporters and importers). In Table 3.A.F.10, we provide a robustness check including them as new EU members throughout, which yields similar results. Tables 3.A.F.11 and 3.A.F.12 reproduce Table 3.A.F.9 exchanging the base period with 2001 and 2003, respectively, which yields the same conclusions.

3.A.C Details on Average Firm Exports

In this appendix, we provide further details on the response of average firm sales to trade shocks. This forms the basis of our regressions in Section 3.6.2.

Using the Pareto distribution of firm productivities, it is straightforward to show that average firm sales from country i to destination j in period t , $\bar{r}_{ij}^t$, are equal to

$$\bar{r}_{ij}^t = \frac{\sigma\theta}{\theta - (\sigma - 1)} \pi_{ij}^{v,t}(\varphi_{ij}^t), \quad (3.22)$$

where, recall, $\pi_{ij}^{v,t}(\varphi_{ij}^t)$ are variable profits of the least productive firm in i that is serving market j in period t

$$\pi_{ij}^{v,t}(\varphi_{ij}^t) = \frac{1}{\sigma} \left(\frac{\sigma}{\sigma - 1} \right)^{1-\sigma} (\tau_{ij}^t w_i^t)^{1-\sigma} (P_j^t)^{\sigma-1} X_j^t (\varphi_{ij}^t)^{\sigma-1}.$$

Suppose there is a trade shock at time t . Analogous to the proof of Proposition 3 in Appendix 3.A.D.e, we can define a hypothetical tenure level $\hat{\lambda}_{ij}^t \in [0, \lambda_{ij}^t(i, \varphi_{ij}^t)]$ such that

$$\pi_{ij}^{v,t}(\varphi_{ij}^t) = f_{ij} w_i^t \tilde{f}(\hat{\lambda}_{ij}^t). \quad (3.23)$$

Using equation (3.23) in equation (3.22) yields

$$\bar{r}_{ij}^t = \frac{\sigma\theta}{\theta - (\sigma - 1)} f_{ij} w_i^t \tilde{f}(\hat{\lambda}_{ij}^t). \quad (3.24)$$

$\hat{\lambda}_{ij}^t$ in equation (3.24) measures the importance of tenure for i 's exports to j .

If, in response to the shock, some firms from i start newly serving j , $\hat{\lambda}_{ij}^t = 0$ and tenure has no effect because all firms would have also served j with zero tenure. Otherwise, if some firms stop serving j , $\hat{\lambda}_{ij}^t = \lambda_j^t(i, \underline{\varphi}_{ij}^t)$, i.e., $\hat{\lambda}_{ij}^t$ is equal to the tenure of the least productive active firm, and even a marginal decline in tenure would have resulted in fewer exports. In general, however, the incumbency effects may also imply that the set of firms which serve j from i remains unchanged in response to the shock, and we have $\hat{\lambda}_{ij}^t \in [0, \lambda_j^t(i, \underline{\varphi}_{ij}^t)]$. Still, $\hat{\lambda}_{ij}^t$ measures the importance of incumbency effects from i to j in all the above cases, as discussed in Appendix 3.A.D.e.

Now, suppose that at time t , τ_{ij}^t decreases. $\underline{\varphi}_{ij}^{t-1}$ is the productivity of the least productive firm serving j from i prior to the shock. Holding constant w_i^t and the market size and competition in j , the decline in τ_{ij}^t implies that $\pi_{ij}^{v,t}(\underline{\varphi}_{ij}^{t-1})$ increases, and hence

$$\pi_{ij}^{v,t}(\underline{\varphi}_{ij}^{t-1}) > f_{ij} w_i^t \tilde{f}(\hat{\lambda}_{ij}^{t-1}).$$

Then, there are two possibilities. Either (i) there is a $\hat{\lambda}_{ij}^t \in [0, \hat{\lambda}_{ij}^{t-1})$ such that

$$\pi_{ij}^{v,t}(\underline{\varphi}_{ij}^{t-1}) = f_{ij} w_i^t \tilde{f}(\hat{\lambda}_{ij}^t),$$

in which case $\underline{\varphi}_{ij}^{t-1} = \underline{\varphi}_{ij}^t$, i.e., there is no market entry by firms from i in j , and $\tilde{f}(\hat{\lambda}_{ij}^t) > \tilde{f}(\hat{\lambda}_{ij}^{t-1})$. Or (ii) there is market entry by firms from i in j , implying that $\underline{\varphi}_{ij}^{t-1} > \underline{\varphi}_{ij}^t$, $0 = \hat{\lambda}_{ij}^t \leq \hat{\lambda}_{ij}^{t-1}$, and $\tilde{f}(\hat{\lambda}_{ij}^t) \geq \tilde{f}(\hat{\lambda}_{ij}^{t-1})$. In either case, $\tilde{f}(\hat{\lambda}_{ij}^t)$ weakly *increases* and, hence, the importance of incumbency effects for exports from i to j weakly *declines*. In turn this implies that average firm sales from i to j weakly increase as stated in Section 3.6.2.

3.A.D Proofs

3.A.D.a Proof of Lemma 1

To derive the desired expression, note that aggregate expenditures in country i on the CES aggregator are equal to household consumption plus firm

expenditure to cover the fixed cost of market access, i.e.,

$$X_i^t = X_i^{h,t} + X_i^{f,t}, \quad (3.25)$$

where $X_i^{h,t} := L_i x_i^t$ denotes aggregate household expenditure and $X_i^{f,t}$ aggregate firm expenditure. The remainder of the proof proceeds in three steps. We first characterize household expenditure, then firm expenditure, and finally derive aggregate demand.

Step 1. Households in country i can invest in an aggregate asset comprising ownership of all domestic firms (entrants and operating firms).²⁶ Hence, asset markets are in equilibrium if we have

$$L_i a_i^t = V_i^t \quad (3.26)$$

at all time t , where a_i^t denotes asset holdings of the representative household at the beginning of the period—after announcement of any potential changes to the trade environment but before payments of the per-period interest rate. V_i^t is the total value of all firms in country i as detailed momentarily. Using equation (3.26) in equation (3.3) yields

$$X_i^{h,t} = L_i w_i^t - \left[\mathbb{E}_t \left[\frac{V_i^{t+1}}{1 + r_i^{t+1}} \right] - V_i^t \right], \quad (3.27)$$

i.e., in equilibrium households exactly absorb any changes in the value of the aggregate asset. With the rate of return on the aggregate asset given by equation (3.4), this is optimal.

The total value of all firms in country i at time t is given by

$$V_i^t \equiv \sum_{s=0}^t M_{ei}^s \int_{\varphi \in \Phi} v^t(i, s, \varphi) g(\varphi) d\varphi, \quad (3.28)$$

²⁶We assume that households can invest in domestic firms only. This is to simplify the exposition, but not essential. In particular, in equilibrium, the interest rate is the same in all countries, which immediately implies that our equilibrium would also apply to a case with investment in a global asset.

where $v^t(i, s, \varphi)$ denotes the value of firm (i, s, φ) , i.e., of a firm with productivity φ that entered country i in period $s \leq t$. M_{ei}^s denotes the mass of firms that entered in s .²⁷ This value is given by the present value of expected future earnings net of fixed costs

$$v^t(i, s, \varphi) = \mathbb{E}_t \left[\sum_{\varsigma=t}^{\infty} \left[-f_{ei} w_i^{\varsigma} + \sum_{j \in \mathcal{J}^{\varsigma}(i, s, \varphi)} \pi_j^{\varsigma}(i, s, \varphi) \right] R_i^{t, \varsigma} \right], \quad (3.29)$$

where, recall, $R_i^{t, \varsigma}$ is the discount factor from period $\varsigma \geq t$ to the beginning of period t , $\mathcal{J}^{\varsigma}(i, s, \varphi)$ denotes the set of markets that the firm serves in $\varsigma \geq s$, and

$$\pi_j^{\varsigma}(i, s, \varphi) := \pi_{ij}^{v, \varsigma}(\varphi) - f_{ij}^{\varsigma}(\lambda_j^{\varsigma}(i, s, \varphi)) \quad (3.30)$$

are the firm's profits net of fixed market access cost from serving destination j in period ς . $\lambda_j^{\varsigma}(i, s, \varphi)$ denotes the firm's tenure in destination j at time ς . Equation (3.29) can be written recursively as

$$\begin{aligned} \mathbb{E}_t \left[\frac{v^{t+1}(i, s, \varphi)}{1 + r_i^{t+1}} \right] - v^t(i, s, \varphi) &= \sum_{\varsigma=t+1}^{\infty} \left\{ \mathbb{E}_t \left[\left(-f_{ei} w_i^{\varsigma} + \sum_{j \in \mathcal{J}^{\varsigma}(i, s, \varphi)} \pi_j^{\varsigma}(i, s, \varphi) \right) R_i^{t, \varsigma} \right] \right. \\ &\quad \left. - \mathbb{E}_t \left[\left(-f_{ei} w_i^{\varsigma} + \sum_{j \in \mathcal{J}^{\varsigma}(i, s, \varphi)} \pi_j^{\varsigma}(i, s, \varphi) \right) R_i^{t, \varsigma} \right] \right\} \\ &\quad - \left[-f_{ei} w_i^t + \sum_{j \in \mathcal{J}^t(i, s, \varphi)} \pi_j^t(i, s, \varphi) \right] \\ &= - \left[-f_{ei} w_i^t + \sum_{j \in \mathcal{J}^t(i, s, \varphi)} \pi_j^t(i, s, \varphi) \right]. \end{aligned} \quad (3.31)$$

The difference between the expected firm value tomorrow and today's firm value (after realization of any potential changes in the trade environment in

²⁷The value of the firm depends on the time of entry, s , because firms in country i with productivity φ that entered in different periods potentially face different fixed costs of market access and—as a consequence—may find it optimal to serve different sets of destinations.

the current period) are just the dividend payments in the current period.²⁸ Combining equations (3.27), (3.28), and (3.31) yields

$$X_i^{h,t} = L_i w_i^t - \sum_{s=0}^t \left[M_{ei}^s \left[f_{ei} w_i^t - \int_{\varphi \in \Phi} \sum_{j \in \mathcal{J}^t(i,s,\varphi)} \pi_j^t(i,s,\varphi) g(\varphi) d\varphi \right] \right]. \quad (3.32)$$

Step 2. The period- t demand for the CES aggregator of firm (i, s, φ) is

$$\sum_{j \in \mathcal{J}^t(i,s,\varphi)} (f_{ij}^t(\lambda_j^t(i,s,\varphi)) - f_{ij} w_i^t).$$

Aggregating over all firms yields

$$X_i^{f,t} = \sum_{s=0}^t \left[M_{ei}^s \int_{\varphi \in \Phi} \sum_{j \in \mathcal{J}^t(i,s,\varphi)} (f_{ij}^t(\lambda_j^t(i,s,\varphi)) - f_{ij} w_i^t) g(\varphi) d\varphi \right]. \quad (3.33)$$

Step 3. Combining equations (3.25), (3.30), (3.32), and (3.33) and simplifying terms yields

$$X_i^t = L_i w_i^t - \sum_{s=0}^t \left[M_{ei}^s \left[f_{ei} w_i^t - \int_{\varphi \in \Phi} \sum_{j \in \mathcal{J}^t(i,s,\varphi)} (\pi_{ij}^{v,t}(\varphi) - f_{ij} w_i^t) g(\varphi) d\varphi \right] \right].$$

²⁸Note that potential valuation gains (or losses, for that matter) due to unanticipated changes in the trade environment are reflected in V_i^t . In particular by equation (3.29) it holds

$$\begin{aligned} v^t(i,s,\varphi) - \mathbb{E}_{t-1} [v^t(i,s,\varphi)] &= \sum_{\varsigma=t}^{\infty} \left\{ \mathbb{E}_t \left[\left(-f_{ei} w_i^\varsigma + \sum_{j \in \mathcal{J}} \pi_j^\varsigma(i,s,\varphi) \right) R_i^{t,\varsigma} \right] \right. \\ &\quad \left. - \mathbb{E}_{t-1} \left[\left(-f_{ei} w_i^\varsigma + \sum_{j \in \mathcal{J}} \pi_j^\varsigma(i,s,\varphi) \right) R_i^{t,\varsigma} \right] \right\}, \end{aligned}$$

i.e., the unanticipated change in V_i^t in response to a trade shock is equal to the present value of the change in expected future earnings. Hence, the valuation gain is one-for-one reflected in households' asset holdings, and it will not impact household consumption beyond effects implied by equation (3.31).

Using equations (3.6), (3.7), labor market clearing in i

$$L_i = \sum_{s=0}^t \left[M_{ei}^s \left[f_{ei} + \int_{\varphi \in \Phi} \sum_{j \in \mathcal{J}^t(i, s, \varphi)} \left(\frac{y_{ij}^t(\varphi) \tau_{ij}^t}{\varphi} + f_{ij} \right) g(\varphi) d\varphi \right] \right],$$

and total sales in j —conditional on serving the market—of a firm from i with productivity φ

$$y_{ij}^t(\varphi) p_{ij}^t(\varphi) = p_{ij}^t(\varphi)^{1-\sigma} (P_j^t)^{\sigma-1} X_j^t,$$

yields after rearranging terms the expression in Lemma 1.

□

3.A.D.b Proof of Lemma 2

Let there be a trade shock at time t . To show the desired result, we proceed in three steps. We first show that if (i) any potential firm entry in response to the shock is at time t , and (ii) all firms revise their market access decisions at time t and then continue following this decision, equilibrium wages, prices, and aggregate demand are constant over time. We then show that if all potential entry in response to the shock is at time t , all firms will indeed perpetually follow their initial market access decision. We finally show that in such case no firm has an incentive to enter at a later stage.

Step 1. With the trade shock at time t and no future changes in the trade environment, we have $\tau_{ij}^\varsigma = \tau_{ij}^t$ for every i, j , and $\varsigma \geq t$. Similarly, with all firm entry in response to the shock (if any) at time t and all market access decisions fixed at time t , we have $M_i^\varsigma = M_i^t := \sum_{s=0}^t M_{ei}^s$, and $\mathcal{J}^\varsigma(i, s, \varphi) = \mathcal{J}^t(i, s, \varphi) \forall (i, s, \varphi)$,

$$Y_i^\varsigma = \sum_{s=0}^t \left[M_{ei}^s \int_{\varphi \in \Phi} \sum_{j \in \mathcal{J}^t(i, s, \varphi)} \left(\frac{\sigma}{\sigma-1} \right)^{1-\sigma} (\tau_{ij}^t w_i^\varsigma)^{1-\sigma} (P_j^\varsigma)^{\sigma-1} X_j^\varsigma \varphi^{\sigma-1} g(\varphi) d\varphi \right], \quad (3.34)$$

where $(P_j^\varsigma)^{\sigma-1}$ is given by

$$(P_j^\varsigma)^{\sigma-1} = \left[\sum_{i \in \mathcal{J}} \sum_{s=0}^t M_{ei}^s \int_{\varphi \in \Phi} \mathbb{1}[j \in \mathcal{J}^t(i, s, \varphi)] \left(\frac{\sigma}{\sigma-1} \frac{\tau_{ij}^t w_i^\varsigma}{\varphi} \right)^{1-\sigma} g(\varphi) d\varphi \right]^{-1}. \quad (3.35)$$

$\mathbb{1}[\cdot]$ is an indicator function that takes on value one if the term in brackets is correct, and zero otherwise. Moreover, Lemma 1 implies

$$X_i^\varsigma = \frac{\sigma}{\sigma-1} w_i^\varsigma \left[L_i - \sum_{s=0}^t \left[M_{ei}^s \left[f_{ei} + \int_{\varphi \in \Phi} \sum_{j \in \mathcal{J}^t(i, s, \varphi)} f_{ij} g(\varphi) d\varphi \right] \right] \right]. \quad (3.36)$$

Invoking balanced trade, equations (3.34) to (3.36) can be reduced to a system of I equations in the I wage rates. This system of equations is the same in every period $\varsigma \geq t$, i.e., indeed wages and, hence, prices, interest rates, and aggregate demand in each country are constant over time.

Step 2. With all prices and aggregate demand constant over time, so are variable profits of all firms in country i with productivity φ serving country j . Moreover, all wages and interest rates constant implies that $\tilde{f}_{ij}(\cdot)$ is weakly decreasing over time for all firms serving a market (i.e., for whom λ is increasing), while it is constant over time for any given λ . The former implies that all firms who found it optimal to start serving a market at time t also find it optimal to continue doing so in future periods. The latter implies that for all firms for whom it was optimal *not* to serve market j at t , it is also optimal not to do so in future periods. Hence, indeed, all firms perpetually follow their market access decision.

Step 3. Finally, free entry implies that

$$\sum_{j \in \mathcal{J}} \int_{\varphi_{ij}^{0,t}}^{\infty} (\tilde{\pi}_{ij}^{v,t}(\varphi) - \tilde{f}_{ij}^t(0)) g(\varphi) d\varphi \leq \tilde{f}_{ei}^t, \quad (3.37)$$

where in case of strictly positive entry in response to the shock the condition holds with equality. Condition (3.37) also remains unchanged in all periods $\varsigma \geq t$. Therefore, the fact that additional entry is not profitable at the time of the shock implies that it is also not profitable in all subsequent periods.

□

3.A.D.c Proof of Lemma 3

To show the result, we first derive the mass of entrants in the first period, equation (3.12), and then proceed by contradiction.

Step 1. Labor market clearing at time $t = 0$ requires

$$M_{ei}^0 \left[\sum_{j \in \mathcal{J}} \int_{\underline{\varphi}_{ij}^{0,0}}^{\infty} \left(\frac{y_{ij}^0(\varphi) \tau_{ij}^0}{\varphi} + f_{ij} \right) g(\varphi) d\varphi + f_{ei} \right] = L_i,$$

which, using equations (3.6) and (3.7), and the fact that variable profits are a constant fraction of revenues, can be rewritten as

$$M_{ei}^0 \left[\sum_{j \in \mathcal{J}} \int_{\underline{\varphi}_{ij}^{0,0}}^{\infty} \left(\frac{\pi_{ij}^{v,0}(\underline{\varphi}_{ij}^{0,0})}{w_i^0} \left(\frac{\varphi}{\underline{\varphi}_{ij}^{0,0}} \right)^{\sigma-1} (\sigma - 1) + f_{ij} \right) g(\varphi) d\varphi + f_{ei} \right] = L_i. \quad (3.38)$$

By Lemma 2 we know that—in the absence of shocks—the equilibrium is constant. Hence, equation (3.9) implies

$$\begin{aligned} \pi_{ij}^{v,0}(\underline{\varphi}_{ij}^{0,0}) &= \tilde{f}_{ij}^0(0) \\ &= f_{ij} w_i^0 \tilde{f}(0), \end{aligned} \quad (3.39)$$

where in the second line we used

$$\tilde{f}(0) := \left[\sum_{\lambda=0}^{\infty} f(\lambda) \left(\frac{1}{1+r} \right)^{\lambda} \right] \frac{r}{1+r}.$$

$r := \frac{1}{\beta} - 1$ is the equilibrium interest rate, which is the same in all countries. Using equation (3.39) in (3.38) yields

$$M_{ei}^0 \left[\sum_{j \in \mathcal{J}} \int_{\underline{\varphi}_{ij}^{0,0}}^{\infty} \left[\left(\frac{\varphi}{\underline{\varphi}_{ij}^{0,0}} \right)^{\sigma-1} (\sigma - 1) \tilde{f}(0) + 1 \right] f_{ij} g(\varphi) d\varphi + f_{ei} \right] = L_i. \quad (3.40)$$

Free entry—equation (3.10)—implies

$$\sum_{j \in \mathcal{J}} \int_{\underline{\varphi}_{ij}^{0,0}}^{\infty} \left[\left(\frac{\varphi}{\underline{\varphi}_{ij}^{0,0}} \right)^{\sigma-1} - 1 \right] f_{ij} \tilde{f}(0) g(\varphi) d\varphi = f_{ei}. \quad (3.41)$$

Using the Pareto distribution of productivities in equations (3.40) and (3.41), and rearranging terms yields the expression shown in equation (3.12).

Step 2. From Lemma 2, we know that—in the absence of trade shocks—there will be no firm entry after $t = 0$. To show that this is also the case after a trade shock, we proceed by contradiction.²⁹

Suppose, by way of contradiction, that in response to a trade shock at time $t > 0$ a mass $M_{ei}^t > 0$ of firms entered in country i . Let $\mathcal{M}_i^t$ denote the set of all firms born in i up to and including time t . Firms in this set can uniquely be identified by a triple (i, s, φ) . Let

$$l^t(i, s, \varphi) := f_{ei} + \sum_{j \in \mathcal{J}^t(i, s, \varphi)} \left(\frac{y_{ij}^t(\varphi) \tau_{ij}^t}{\varphi} + f_{ij} \right)$$

denote the total time- t demand for labor of firm (i, s, φ) . Then, for every pair of firms $(i, 0, \varphi), (i, t, \varphi) \in \mathcal{M}_i^t$, i.e., for every pair of firms with equal productivity but born in periods 0 and t , respectively, it holds that $l^t(i, 0, \varphi) \geq l^t(i, t, \varphi)$. In words, the firm born at $t = 0$ demands weakly more labor. This is because (i)

²⁹From Step 1 we know that the mass of entrants at $t = 0$ is independent of the trade environment in the initial period. We note that in our case this is not sufficient to conclude that there will be no entry upon a future trade shock. This is because—due to the tenure effects—the zero-profit cutoff condition is not necessarily binding for the incumbent firm with lowest productivity after a trade shock. We will get back to this point when deriving the general gravity equation for sequential trade liberalizations in Appendix 3.A.D.e.

conditional on serving a market j , $(i, 0, \varphi)$ and (i, t, φ) make the same revenues and, hence, variable profits in market j . (ii) This implies that conditional on serving j , firm (i, t, φ) demands the same amount of labor to serve j as firm $(i, 0, \varphi)$. (iii) For every market j , it holds that $\lambda_j^t(i, 0, \varphi) \geq \lambda_j^t(i, t, \varphi)$ and, hence, $\mathcal{J}^t(i, t, \varphi) \subseteq \mathcal{J}^t(i, 0, \varphi)$. That is, firm (i, t, φ) serves a subset of the markets that firm $(i, 0, \varphi)$ serves, due to (potential) tenure effects.

Now, equations (3.39) and (3.41) hold for the new entrants with their respective cutoffs, i.e.,

$$\begin{aligned} \pi_{ij}^{v,t}(\varphi_{ij}^{0,t}) &= f_{ij} w_i^t \tilde{f}(0) \\ f_{ei} &= \sum_{j \in \mathcal{J}} \int_{\varphi_{ij}^{0,t}}^{\infty} \left[\left(\frac{\varphi}{\varphi_{ij}^{0,t}} \right)^{\sigma-1} - 1 \right] f_{ij} \tilde{f}(0) g(\varphi) d\varphi. \end{aligned}$$

Derivations along the lines of Step 1 then imply that the labor market in country i would clear if (i) we were in the limiting case where $\mathcal{J}^t(i, t, \varphi) = \mathcal{J}^t(i, 0, \varphi)$ and, hence, $l^t(i, t, \varphi) = l^t(i, 0, \varphi)$ for every $(i, 0, \varphi), (i, t, \varphi) \in \mathcal{M}_i^t$, and (ii) if $M_i^t = M_{ei}^0$. This, however, contradicts $M_{ei}^t > 0$. □

3.A.D.d Proof of Proposition 2

From Lemma 3 we know that all firms enter at $t = 0$. Lemma 2 implies that they keep constant over time their market-access strategies. Hence, all firms in i with productivity $\varphi \geq \varphi_{ij}^{0,0}$ serve market j at time t , and total exports of i to j are given by

$$X_{ij}^t = M_{ei}^0 \int_{\varphi_{ij}^{0,0}}^{\infty} \left(\frac{\sigma}{\sigma-1} \right)^{1-\sigma} (\tau_{ij} w_i^t)^{1-\sigma} (P_j^t)^{\sigma-1} X_j^t \varphi^{\sigma-1} g(\varphi) d\varphi.$$

Solving the integral with the Pareto distribution of firm productivities, using equation (3.12) for the mass of entrants, and rearranging terms, we get

$$\frac{X_{ij}^t}{X_j^t} = \frac{\frac{L_i}{f_{ei}} (w_i^t \tau_{ij})^{1-\sigma} (\varphi_{ij}^{0,0})^{\sigma-1-\theta}}{\sum_{k \in \mathcal{J}} \frac{L_k}{f_{ek}} (w_k^t \tau_{kj})^{1-\sigma} (\varphi_{kj}^{0,0})^{\sigma-1-\theta}}. \quad (3.42)$$

Equations (3.7), (3.9), and the fact that $\tilde{f}_{ij}^t(0) = f_{ij} w_i^t \tilde{f}(0)$ imply

$$(\varphi_{ij}^{0,0})^{\sigma-1-\theta} = \left[(w_i^t)^\sigma (\tau_{ij})^{\sigma-1} f_{ij} \tilde{f}(0) \sigma \left(\frac{\sigma}{\sigma-1} \right)^{\sigma-1} (P_j^t)^{1-\sigma} (X_j^t)^{-1} \right]^{\frac{\sigma-1-\theta}{\sigma-1}}.$$

Using this expression in equation (3.42) and simplifying terms yields equation (3.13). By Lemma 2, in the absence of further trade shocks all prices and firms' market access strategies are constant over time and, hence, so are bilateral trade shares. □

3.A.D.e Proof of Proposition 3

We proceed in two steps. We first characterize firms' market access strategies and then use these to derive the gravity equation.

Step 1. By Lemma 3, all firms are born at time $t = 0$. This implies that for each $(i, j) \in \mathcal{J} \times \mathcal{J}$ there is a unique productivity φ_{ij}^t such that all firms in i with productivity $\varphi \geq \varphi_{ij}^t$ serve j at t , while all firms with $\varphi < \varphi_{ij}^t$ do not.³⁰ Let $\lambda_j^t(i, \varphi_{ij}^t) \geq 0$ denote the time- t tenure of the least productive firm in i serving j , where here and below we make use of Lemma 3 and simplify the notation by omitting the entry period s from the firm identifier. Accordingly, we henceforth identify firms by a pair (i, φ) . The cut-off firm must make non-negative profits over its lifetime, i.e.,

$$\pi_{ij}^{v,t}(\varphi_{ij}^t) \geq w_i^t f_{ij} \tilde{f}(\lambda_j^t(i, \varphi_{ij}^t)), \quad (3.43)$$

where we used the fact that by Lemma 2 future profits are constant, and where $\tilde{f}(\lambda) := \left[\sum_{\varsigma=\lambda}^{\infty} f(\varsigma) \left(\frac{1}{1+r} \right)^\varsigma \right] \frac{r}{1+r}$. Due to the tenure effect, condition (3.43) may be strictly non-binding. Let us therefore define a hypothetical tenure

³⁰This is true for an arbitrary sequence of trade liberalizations because at all times firms with a higher productivity earn higher variable profits and, hence, they must have a weakly longer tenure in that market.

level, $\hat{\lambda}_{ij}^t$, at which condition (3.43) holds with equality, i.e.,³¹

$$\pi_{ij}^{v,t}(\varphi_{ij}^t) = w_i^t f_{ij} \tilde{f}(\hat{\lambda}_{ij}^t). \quad (3.44)$$

For all $(i, j, t) \in \mathcal{I} \times \mathcal{J} \times [0, \infty)$, it must hold that $\hat{\lambda}_{ij}^t \in [0, \lambda_j^t(i, \varphi_{ij}^t)]$. This follows by contradiction. In particular, (i) if $\hat{\lambda}_{ij}^t > \lambda_j^t(i, \varphi_{ij}^t)$ it would not be optimal for (i, φ_{ij}^t) to serve j . And (ii) if condition (3.44) would be strictly non-binding for $\hat{\lambda}_{ij}^t = 0$, it would be optimal for some firms with productivity $\varphi < \varphi_{ij}^t$ to start serving j , and (i, φ_{ij}^t) would not be the least productive such firm.

Step 2. The remainder of the proof is analogous to the proof of Proposition 2. In particular, using the productivity cutoff φ_{ij}^t , the Pareto distribution of firm productivities, equation (3.12) for the mass of entrants, and rearranging terms yields

$$\frac{X_{ij}^t}{X_j^t} = \frac{\frac{L_i}{f_{ei}} (w_i^t \tau_{ij}^t)^{1-\sigma} (\varphi_{ij}^t)^{\sigma-1-\theta}}{\sum_{k \in \mathcal{J}} \frac{L_k}{f_{ek}} (w_k^t \tau_{kj}^t)^{1-\sigma} (\varphi_{kj}^t)^{\sigma-1-\theta}}. \quad (3.45)$$

Equations (3.7), and (3.44) imply

$$(\varphi_{ij}^t)^{\sigma-1-\theta} = \left[(w_i^t)^\sigma (\tau_{ij}^t)^{\sigma-1} f_{ij} \tilde{f}(\hat{\lambda}_{ij}^t) \sigma \left(\frac{\sigma}{\sigma-1} \right)^{\sigma-1} (P_j^t)^{1-\sigma} (X_j^t)^{-1} \right]^{\frac{\sigma-1-\theta}{\sigma-1}}.$$

Using this expression in equation (3.45) and simplifying terms yields equation (3.15). Lastly, Lemma 2 implies again that—in the absence of further shocks—bilateral trade shares are constant over time. □

³¹If condition (3.43) holds with equality for more than one λ because $\tilde{f}(\cdot)$ is constant for these λ , we take $\hat{\lambda}_{ij}^t$ to be the smallest λ for which condition (3.43) holds with equality.

3.A.E Additional Figures

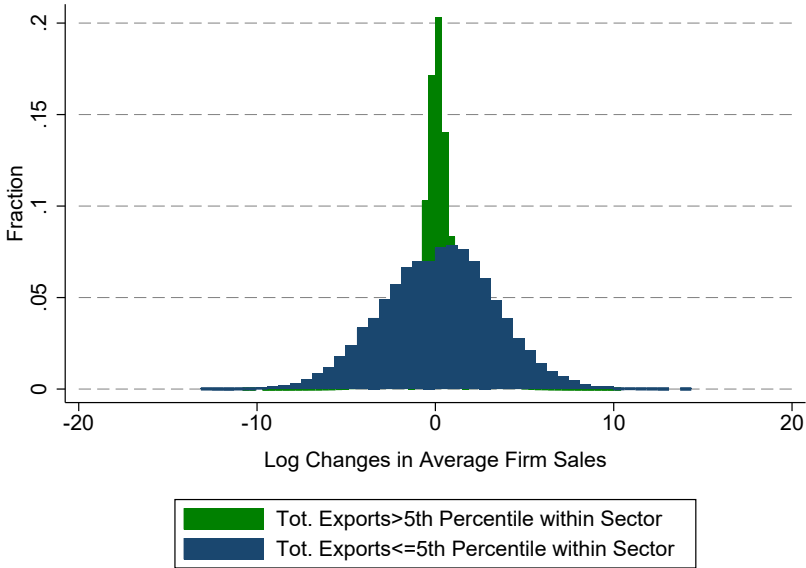


Figure 3.A.E.1: Log-Changes in Average Firm Sales – Very Small Country Pairs vs. Rest

Notes. This figure depicts log-changes in average firm sales from year t to $t+2$ for very small country pairs (i.e., those with total exports belonging to the bottom 5% within a sector) vs. the remaining relations. The sample spans the years 1997–2014. Sectors are defined as the HS2-digit level. Details are given in Section 3.6.

Data source. EDD.

Graph. Authors' representation.

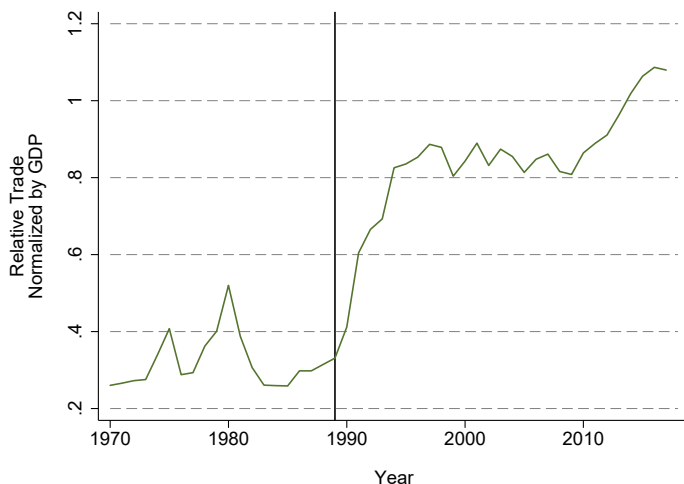


Figure 3.A.E.2: Average Trade Flows between OECD and Iron Curtain Countries relative to Average Intra-OECD Trade

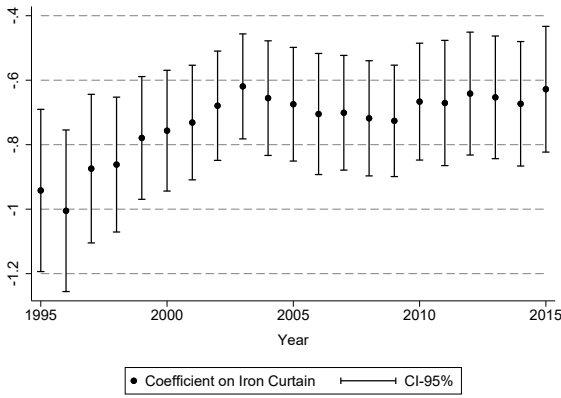
Notes. This figure shows how average trade flows between OECD and Iron Curtain countries evolved over time relative to average within-OECD flows. Trade flows are normalized by the product of importer and exporter GDPs. Iron Curtain countries are listed in footnote 16. Details on the sample selection for this graph are given in Appendix 3.A.B. The vertical line is drawn at 1989.

Data source. [Atlas of Economic Complexity](#) (trade data), [UNdata](#) (GDP data).

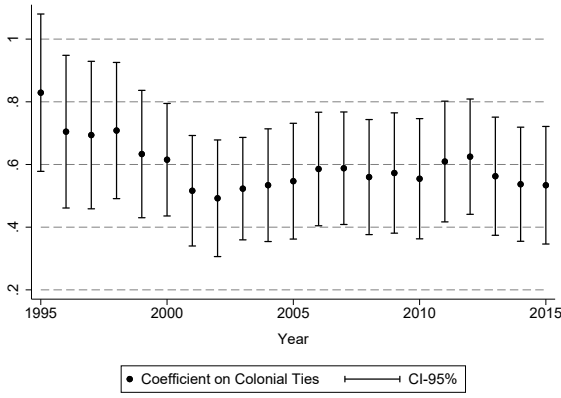
Graph. Authors' representation.

Figure 3.A.E.3: Main Coefficients in Table 3.1 over Time

(a) $\mathbb{1}[\text{Sep. by Iron Curtain}]$



(b) $\mathbb{1}[\text{Colonial Ties}]$



Notes. This figure reproduces the main coefficients in Column (2) of Table 3.1 for 1995 to 2015. Further details are provided in the notes to Table 3.1.
Data source. [Atlas of Economic Complexity](#) (trade data), [CEPII/Head et al. \(2010\)](#) (co-variates).
Graph. Authors' representation.

3.A.F Additional Tables

Table 3.A.F.1: Impact of Firms' Tenure Levels on Market Exit Probabilities – Venezuela Shock

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tenure=2	-0.105*** (0.020)	-0.096*** (0.019)	-0.092*** (0.020)	-0.088*** (0.021)	-0.053*** (0.016)	-0.055*** (0.015)	-0.051*** (0.015)
Tenure=3	-0.232*** (0.026)	-0.210*** (0.028)	-0.211*** (0.027)	-0.211*** (0.027)	-0.164*** (0.031)	-0.180*** (0.030)	-0.172*** (0.031)
Tenure=4	-0.342*** (0.041)	-0.304*** (0.042)	-0.303*** (0.041)	-0.307*** (0.042)	-0.260*** (0.038)	-0.276*** (0.040)	-0.283*** (0.044)
Tenure=5	-0.429*** (0.037)	-0.388*** (0.037)	-0.385*** (0.037)	-0.384*** (0.038)	-0.281*** (0.037)	-0.291*** (0.038)	-0.289*** (0.043)
Tenure=6	-0.515*** (0.044)	-0.481*** (0.046)	-0.478*** (0.045)	-0.487*** (0.045)	-0.288*** (0.056)	-0.321*** (0.053)	-0.339*** (0.050)
Tenure=7	-0.519*** (0.060)	-0.461*** (0.060)	-0.459*** (0.059)	-0.468*** (0.057)	-0.235*** (0.058)	-0.253*** (0.061)	-0.242*** (0.064)
Tenure=8	-0.447*** (0.056)	-0.388*** (0.058)	-0.382*** (0.061)	-0.386*** (0.061)	-0.457*** (0.052)	-0.478*** (0.060)	-0.470*** (0.063)
Tenure=9	-0.545*** (0.050)	-0.476*** (0.052)	-0.451*** (0.053)	-0.443*** (0.057)	-0.419*** (0.058)	-0.446*** (0.057)	-0.424*** (0.065)
Tenure=10	-0.661*** (0.027)	-0.583*** (0.032)	-0.577*** (0.035)	-0.569*** (0.035)	-0.419*** (0.038)	-0.440*** (0.036)	-0.428*** (0.039)
Mean of Dep. Var.	0.715	0.714	0.713	0.713	0.719	0.718	0.718
Nr. of Observations	6,382	6,375	6,286	6,286	4,826	4,732	4,732
Adj. R-squared	0.212	0.243	0.251	0.255	0.560	0.556	0.566
Size by p	No	Yes	No	No	Yes	No	No
Size by pq	No	No	No	Yes	No	No	Yes
FEs	-	p	pq	pq	p, f	pq, f	pq, f

Notes. This table summarizes results from estimating equation (3.1) using OLS. Standard errors clustered at the product and firm levels are reported in parentheses. “size by p” (“size by pq”) indicates whether the regression allows for heterogeneous effects of log sales at the product (product-sales quintile) level. *p* indicates product fixed effects, *pq* product-sales quintile fixed effects, where the quintiles are computed at the product level. *f* denotes firm fixed effects. All tenure levels of 10 years and above have been summarized in one category “tenure ≥ 10 .” * significant at 10%-level; ** significant at 5%-level; *** significant at 1%-level.

Data source. ADUANAS (DIAN)

Results. Authors' computations.

Table 3.A.F.2: Impact of Firms' Tenure Levels on Market Exit Probabilities – Venezuela Shock – Robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tenure=2	-0.111*** (0.023)	-0.099*** (0.022)	-0.103*** (0.022)	-0.102*** (0.022)	-0.052** (0.020)	-0.069*** (0.020)	-0.053** (0.020)
Tenure=3	-0.281*** (0.026)	-0.237*** (0.027)	-0.239*** (0.026)	-0.228*** (0.025)	-0.165*** (0.030)	-0.185*** (0.030)	-0.168*** (0.032)
Tenure=4	-0.379*** (0.039)	-0.339*** (0.038)	-0.355*** (0.038)	-0.349*** (0.038)	-0.271*** (0.037)	-0.313*** (0.038)	-0.289*** (0.040)
Tenure=5	-0.468*** (0.028)	-0.422*** (0.028)	-0.418*** (0.029)	-0.413*** (0.030)	-0.291*** (0.039)	-0.307*** (0.040)	-0.291*** (0.040)
Tenure=6	-0.568*** (0.044)	-0.516*** (0.047)	-0.515*** (0.047)	-0.525*** (0.048)	-0.335*** (0.044)	-0.365*** (0.047)	-0.347*** (0.047)
Tenure=7	-0.554*** (0.048)	-0.489*** (0.046)	-0.494*** (0.047)	-0.492*** (0.048)	-0.327*** (0.055)	-0.352*** (0.063)	-0.319*** (0.071)
Tenure=8	-0.512*** (0.052)	-0.436*** (0.055)	-0.439*** (0.056)	-0.421*** (0.057)	-0.415*** (0.060)	-0.431*** (0.063)	-0.379*** (0.062)
Tenure=9	-0.599*** (0.043)	-0.512*** (0.044)	-0.510*** (0.043)	-0.501*** (0.047)	-0.398*** (0.067)	-0.438*** (0.068)	-0.431*** (0.071)
Tenure=10	-0.668*** (0.021)	-0.574*** (0.025)	-0.569*** (0.026)	-0.558*** (0.027)	-0.379*** (0.038)	-0.420*** (0.038)	-0.390*** (0.043)
Mean of Dep. Var.	0.629	0.629	0.627	0.627	0.634	0.633	0.633
Nr. of Observations	6,377	6,371	6,281	6,281	4,804	4,705	4,705
Adj. R-squared	0.214	0.247	0.254	0.260	0.553	0.551	0.562
Size by p	No	Yes	No	No	Yes	No	No
Size by pq	No	No	No	Yes	No	No	Yes
FES	-	p	pq	pq	p, f	pq, f	pq, f

Notes. This table reproduces Table 3.A.F.1 changing the coding of the dependent variable. The dependent variable is equal to one if a firm exported a product to Venezuela in the period August 2008 to July 2009 but not in the period from August 2009 to July 2010, and equal to zero if the firm had positive exports in both of these periods. Consult the notes to Table 3.A.F.1 for further details.

Table 3.A.F.3: Impact of Firms' Tenure Levels on Market Exit Probabilities

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tenure=2	-0.229*** (0.011)	-0.161*** (0.012)	-0.158*** (0.012)	-0.154*** (0.013)	-0.099*** (0.008)	-0.098*** (0.008)	-0.091*** (0.010)
Tenure=3	-0.346*** (0.014)	-0.245*** (0.017)	-0.245*** (0.018)	-0.239*** (0.021)	-0.154*** (0.011)	-0.155*** (0.012)	-0.145*** (0.014)
Tenure=4	-0.407*** (0.014)	-0.289*** (0.018)	-0.287*** (0.019)	-0.278*** (0.022)	-0.182*** (0.012)	-0.182*** (0.013)	-0.168*** (0.015)
Tenure=5	-0.450*** (0.016)	-0.315*** (0.020)	-0.311*** (0.021)	-0.302*** (0.024)	-0.201*** (0.013)	-0.202*** (0.014)	-0.188*** (0.017)
Tenure=6	-0.482*** (0.017)	-0.339*** (0.020)	-0.334*** (0.021)	-0.326*** (0.024)	-0.212*** (0.014)	-0.216*** (0.014)	-0.198*** (0.017)
Tenure=7	-0.494*** (0.016)	-0.341*** (0.022)	-0.337*** (0.023)	-0.326*** (0.027)	-0.214*** (0.016)	-0.216*** (0.016)	-0.198*** (0.019)
Tenure=8	-0.519*** (0.015)	-0.356*** (0.022)	-0.354*** (0.023)	-0.341*** (0.027)	-0.213*** (0.016)	-0.217*** (0.016)	-0.193*** (0.018)
Tenure=9	-0.528*** (0.015)	-0.363*** (0.022)	-0.356*** (0.023)	-0.345*** (0.026)	-0.218*** (0.015)	-0.222*** (0.015)	-0.201*** (0.019)
Tenure=10	-0.550*** (0.015)	-0.366*** (0.023)	-0.361*** (0.024)	-0.348*** (0.028)	-0.216*** (0.016)	-0.227*** (0.016)	-0.207*** (0.020)
Mean of Dep. Var.	0.475	0.470	0.464	0.464	0.376	0.360	0.360
Nr. of Observations	644,289	619,943	543,695	543,690	413,359	337,422	337,422
Adj. R-squared	0.160	0.231	0.233	0.236	0.386	0.387	0.389
Size by jpy	No	Yes	No	No	Yes	No	No
Size by jpyq	No	No	No	Yes	No	No	Yes
FEs	-	jpy	jpyq	jpyq	jpy, fpy	jpyq, fpy	jpyq, fpy

Notes. This table shows results analogous to Table 3.A.F.1, but using data for all destination countries and years 1994 to 2013. Standard errors clustered at the destination, product, and firm levels are reported in parantheses. “size by jpy” (“size by jpyq”) indicates whether the regression allows for heterogeneous size effects at the destination-product-year (destination-product-year-sales quintile) level. *cpy* indicates country-product-year fixed effects, *cpyq* country-product-year-sales quintile fixed effects, where the quintiles are computed at the country-product-year level. *f* (*fpy*) denotes firm(-product-year) fixed effects. Consult the notes to Table 3.A.F.1 for further details.

Table 3.A.F.4: Gravity Regressions, Year 2015, OLS

	Main Specification		Robustness		
			+Tariffs & NTMs	+CIF/FOB ratio	Comext Data
Sep. by Iron Curtain	-0.572*** (0.096)	-0.628*** (0.099)	-0.572*** (0.090)	-0.553*** (0.087)	-0.930*** (0.142)
Colonial Ties	0.781*** (0.095)	0.534*** (0.096)	0.507*** (0.096)	0.497*** (0.098)	0.935*** (0.240)
Log of Distance	-1.341*** (0.044)	-1.210*** (0.054)	-1.162*** (0.054)	-1.173*** (0.056)	-1.658*** (0.252)
Contiguity	0.849*** (0.135)	0.621*** (0.131)	0.333*** (0.126)	0.505*** (0.144)	-0.802*** (0.209)
Common Language	0.409*** (0.080)	0.409*** (0.080)	0.306*** (0.077)	0.302*** (0.079)	0.447*** (0.165)
Common Currency	-0.078 (0.167)	-0.078 (0.167)	-0.063 (0.162)	-0.160 (0.153)	
RTA	0.406*** (0.076)	0.406*** (0.076)	0.166* (0.087)	0.175* (0.089)	
Common Legal Origin	-0.012 (0.045)	-0.012 (0.045)	0.018 (0.044)	0.008 (0.043)	0.335*** (0.114)
Abs. Log-Diff. in GDP pc	-0.128*** (0.033)	-0.128*** (0.033)	-0.119*** (0.034)	-0.112*** (0.034)	-0.101 (0.085)
Average Tariffs			-7.405*** (1.235)	-7.862*** (1.402)	
CIF/FOB Ratio				5.454*** (1.800)	
Mean of Dep. Var.	2.469	2.682	2.742	2.918	2.024
Adj. R-squared	0.737	0.765	0.782	0.796	0.799
Nr. of Observations	19,847	17,153	16,195	14,511	6,849

Notes. This table reproduces Table 3.1, showing coefficients for all controls but for the 29 NTM controls. The coefficients on NTMs in column (3) are provided in Table 3.A.F.5. Consult the notes to Table 3.1 for details.

Table 3.A.F.5: Gravity Regressions, OLS, Detailed NTMs

	+ Tariffs & NTMs	
Sep. by Iron Curtain	-0.572***	(0.090)
Colonial Ties	0.507***	(0.096)
Log of Distance	-1.162***	(0.054)
Contiguity	0.333***	(0.126)
Common Language	0.306***	(0.077)
Common Currency	-0.063	(0.162)
RTA	0.166*	(0.087)
Common Legal System	0.018	(0.044)
Abs. Log-Diff. in GDP pc	-0.119***	(0.034)
Simple Average of Tariffs	-7.405***	(1.235)
NTM: Sanitary and phytosanitary measure	-0.265***	(0.090)
NTM: Technical barriers to trade	0.282***	(0.106)
NTM: Capital control measures	0.000	(0.000)
NTM: Antidumping	-0.148***	(0.057)
NTM: Countervailing measure	-0.283**	(0.137)
NTM: General (multilateral) safeguard	-0.106	(0.089)
NTM: Non-automatic import licensing procedures	0.114**	(0.054)
NTM: Quotas	0.105	(0.074)
NTM: Prohibitions other than for SPS and TBT reasons	0.046	(0.058)
NTM: Tariff-rate quotas (TRQ)	-0.076	(0.056)
NTM: Internal taxes and charges levied on imports	-0.035	(0.071)
NTM: FDI measures	0.000	(0.000)
NTM: Finance measures	0.119	(0.086)
NTM: Local content measures	0.015	(0.094)
NTM: Trade-balancing measures	-0.107	(0.135)
NTM: Instrument unclear	0.017	(0.061)
NTM: Subsidies (excl. export subsidies)	0.227***	(0.049)
NTM: Government procurement market access restrictions	0.081	(0.107)
NTM: Government procurement domestic price preference	0.107	(0.106)
NTM: Government procurement local content requirement	-0.128**	(0.059)
NTM: Migration measures	0.000	(0.000)
NTM: Intellectual property	0.000	(0.000)
NTM: Export prohibition	-0.042	(0.102)
NTM: Export quotas	-0.116	(0.088)
NTM: Licensing- or permit requirements to export	0.184***	(0.052)
NTM: Export taxes and charges	0.179**	(0.078)
NTM: Export subsidies	0.013	(0.042)
NTM: Export credits	-0.112*	(0.062)
NTM: Export measures, n.e.s.	0.136	(0.086)
Mean of Dep. Var.	2.742	
Adj. R-squared	0.782	
Nr. of Observations	16,195	

Notes. This table reproduces the third column of Table 3.1 and reports the coefficients on the non-tariff measures. Consult the notes to Table 3.1 for details.

Table 3.A.F.6: Gravity Regressions, Year 2015, PPML

Main Specification		Robustness		
	+Tariffs & NTMs	+CIF/FOB ratio	Conext Data	Air or Sea Transport
Sep. by Iron Curtain	-0.469*** (0.107)	-0.477*** (0.088)	-0.433*** (0.089)	-0.427*** (0.089)
Colonial Ties	0.275 (0.205)	0.319** (0.156)	0.201 (0.145)	0.380*** (0.131)
Log of Distance	-0.854*** (0.065)	-0.743*** (0.060)	-0.719*** (0.056)	-1.508*** (0.223)
Contiguity	0.475*** (0.108)	0.400*** (0.084)	0.284*** (0.078)	0.329 (0.248)
Common Language	-0.115 (0.080)	-0.019 (0.082)	-0.026 (0.083)	-0.072 (0.184)
Common Currency	-0.093 (0.150)	-0.023 (0.139)	-0.022 (0.139)	-0.119 (0.216)
RTA	0.485*** (0.058)	0.358*** (0.099)	0.366*** (0.100)	
Common Legal Origin	0.167*** (0.067)	0.213*** (0.055)	0.216*** (0.055)	0.272*** (0.048)
Abs. Log-Diff. in GDP pc	-0.005 (0.034)	-0.018 (0.032)	-0.016 (0.032)	-0.200** (0.085)
Average Tariffs		-1.911 (1.404)	-1.661 (1.425)	-0.244** (0.104)
CIF/FOB Ratio			-0.280 (1.516)	
Mean of Dep. Var.	348	505	787	677
Nr. of Observations	43,890	29,756	26,059	7,644

Notes. This table reproduces Table 3.A.F.4 using PPML instead of OLS. Dependent variable is absolute trade flows. Consult the notes to Table 3.1 for details.

Table 3.A.F.7: Impact of Tariff Changes on Average Firm Sales – Weighted Average of Tariffs

Dependent Var.: Log-Change in Average Firm Sales from t to $t+2$				
Control Group: No Tariff Cuts from $t-1$ to $t+2$				
	Main Specification		Robustness	
	90% Wins.		Pair-Year FE	
1 [Top Quartile of Tariff Cuts from t to $t+1$]	0.169*** (0.016)	0.168*** (0.017)	0.153*** (0.037)	0.104*** (0.027)
Sector-Year FE	✓			
Exp-Sector-Year FE		✓	✓	✓
Imp-Sector-Year FE		✓	✓	✓
Mean of Dep. Var.	0.096	0.094	0.089	0.087
Adj. R-squared	0.027	0.100	0.108	0.115
Nr. of Observations	137,315	132,279	93,431	93,431
Dependent Var.: Log-Change in Average Firm Sales from t to $t+\bar{t}$				
Control Group: No Tariff Cuts from $t-\bar{t}$ to $t+\bar{t}$				
	Robustness			
$\bar{t}=2, \underline{t}=0$	$\bar{t}=2, \underline{t}=2$	$\bar{t}=3, \underline{t}=1$	$\bar{t}=3, \underline{t}=2$	$\bar{t}=3, \underline{t}=3$
1 [Top Quartile of Tariff Cuts from t to $t+1$]	0.166*** (0.029)	0.100** (0.045)	0.215*** (0.056)	0.226*** (0.070)
Exp-Sector-Year FE	✓	✓	✓	✓
Imp-Sector-Year FE	✓	✓	✓	✓
Mean of Dep. Var.	0.099	0.081	0.112	0.095
Adj. R-squared	0.105	0.106	0.119	0.117
Nr. of Observations	125,042	75,909	65,712	55,208
				48,157

Notes. This table reproduces Table 3.2 but using the weighted average of tariffs within a sector instead of the simple average. The weights are the bilateral trade shares for a given product within a sector. Consult the notes to Table 3.2 for further details.

Table 3.A.F.8: Impact of Tariff Changes on Average Firm Sales – Robustness

	Dependent Var.: Log-Change in Average Firm Sales from t to $t+2$			
	Control Group: No Tariff Cuts from $t-1$ to $t+2$		+ Lagged Tariffs + 1[Preferential Terms]	
	Above-Median Tariff Cuts	No Rebound of Tariff Cuts	Up to 50% Rebound of Tariff Cuts	
1[Top Tariff Cuts from t to $t+1$]	0.075*** (0.022)	0.061* (0.035)	0.077** (0.033)	0.085** (0.035)
Exp-Sector-Year FE	✓	✓	✓	✓
Imp-Sector-Year FE	✓	✓	✓	✓
Mean of Dep. Var.	0.097	0.091	0.092	0.091
Adj. R-squared	0.090	0.090	0.091	0.091
Nr. of Observations	164,706	150,844	151,817	151,694
	No Minimum Tot. Exports	Drop Bottom 10% within Sector	Corr. w/ Aggr. Data Unchecked	Corr. w/ Aggr. Data ≥ 0.8
1[Top Tariff Cuts from t to $t+1$]	0.096*** (0.036)	0.078** (0.032)	0.067** (0.031)	0.077** (0.037)
Exp-Sector-Year FE	✓	✓	✓	✓
Imp-Sector-Year FE	✓	✓	✓	✓
Mean of Dep. Var.	0.096	0.090	0.076	0.097
Adj. R-squared	0.079	0.093	0.076	0.101
Nr. of Observations	154,389	145,807	198,842	117,183

Notes. This table reproduces the specification in the third column of the top panel in Table 3.2 for different definitions of the “treatment” and for variations of the data cleaning steps. Consult the notes to Table 3.2 for details.

Top panel. The first column allocates observations with above-median tariff cuts to the treatment group. The second and third ones allow for no or up to a 50% rebound of the tariff cut between $t+1$ and $t+2$, respectively. The fourth column controls for the initial tariff level and for whether the tariffs are preferential.

Bottom panel. The first (second) column sets the minimum total exports a country pair should have to zero (the 10th percentile within a sector). The third (fourth) column sets the correlation cutoff between total exports of an exporter-importer-sector pair in the aggregate data and the firm-level statistics to -1 (0.8). * significant at 10%-level; ** significant at 5%-level; *** significant at 1%-level.

Table 3.A.F.9: Impact of EU Eastern Enlargement on Average Firm Sales

	Dependent Variable: Log-Change in Average Firm Sales from 2002 to $\bar{t}$							
	$\bar{t}=2005$	$\bar{t}=2005$	$\bar{t}=2006$	$\bar{t}=2006$	$\bar{t}=2007$	$\bar{t}=2007$	$\bar{t}=2008$	$\bar{t}=2008$
EU Pair	0.059* (0.031)	-0.024 (0.083)	0.085 (0.073)	0.026 (0.108)	0.128*** (0.017)	0.042 (0.084)	0.089*** (0.026)	-0.008 (0.090)
New EU Pair	0.204*** (0.031)	0.256*** (0.079)	0.283*** (0.039)	0.287*** (0.070)	0.294*** (0.061)	0.269** (0.122)	0.330*** (0.096)	0.358*** (0.114)
Exporter FE	✓	✓	✓	✓	✓	✓	✓	✓
Importer FE		✓		✓		✓		✓
Mean of Dep. Var.	0.414	0.414	0.556	0.556	0.667	0.668	0.830	0.831
Adj. R-squared	0.044	0.090	0.041	0.097	0.047	0.093	0.027	0.089
Nr. of Observations	2,470	2,464	2,437	2,433	2,317	2,312	2,332	2,329

Notes. This table reports OLS results from regressing log-changes in average firm sales on indicators for EU pairs in general and EU pairs that were newly created in the EU enlargement of 2004, as described in Appendix 3.A.B. The regression is outlined in Equation (3.21). Standard errors clustered at the importer and exporter levels are reported in parentheses. “EU Pair” is a dummy variable equal to one if both the exporter and the importer are EU members, and zero otherwise. “New EU Pair” is a dummy variable equal to one if both the exporter and the importer are EU members and at least one of them entered the EU in 2004, and zero otherwise. The sample excludes Bulgaria and Romania as exporters and importers.

* significant at 10%-level;** significant at 5%-level;*** significant at 1%-level.

Data source. TEC and EDD.

Results. Authors’ computations.

Table 3.A.F.10: Impact of EU Eastern Enlargement on Average Firm Sales – incl. BGR & ROU

	Dependent Variable: Log-Change in Average Firm Sales from 2002 to $\bar{t}$							
	$\bar{t}=2005$	$\bar{t}=2005$	$\bar{t}=2006$	$\bar{t}=2006$	$\bar{t}=2007$	$\bar{t}=2007$	$\bar{t}=2008$	$\bar{t}=2008$
EU Pair	0.086* (0.043)	0.056 (0.079)	0.071 (0.081)	0.024 (0.101)	0.126*** (0.017)	0.041 (0.083)	0.090*** (0.028)	0.000 (0.087)
New EU Pair	0.151*** (0.008)	0.100 (0.071)	0.318*** (0.052)	0.282*** (0.087)	0.314*** (0.054)	0.239** (0.111)	0.341*** (0.081)	0.318*** (0.104)
Exporter FE	✓	✓	✓	✓	✓	✓	✓	✓
Importer FE		✓		✓		✓		✓
Mean of Dep. Var.	0.417	0.417	0.551	0.551	0.672	0.673	0.841	0.842
Adj. R-squared	0.040	0.081	0.041	0.093	0.047	0.095	0.031	0.094
Nr. of Observations	2,613	2,607	2,581	2,577	2,348	2,343	2,398	2,395

Notes. This table reports OLS results from regressing log-changes in average firm sales on indicators for EU pairs in general and EU pairs that were newly created in the EU enlargements of 2004 and 2007, as described in Appendix 3.A.B. The regression is outlined in Equation (3.21). Standard errors clustered at the importer and exporter levels are reported in parentheses. “EU Pair” is a dummy variable equal to one if both the exporter and the importer are EU members by the year 2007, and zero otherwise. “New EU Pair” is a dummy variable equal to one if both the exporter and the importer are EU members by the year 2007 and at least one of them entered the EU after 2003, and zero otherwise.

* significant at 10%-level,** significant at 5%-level,*** significant at 1%-level.

Data source. TEC and EDD.

Results. Authors’ computations.

Table 3.A.F.11: Impact of EU Eastern Enlargement on Average Firm Sales – Base Year 2001

	Dependent Variable: Log-Change in Average Firm Sales from 2001 to t							
	$\bar{t}=2005$	$\bar{t}=2005$	$\bar{t}=2006$	$\bar{t}=2006$	$\bar{t}=2007$	$\bar{t}=2007$	$\bar{t}=2008$	$\bar{t}=2008$
EU Pair	0.086 (0.079)	0.091 (0.137)	0.113 (0.122)	0.119 (0.199)	0.190*** (0.040)	0.137 (0.140)	0.149*** (0.022)	0.114 (0.099)
New EU Pair	0.216*** (0.011)	0.323*** (0.109)	0.315*** (0.050)	0.412*** (0.144)	0.278*** (0.047)	0.397*** (0.120)	0.342*** (0.057)	0.504*** (0.115)
Exporter FE	✓	✓	✓	✓	✓	✓	✓	✓
Importer FE		✓		✓		✓		✓
Mean of Dep. Var.	0.433	0.432	0.567	0.568	0.645	0.646	0.804	0.804
Adj. R-squared	0.076	0.144	0.061	0.146	0.045	0.108	0.032	0.103
Nr. of Observations	1,769	1,762	1,740	1,734	1,628	1,622	1,642	1,639

Notes. This table reproduces Table 3.A.F.9 changing the base period from 2002 to 2001. Consult the notes to Table 3.A.F.9 for details.

Table 3.A.F.12: Impact of EU Eastern Enlargement on Average Firm Sales – Base Year 2003

	Dependent Variable: Log-Change in Average Firm Sales from 2003 to $\bar{t}$							
	$\bar{t}=2005$	$\bar{t}=2005$	$\bar{t}=2006$	$\bar{t}=2006$	$\bar{t}=2007$	$\bar{t}=2007$	$\bar{t}=2008$	$\bar{t}=2008$
EU Pair	0.135*** (0.039)	0.037 (0.061)	0.162* (0.083)	0.052 (0.099)	0.193*** (0.059)	0.054 (0.095)	0.157* (0.080)	0.063 (0.116)
New EU Pair	0.098 (0.098)	0.106 (0.105)	0.185** (0.088)	0.213* (0.122)	0.202* (0.110)	0.241* (0.136)	0.297*** (0.098)	0.380*** (0.119)
Exporter FE	✓	✓	✓	✓	✓	✓	✓	✓
Importer FE		✓		✓		✓		✓
Mean of Dep. Var.	0.280	0.280	0.430	0.432	0.563	0.565	0.705	0.708
Adj. R-squared	0.016	0.038	0.022	0.065	0.027	0.057	0.020	0.069
Nr. of Observations	2,833	2,822	2,760	2,751	2,638	2,628	2,684	2,675

Notes. This table reproduces Table 3.A.F.9 changing the base period from 2002 to 2003. Consult the notes to Table 3.A.F.9 for details.

Chapter 4

Disappearing Products and Jumbo Clusters: Addressing the HS Vintage Conversion Problem

Joint with Piotr Lukaszuk

4.1 Introduction

Empirical international trade research has expanded from the aggregate analysis of trade flows to more granular data at the sector, product, firm, and even transaction level. Given that trade figures by default include flows between two different jurisdictions, these granular data require an internationally har-

monized classification of the shipped products. Since 1988, the World Customs Organization (WCO) has developed the Harmonized Commodity Description and Coding System (HS) in order to globally standardize the classification of products into over 5,000 6-digit HS codes.¹

Although the harmonization in 1988 created a standard allowing for cross-country comparisons of trade figures going beyond trade flows to, e.g., tariff lines, it was unable to avoid necessary product code adjustments over time. Hence, on average every five years the WCO updates its list of HS product codes and publishes a new vintage.² The amendments have been necessary, *inter alia*, due to technological progress (creating new products and making some void), shifting trade patterns (increasing or declining relevance of specific products), adaptation to reflect trade practice (unclear product classifications or bundling of product parts by firms), as well as amendments related to social and environmental fields (e.g., to improve the tracking of endangered plants and animals).³

While these adjustments are meant to facilitate trade, they often result in products being split up (“1:n”), merged (“m:1”), or generally reassigned (“m:n”). This, in turn, complicates a comparison of trade flows both across time and across countries. The former comparison is difficult because it is unclear how to convert trade flows from one HS classification to another whenever a link is 1:n or m:n, while the fact that some jurisdictions adopt the new HS code version (“vintage”) with a delay hinders cross-sectional analyses.⁴ Below, we revisit commonly used conversion methods and develop an alternative to tackle several issues that arise when applying these existing techniques.

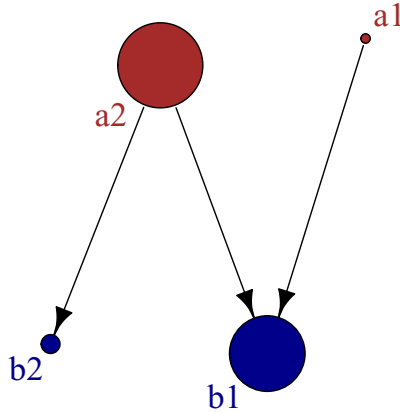
Given that some HS code adjustments are interlinked, they may not be

¹Each country may determine and adjust its HS codes at a more granular level. For instance, the United States classifies its products at the 10-digit level in its so-called HTS classification. However, the HS codes remain globally harmonized at the 6-digit level. Before the introduction of HS codes, goods were largely classified under the so-called SITC system.

²More precisely, revised HS vintages were introduced in 1992, 1996, 2002, 2007, 2012, and 2017. We will refer to a vintage introduced in a specific year as, e.g., HS 2007.

³For further details see Appendix 4.A.A.

⁴For instance, according to article 4(2) of the HS Convention of the WCO, developing countries—representing three-quarters of the WCO’s 183 members—may adopt the newest vintage with a delay of up to five years.

Figure 4.1: Network Visualization of an Exemplary HS Code Group

Notes: This figure illustrates an exemplary HS code group in the form of a network diagram. In this example, code $a1$ was reassigned to code $b1$ (therefore receiving a conversion weight equal to one), while code $a2$ was reassigned to codes $b1$ and $b2$, which means that we do not know a priori how much trade should be allocated to which new code. Larger circles depict larger product trade shares. We refer to such networks of linked products as product groups.

viewed as individual changes but rather as a connected network or *group* of adjusted HS codes. Put differently, whenever two codes of an “initial” vintage are assigned to the same code of a “target” vintage, they are (indirectly) linked, and will thus belong to the same *group*. Figure 4.1 visualizes an exemplary group of HS codes in the form of a network chart, where bigger circles depict larger trade shares within this group. In this example, the “initial” codes $a1$ and $a2$ are reassigned to “target” codes $b1$ and $b2$. Therefore, we want to know how we can convert trade flows reported in the a -vintage into the b -vintage. Since product $a1$ has only a link to product $b1$, all trade flows reported in product category $a1$ will be added to product $b1$ when converting to this new HS vintage. By contrast, because code $a2$ has a link with both $b1$ and $b2$, it is a priori unclear which part of trade flows classified in the initial vintage should be allocated to $b1$ or to $b2$. We will refer to the share of $a2$ flows being

allocated to $b1$ as the *conversion weight* from $a2$ to $b1$, given that overall trade in $b1$ will be a weighted sum of trade in the linked initial products (here, $a1$ and $a2$). The conversion weight from $a1$ to $b1$ is equal to 1, and, accordingly, equal to 0 for $a1$ to $b2$, while the conversion weights from $a2$ to $b1$ and $b2$ are unknown.

While in each amendment most HS codes remain unchanged, the non-1:1 adjustments accumulate when the conversion is performed over multiple vintages. As indicated in Table 4.1, the majority of HS codes remain the same and have a 1:1 assignment when converting the HS 2017 codes into HS 2012. However, the further one extends the conversion back to earlier vintages, the larger the share of conversions with undefined weights becomes. When trying to convert HS 2017 codes into HS 1992 codes, the share of products with unknown conversion weights increases to about 16%.⁵ More generally, out of 6,539 HS codes applied between 1992 and 2017, 48% have faced a non-1:1 adjustment at least in one of the classification updates. In terms of 2018 trade, these codes accounted for 43.7% of total trade.

Widely-used methods address the HS code conversion issue by either forcing a one-to-one mapping between product codes across vintages (UNSD, 2017), and thus ignoring remaining product links, or by pooling adjusted codes into synthetic codes (Pierce and Schott, 2012). The former method leads to intertemporal inconsistencies in HS vintage amendment years. For instance, when converting HS 2007 codes into the HS 2002 vintage, the number of non-traded products increases from 0 in the year 2006 to 299 in the year 2007, which represents 42% of codes where the assignment is not pre-determined. Meanwhile, the Pierce and Schott (2012) method increases the share of the largest product code in overall trade of the year 2018 from about 7% to 26%, when converting the HS 2017 codes into HS 1992 codes. Not only do we then obtain, by construction, more aggregated data, the distribution also appears to become more skewed due to the fact that unchanged codes are never lumped into these synthetic clusters. Researchers may circumvent these issues either by using only data at coarser levels, as in gravity regressions at the country

⁵Note that one may instead want to convert HS 1992 to HS 2017 codes. In that case, those HS codes summarized in Table 4.1 as “m:1” will have unknown conversion weights.

Table 4.1: Distribution of Product Links between HS 2017 and Other Vintages

Relationship with HS 2017	HS 2012	HS 2002	HS 1992	Conversion Weight
1:1	4,744	3,966	3,332	1
$m:1$	578	948	1332	1
$1:n$	36	262	360	?
$m:n$	60	316	538	?

Notes: This table depicts how many product links between HS 2017 codes and other vintages are of a specific type. As illustrated by the question marks, whenever a link is “1:n” (i.e., the code has to be split into multiple codes of the other reporting standard) or “m:n” (i.e., m HS 2017 codes are interlinked with n codes of the other HS vintage), the correct conversion is unknown.

Data source: [UN Comtrade Database](#).

level (cf. [Head and Mayer \(2014\)](#) for an overview), or by focusing solely on products with a 1:1 mapping (e.g., [Boehm et al., 2020](#)). While the former approach eliminates interesting variation, the latter one substantially reduces the sample size for longer time horizons.⁶

In light of the available methods, this paper aims to contribute to the existing literature by proposing an alternative conversion technique between different vintages of HS product codes. The key feature of trade data we exploit is their persistence over time. Specifically, we provide an algorithm that through the choice of conversion weights minimizes the squared deviations between pre- and post-classification-change country-pair data. Our minimization imposes a set of constraints which respects the product links. The algorithm solves the aforementioned issues of existing methods, and the calculated weights are robust to altering the set of years and country-pair observations.

⁶In the paper we exclusively present results based on converting backwards (i.e., from newer vintages to older ones), because Comtrade provides only backward conversions. However, our algorithm can also be applied to forward conversion. For illustrative purposes, we focus on the example of converting HS 2007 to HS 2002, since that change lies in the middle of our sample and happened before the Great Trade Collapse. Results for other conversions look very similar.

In a simple descriptive application, we illustrate how our method may facilitate answering simple questions like: How many different varieties did the U.S. import each year over the past three decades? We show that using data converted to HS 1992 by Comtrade suggests that the number of products with positive U.S. imports has declined by about 10% from 1992 to 2018. By contrast, data converted into HS 1992 using our weights shows that this number has been fairly stable over that time span. [Pierce and Schott’s \(2012\)](#) method has a level effect in that there are a quarter less product categories available in the sample, but the trend looks similar to that suggested by our conversion. Such differences matter, e.g., when we want to investigate the evolution of product varieties à la [Broda and Weinstein \(2006\)](#), where a variety is defined as an exporter-product combination. For instance, when looking at the aftermath of the Great Trade Collapse, data converted using our technique suggest that the number of varieties imported by the U.S. had recovered to its pre-crisis level by 2011. The Comtrade-converted data, however, hint at a more sluggish regeneration, as there the number of imported varieties reaches its pre-crisis level only by 2015.

The following section summarizes existing methods. Section [4.3](#) then describes our conversion algorithm. Section [4.4](#) investigates the performance of our method, while Section [4.5](#) contextualizes it in simple descriptive exercises. Section [4.6](#) explores the heterogeneity and robustness of the conversion weights. Section [4.7](#) concludes.

4.2 Related Literature

Thus far there have been two main approaches to address the HS code conversion issue: the method applied by the UN Statistics Department which generates the Comtrade trade database (“Comtrade method”), and an algorithm developed by [Pierce and Schott \(2012\)](#).

According to the Comtrade method, “the data conversions [...] assign one single code (subheading) of an earlier HS edition to each HS 2017 subheading” ([UNSD, 2017](#), p.1). This method thus always assigns a weight of 1 to a single

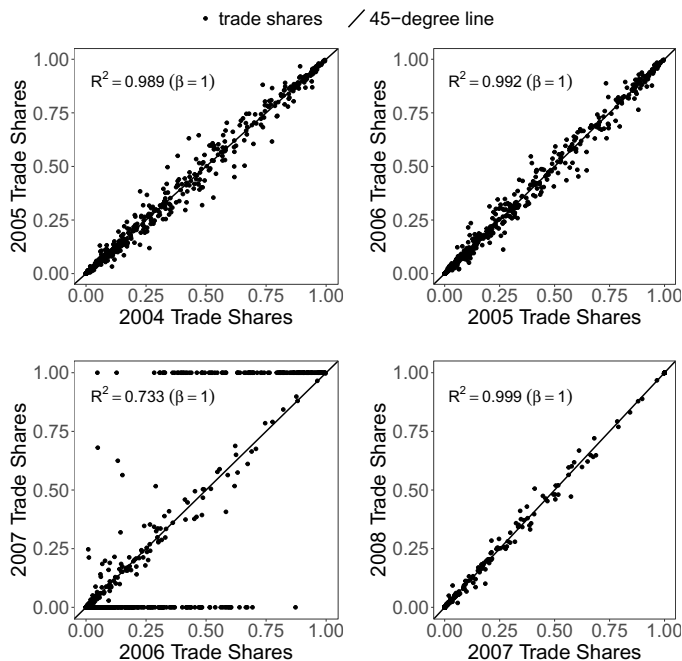
target HS code, and ignores the remaining target HS codes. In terms of the conversion problem depicted in Figure 4.1, this would, e.g., imply that the link between $a2$ and $b1$ is ignored with all $a2$ trade now converted to $b2$, thus imposing conversion weights of 0 and 1, respectively. In general, the Comtrade approach therefore tends to assign (too) much weight to one single code, and thus results in some codes not receiving any trade at all. Furthermore, using this method may lead to significant jumps at the extensive margin in HS code amendment years (Cebeci, 2012, p.5).

In terms of data continuity, the Comtrade assignment also leads to jumps at the intensive margin of trade. As depicted in Figure 4.2, when we look at trade data provided in HS 2002, the relative trade shares within conversion-dependent groups jump in a conversion year (here, in 2007), as the Comtrade method simplifies product links. By contrast, these product hierarchies (i.e., the product trade shares within groups) are highly persistent over time, as Figure 4.2 shows. Accordingly, there is no a priori reason to expect (large) jumps in product shares within their groups that are not due to conversion. The share of zeros jumps from 0% in 2006 to 42.4% in 2007. Note that this figure includes only groups (or, networks) with at least one $1:n$ or $m:n$ assignment from HS 2007 to HS 2002.

Alternatively to this method, Pierce and Schott (2012, p.58) developed an algorithm which avoids the data continuity issues by pooling the adjusted HS codes together and “assigning a common synthetic code to all HS codes in a growing family tree.” A family tree is an HS product group in our notation. While this method has been widely used (e.g., Felbermayr et al. (2019), Wagner and Zahler (2015), Piveteau and Smagghue (2019), Fort et al. (2018)), pooling linked codes together becomes an issue when analyzing trade data reported in many HS code vintages.

In fact, the longer the time series, the more HS codes are piled together. For instance, when converting 2018 data back to HS 1992 using this method, one synthetic cluster of HS codes ends up covering over a quarter of overall

Figure 4.2: Product Trade Shares within HS Product Groups over Time (Comtrade Method)



Notes: This figure plots for each product its trade share within its specific group across several years. We include only groups which contain at least one 1:n or m:n assignment from HS 2007 to HS 2002. The figure uses Comtrade data provided in HS 2002, i.e., 2007 and 2008 trade data have been converted into HS 2002, while previous years were originally reported in HS 2002. The data are restricted to 70 importers reporting in the latest available HS code version in a given year. Product groups are explained in the main text. The provided R-squared assumes a fit along the 45-degree line.

Data source: [UN Comtrade Database](#).

trade (cf. Table 4.2).⁷ Apart from creating synthetic codes not applied in practice, pooling that much trade across sectors may have significant conse-

⁷This largest synthetic group covers 607 different HS 2017 codes and includes sectors ranging from chemicals (ca. 27% of the codes) through paper products (10%) to machinery and electrical equipment (40%). The second largest HS code group which accounts for almost exactly a fifteenth of 2018 trade is crude oil and the corresponding HS code 270900 remains unchanged from 1992 to 2017, as can be seen by the unchanged share of 6.66% in each row of Table 4.2.

quences for what kind of variation is (not) picked up by product fixed effects. Moreover, Table 4.2 clearly illustrates that this method tends to artificially skew the distribution, especially when we consider that 1:1 links are never piled together.⁸

At the same time, one should point out that the [Pierce and Schott \(2012\)](#) method was initially developed to address monthly adjustments of individual 10-digit HTS codes (the U.S. version of the detailed HS codes), and was extended to the HS code conversion. For the initial purpose, their method seems adequate provided a small number of irregular code amendments. However, as indicated in the paper (p. 56), most of the HTS code adjustments are clustered around the general HS code vintage updates.

Similar methods to that of [Pierce and Schott \(2012\)](#) have been proposed by [Cebeci \(2012\)](#) and [Diodato \(2018\)](#). The former also pools product codes together but applies the method directly to HS codes and accounts for the fact that some developing countries may implement the HS code adjustment at a later stage.

On the other hand, [Diodato \(2018\)](#) analyzes product linkages across vintage versions using a network-adjusted structure of codes. Following [Pierce and Schott's \(2012\)](#) terminology, each network can be interpreted as a “family of related products.” Displaying the linkages as networks allows for a better visualization of the clusters and isolating individual linkages responsible for creating large clusters (“bridging nodes”). His method removes such nodes—using a community discovery algorithm, or manually inspecting the networks and removing such nodes—to deconstruct the larger networks. Although [Diodato \(2018\)](#)’s method helps alleviating the issue faced by bundling many products, it still pools numerous codes into synthetic groups. Furthermore, by removing bridging nodes, the method ignores by construction certain dependencies between product codes across vintages which may be of importance for particular products.

Below, we develop a simple conversion algorithm to alleviate the issues encountered when applying the above-mentioned approaches.

⁸The bottom three rows in Table 4.2 also illustrate that the Comtrade conversion method leads to bigger product groups but to a much lesser extent.

Table 4.2: Largest Trade Shares by Actual/Synthetic HS Code in 2018 Trade Data

Category	Covered vintages	Largest code	2nd largest	3rd largest	4th largest	5th largest
Actual	HS 2017	6.66%	2.17%	1.92%	1.61%	1.59%
Synthetic (P&S method)	HS 2012–HS 2017	13.43%	6.66%	4.17%	2.17%	1.92%
Synthetic (P&S method)	HS 2002–HS 2017	16.00%	6.66%	4.17%	3.43%	2.35%
Synthetic (P&S method)	HS 1992–HS 2017	25.74%	6.66%	4.17%	3.43%	1.92%
Comtrade converted	HS 2012	6.66%	2.17%	1.92%	1.61%	1.61%
Comtrade converted	HS 2002	6.66%	4.17%	2.46%	2.25%	1.92%
Comtrade converted	HS 1992	6.66%	4.19%	3.44%	2.46%	1.91%

Notes: This table depicts trade shares in overall trade of the largest products in the year 2018. The first row displays the trade share of the largest HS 2017 products. The following three rows depict the trade shares of the largest products after applying [Pierce and Schott's \(2012\)](#) method to standardize the vintages from 2017 to those of 2012, 2002, and 1992, respectively. The bottom three rows summarize the shares of the top five products after the Comtrade conversion of 2018 trade data into HS 2012, HS 2002, and HS 1992.

Data source: [UN Comtrade Database](#).

4.3 Methodology

This section briefly introduces the data we use, before explaining our conversion algorithm.

Data. The data used in this paper have been extracted from UN Comtrade and include all import data reported by UN members between 1988 and 2018. This data source is standard to the international trade literature. The import figures are corrected for re-import numbers, as the raw import data includes re-imports. The analysis focuses on those countries that in each year applied

the latest available HS vintage, and are thus considered “compliers.”⁹ An exemplary phasing in of the HS 2007 vintage is shown in Figure 4.A.B.1 of Appendix 4.A.B. The bilateral country pairs include complying importers and all of the partners with which they had at least one non-zero trade flow in a given year.

Furthermore, the analysis includes import data that have been converted by Comtrade to a given HS vintage (e.g., the 2007 and 2008 numbers converted by Comtrade into HS 2002 used in Figure 4.2).

Persistence. The key feature of trade data we exploit is the persistence of trade flows over time. The data show that, e.g., if a specific product was relatively important for a country pair in one year, it is very likely to be important for that country pair in surrounding years. Similarly, if a country is a large importer of a specific product in one year, it is likely to also have relatively high imports of that product in other years. Accordingly, when we look at product hierarchies within groups over time (for a given HS vintage), they appear to be fairly stable (cf. Figure 4.2). Product-level trade flows at the bilateral (i.e., country pair) level also reveal a relatively high persistence over time. Table 4.3 shows this by regressing the share of bilateral product-level flows in their groups on their first lags (cf. also Figure 4.A.B.2). Therefore, we use information on bilateral flows reported in one HS vintage to predict how trade flows should be allocated within a target vintage in cases where the allocation is not a priori clear (i.e., neither 1:1 nor $m:1$). The reasoning behind our method is that we do not expect (substantial) changes in product hierarchies within groups due to a change in classification standards. Moreover, using disaggregated rather than aggregated product-level trade flows is appealing for two reasons. Not only does this add information on how trade flows are distributed across country pairs, it also avoids encountering regressions where we have less observations than unknowns.

⁹For instance, Vietnam reported its 2017 trade figures using the HS 2012 version and would thus not be included when estimating conversion weights between HS 2012 and HS 2017. However, it did change to the latest available version in 2012, so it would be included when converting HS 2012 to HS 2007.

Table 4.3: Persistence of Bilateral Product-Level Trade Flows

Year	2005	2006	2007	2008	2009
All HS codes	0.874 (0.009)	0.867 (0.008)	<i>skipped</i> <i>due to</i>	0.854 (0.009)	0.862 (0.009)
Adjusted HS codes	0.876 (0.020)	0.889 (0.018)	<i>vintage</i> <i>change</i>	0.880 (0.024)	0.836 (0.045)

Notes: This table depicts first-order autoregressive coefficients for bilateral trade flows at the product level for the years 2005, 2006, 2008 and 2009 (2007 was left out due to the HS vintage change). The bilateral product-level flows are scaled by overall trade in the product's group in a given year. The data include 70 countries which reported in the latest available HS code version in the years 2005–2009. The first row depicts the coefficients for all trade flows, whereas the second is restricted to HS codes facing a change in 2007 (i.e., which have a non-1:1 assignment between HS 2002 and HS 2007). We regress the bilateral product-level flows on their respective first lags without a constant. Robust standard errors are reported in parentheses.

Data source: [UN Comtrade Database](#).

Algorithm. The algorithm can be split into three components. As a first step, we identify groups of HS codes that are linked together, and keep those groups which contain ex-ante unspecified conversion weights.¹⁰ Note that one group would be pooled to one synthetic cluster when applying [Pierce and Schott's \(2012\)](#) method.

Second, in order to maximize the number of observations, we use trade flows at the country-pair level. Specifically, we prepare the trade data in the form of an $(I \times J) \times K$ -matrix separately for each group and year, where an entry is equal to $X_{ij,k}$, i.e., the exports from country $i \in I$ to country $j \in J$ of product $k \in K$ in a specific group and year. For notational simplicity we

¹⁰ Accordingly, groups that are left out will contain only conversion weights which are a priori known to be equal to zero or one.

suppress time and group subscripts. We restrict the sample to importers that actually switched reporting standards in the years we focus on. Moreover, in order to consider product hierarchies within groups, we scale each pair’s product-level imports by the total imports within a group. Note that this also allows us to control for group-specific trends. This procedure yields our scaled flows, $x_{ij,k}$, defined as

$$x_{ij,k} := \frac{X_{ij,k}}{\sum_{\hat{i}, \hat{j}} \sum_{\hat{k}} X_{\hat{i}\hat{j},\hat{k}}},$$

which will be the inputs in our regressions. Recall that the scaling is computed separately for each group and year. It is worth noting that this could be further adjusted to calculate, e.g., region- or country-group-specific conversion weights, as addressed in Section 4.6.

The third component of the algorithm aims to fit the trade flows which are reported in the “target” reporting standard by using flows reported in the “initial” reporting standard and a constrained least-squares procedure. Note that these flows will, by construction, not belong to the same year. Hence, we (implicitly) assume that product-level bilateral flows relative to overall group flows are fully persistent over time. That is why most figures report a fit along the 45-degree line. In Section 4.6 we relax this assumption.

As an example, consider “initial” HS codes $k \in K$ used in year t_0 , and “target” HS codes $s \in S$ used in year t_1 .¹¹ We need three types of constraints in our optimization. First, the conversion weights for an initial product k need to sum up to one across target products s , otherwise overall trade might artificially increase or decrease across vintages. Second, the weights should be non-negative, else we would subtract product flows, which might improve the fit but would make no intuitive sense.¹² Third, some weights are known to be equal to zero or one from the correlation tables provided by the WCO.¹³ Our

¹¹Note that we can also estimate conversion weights going, e.g., from 2017 to 2016 data, in which case $t_0 > t_1$.

¹²These two constraints together restrict conversion weights to be smaller than or equal to one.

¹³The correlation tables were downloaded from the UN Comtrade webpage.

constrained minimization problem is outlined in Equation (4.1).

$$\begin{aligned} \min_{\{\beta_{k,s}\}} \quad & \sum_s \sum_{i,j} \left(x_{ij,s}^{t_1} - \sum_k x_{ij,k}^{t_0} \beta_{k,s} \right)^2 \\ \text{s.t.} \quad & \beta_{k,s} \geq 0 \quad \forall k, s \\ & \sum_s \beta_{k,s} = 1 \quad \forall k, \end{aligned} \tag{4.1}$$

where $\beta_{k,s}$ is the conversion weight from initial product k to target product s . Note that various $\beta_{k,s}$ are a priori known to be equal to zero or one. This problem can then be estimated separately for each group to reduce the size of the coefficient matrix, given that conversion weights across groups are, by construction, always equal to zero. This is the specification we use for our main results, and variations can be found in Section 4.6.

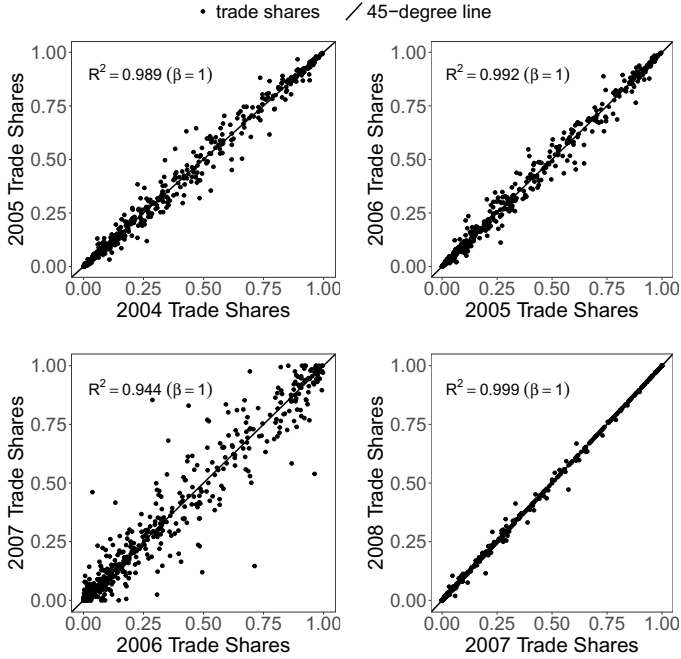
In order to apply our conversion weights to trade data, one has to multiply the trade flow reported in the initial HS vintage by the respective weight, and then sum up across initial codes.¹⁴ Or, put differently, one computes the fitted value, $\hat{x}_{ij,s}^{t_0} = \sum_k x_{ij,k}^{t_0} \hat{\beta}_{k,s}$. Furthermore, in order to convert one vintage to another that is not adjacent, we multiply the weights across conversions. For instance, if we want to convert HS 2017 codes $k \in K$ to an HS 2007 code l , which are linked through HS 2012 codes $s \in S$, we compute $\hat{x}_{ij,l}^{t_0} = \sum_s \sum_k x_{ij,k}^{t_0} \hat{\beta}_{k,s} \hat{\beta}_{s,l}$. The weights $\hat{\beta}_{k,l} (= \sum_s \hat{\beta}_{k,s} \hat{\beta}_{s,l})$ have the same properties as those constructed based on adjacent vintages $\hat{\beta}_{k,s}$ and $\hat{\beta}_{s,l}$.

4.4 Results

In Section 4.2, or in Figure 4.2, we illustrate how product shares within groups jump in 2007 when applying Comtrade's conversion method, despite product hierarchies being strikingly stable over time. When applying our estimated conversion weights, we obtain a substantially smoother transition from 2006 to 2007, as depicted by Figure 4.3. The share of zeros in the graph drops heavily from 42.4% to only 5.6%. We thus manage to reduce the share of HS

¹⁴Also, we recommend rounding the weights to, e.g., the sixth digit after the comma, in order to set weights which are numerically indistinguishable from zero to 0.

Figure 4.3: Product Trade Shares within HS Product Groups over Time (Our Conversion Weights)



Notes: This figure plots for each product its trade share within its specific group across several years. We include only groups which contain at least one 1:n or m:n assignment from HS 2007 to HS 2002, and trade data for 72 countries which used the latest available HS code classification in each year between 2004 and 2008. The figure uses Comtrade data provided in the latest available HS vintage, and transforms 2007 and 2008 data using our weights back into HS 2002. Product groups and the algorithm to obtain our conversion weights are explained in the main text. The provided R-squared assumes a fit along the 45-degree line.

Data source: [UN Comtrade Database](#).

2002 products (in total 5,224) with no trade from 6% to below 1%. Moreover, although we tend to slightly decrease the variation between the years 2007 and 2008, it is important to note that it was already very low in the original vintage. Results for other vintages look very similar. Hence, our method alleviates the aforementioned issues introduced by the Comtrade method. [Pierce and](#)

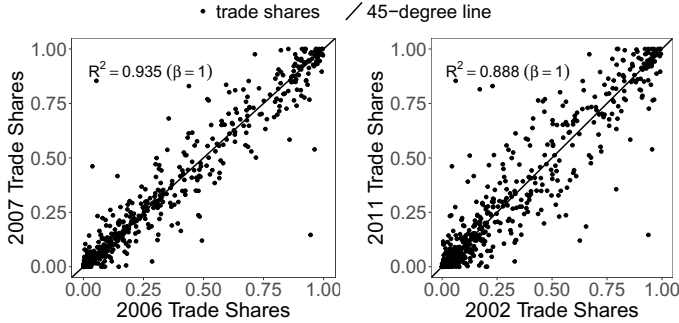
Schott's (2012) method would reduce the number of product categories in the graph by roughly three-quarters from 706 to 183, given that we identify 183 groups in our conversion.

We can verify how these estimated conversion weights perform with other year constellations. Specifically, in Figure 4.4, we convert 2011 data into HS 2002 using the conversion weights from above, and compare the resulting product hierarchies within groups with those in the year 2002. This is the widest year range we can take when comparing HS 2007 with HS 2002. We also re-include the comparison of 2006 shares and converted 2007 shares (left panel). The product shares closely follow the 45-degree line in both panels, with (intuitively) somewhat more variation in the wider year range. In order to compare the product shares in different years but with unconverted data, we can confront the two "initial" as well as the two "target" years. Figure 4.A.B.3 in Appendix 4.A.B illustrates that the similarity should be expected (and desirable) given the persistence of product hierarchies measured in the reported HS vintages. Overall, this suggests that, if our method is deemed appropriate to convert 2007 data into HS 2002, it also allows for a conversion of later years into that HS vintage.

Importantly, we obtain similar outcomes when comparing absolute (i.e., not scaled by group totals) values. In Figure 4.A.B.4 of Appendix 4.A.B we reproduce Figure 4.3 using the log of product-level trade flows. There are only a few products for which we get somewhat large deviations. However, these products tend to be relatively small. This, in turn, suggests that exploiting product group hierarchies in our estimation routine seems to yield conversion weights that are suitable for the analysis of absolute trade flows.

In Section 4.3 we explain how to standardize HS vintages which are non-adjacent by multiplying the conversion weights for the intermediate vintages. In Figure 4.A.B.5 we plot the log of product-level trade flows for products which have at least one non-1:1 assignment over our sample period. The figure shows that this procedure does not create sizable inconsistencies when converting all the way back to HS 1992, even in recent years, although we do eliminate a little less than 1% of the HS 1992 products (5,020 in total)

Figure 4.4: Product Trade Shares within HS Product Groups (Our Conversion Weights)



Notes: This figure plots for each product its trade share within its specific group across several years. We include only groups which contain at least one 1:n or m:n assignment from HS 2007 to HS 2002, and trade data for 63 countries which used the latest available HS code classification in each of the years used in the plot. The figure uses Comtrade data provided in the latest available HS vintage, and transforms 2007 and 2011 data using our weights back into HS 2002. Product groups and the algorithm to obtain our conversion weights are explained in the main text. The conversion weights are calculated using 2006 and 2007 trade data. The provided R-squared assumes a fit along the 45-degree line.

Data source: [UN Comtrade Database](#).

over the years. In comparison, the Comtrade conversion drops almost 10% of the HS 1992 products by 2017. Importantly, note that our technique does not artificially fix the importance of a product over longer time spans. The multiplication of the weights, which are recalculated whenever there is an amendment, implies that long-run trends should be captured by our method, as long as the HS vintage adjustments reflect such trends.

Additionally, as illustrated in Table 4.4, our method seems not to skew the distribution, as opposed to [Pierce and Schott's \(2012\)](#) method (cf. Table 4.2). Accordingly, at least in terms of relative size, the converted product categories remain interpretable as such. The trade shares obtained using our conversion method perform slightly better than the Comtrade method also in this dimension (cf. bottom three rows in Table 4.2).

Table 4.4: Largest Trade Shares by Actual/Converted HS Code in 2018 Trade Data (Our Conversion Weights)

Category	Covered Vintages	Largest Code	2nd largest	3rd largest	4th largest	5th largest
Actual	HS 2017	6.66%	2.17%	1.92%	1.61%	1.59%
Converted	HS 2012	6.66%	2.17%	1.92%	1.73%	1.61%
Converted	HS 2002	6.66%	2.58%	2.25%	1.92%	1.73%
Converted	HS 1992	6.66%	3.43%	2.09%	1.92%	1.73%

Notes: This table depicts trade shares in overall trade of the largest products in the year 2018. The first row displays the trade share of the five largest HS 2017 products. The remaining rows depict the trade shares of the five largest products after applying the algorithm described in this paper and converting 2018 data from HS 2017 to HS 2012, HS 2002, and HS 1992, respectively.

Data source: [UN Comtrade Database](#).

4.5 Descriptive Exercises

Having shown in the previous section that the proposed algorithm addresses the conversion issues caused by prevalent methods, this section analyses the impact our method has on descriptive applications. We provide illustrative examples related to both the demand and supply side. These are meant to demonstrate, e.g., how motivating facts may change depending on the conversion method that is applied.

Import demand trends. One key dimension of assessing welfare gains from trade continues to be the gains from product variety. As shown in [Broda and Weinstein \(2006\)](#), product varieties have significantly increased over the past decades, substantially contributing to the increased role of trade in the global economy ([Hsieh et al., 2020](#)). In order to track product varieties correctly, one ought to have a continued standardised classification of products throughout the sample, requiring thus to correct for the HS code adjustments every five

years.¹⁵

Taking U.S. imports as an example, we illustrate in Figure 4.5 that using our method we obtain a stable flow of 6-digit HS products being imported each year into the United States. Applying instead the Comtrade method to convert all codes back to the HS 1992 vintage leads to artificial reductions of imported products in 2002 and especially in 2007—years in which the United States updated its applied HS vintage. Therefore, over time the number of imported products appears to be shrinking, despite there being no obvious reason for that to be the case.

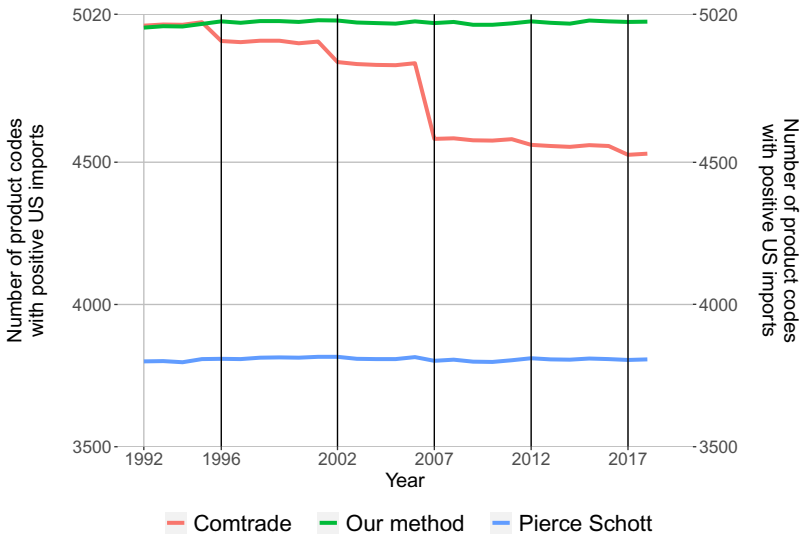
This, in turn, leads to different interpretations when looking at recent trends of U.S. imported product varieties after the Great Trade Collapse (cf. Figure 4.A.B.6 in Appendix 4.A.B). Following Broda and Weinstein (2006), we define a variety as a product imported from a specific origin country. While data based on our method imply a relatively swift return to pre-crisis levels, data based on the Comtrade method would suggest that the number of imported product varieties did not recover until 2015.

Meanwhile, applying the Pierce and Schott (2012) method leads to nearly a quarter less of products (and thus product varieties). In terms of recovery patterns, the trends captured by the Pierce and Schott (2012) data resemble those suggested by our method, as both versions show a strong post-crisis upturn. This pattern is also in line with aggregate U.S. imports (depicted as bars in Figure 4.A.B.6), which had recovered to pre-crisis values by 2011.

Supply-side trends. While the example above is meant to illustrate the benefits of our conversion method to study import demand trends, similar statements may be made on the supply side. For instance, Figure 4.6 investigates entry and exit rates of product-exporter combinations for 71 countries that reported data between 1996 and 2017. For these countries, we find an unwarranted jump nearly doubling the exit rate in the year 2007 when applying

¹⁵In fact, the aggregate price index by Broda and Weinstein (2006) accounts for the possibility of codes being merged and split depending on whether their importance in trade decreases or increases. They do not account, however, for trade being reclassified due to other reasons (cf. Appendix 4.A.A).

Figure 4.5: Number of Products Imported by U.S. according to Different Conversion Methods



Notes: This figure plots for each year the number of 6-digit HS products (in the case of Pierce Schott: synthetic codes) imported by the United States. The three lines depict the different numbers of products based on the Comtrade, the [Pierce and Schott \(2012\)](#), and our conversion method; with the first and the last method used to convert all codes back to the HS 1992 vintage (containing 5,020 codes).

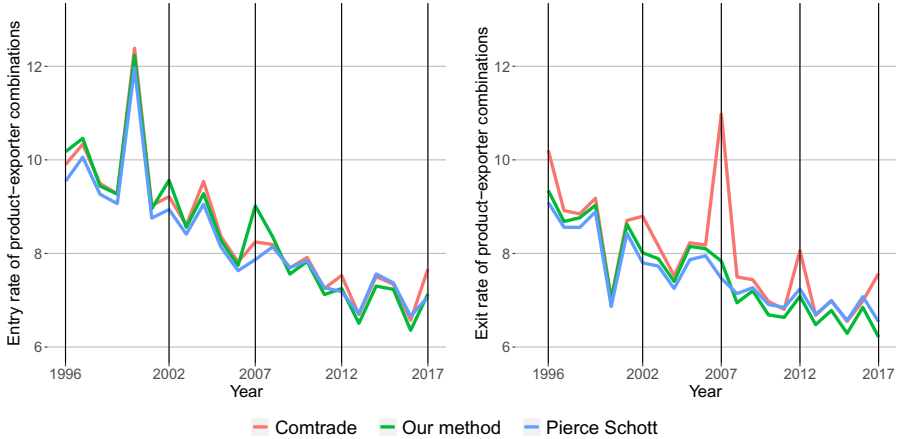
Data source: [UN Comtrade Database](#).

the Comtrade method, in line with findings of [Cebeci \(2012\)](#). Similar artificial jumps—while less pronounced—occur in each of the other years facing an HS code amendment. Otherwise, all three methods seem very much aligned with respect to the exit rate statistics.^{16,17}

¹⁶Note that not all 71 countries each time adjusted the HS vintage in the same year, leading to the fact that the HS-conversion-related jumps may be lagged for some countries. Reducing the pool to the 35 complying countries—i.e., those which each year applied the latest available HS vintage—further sharpens the jumps seen in the vintage adjustment years. However, the overall picture does not change.

¹⁷The Pierce-Schott method results in very similar entry and exit rates as the two other methods (despite having a quarter fewer products) largely due to the share of unchanged HS codes. Restricting the products in Figure 4.A.B.6 to those HS codes that change at least once results in the entry and exit rates of the Pierce-Schott method being roughly three percentage points lower than those based on the other approaches.

Figure 4.6: Entry and Exit Rates (in Percent) for Product-Exporter Combinations using the Different Conversion Methods



Notes: This figure plots for each year the entry and exit rates (in percent) for product-exporter combinations. Products are defined as 6-digit HS codes (in the case of Pierce Schott: synthetic codes) and observations are restricted to those 71 countries which reported trade data for the years 1995 through 2017. The three lines depict the different rates based on the Comtrade, the [Pierce and Schott \(2012\)](#), and our conversion method; with the first and the last method used to convert all codes back to the HS 1992 vintage. The large 2000 increase in the entry rate seen in all three versions relates to a large wave of trade liberalisation, in particular to China.

Data source: [UN Comtrade Database](#).

With regard to the entry rate, all three methods seem more aligned; with the notable exception of using our method in 2007. The jump may indicate that our method generates somewhat too many positive product flows in that year. However, this increase—as well those in 1996 and 2002—are not nearly as pronounced as those of the Comtrade method regarding exit rates.

4.6 Heterogeneity and Robustness of Weights

This section presents variations of our main specification outlined in Section 4.3.

Product-specific trends. One potential concern is that we do not account for product-specific trends in the data. In order to tackle the issue of product categories potentially becoming more or less relevant over time, we can estimate product-specific autoregressive coefficients which are then used to augment our minimization problem. Specifically, using all years in which the “initial” HS vintage was the latest available (i.e., the years where this specific classification standard was officially in place), let us estimate the following regression

$$x_k^t = \rho_k x_k^{t-1} + \varepsilon_k^t,$$

where x_k^t is the (aggregate) trade flow of product k in year t , ρ_k is the product-specific autoregressive coefficient, and ε_k^t is an error term. We can then assign more weight to flows in t_0 where we measured a higher persistence over time. The objective function for both $t_0 < t_1$ or $t_0 > t_1$ is illustrated in Equation (4.2).

$$\sum_s \sum_{i,j} \left(x_{ij,s}^{t_1} - \sum_k (x_{ij,k}^{t_0} \hat{\rho}_k^{t_1-t_0}) \gamma_{k,s} \right)^2, \quad (4.2)$$

where $\gamma_{k,s}$ is the new conversion weight.¹⁸

¹⁸In order for trade shares to still sum up to one in year t_0 , we rescale the flows by total trade within groups after multiplying them with the product-specific autoregressive coefficients. Also, note that this reasoning can be extended to higher-order autoregressive processes. For instance, let $\bar{t}_0$ and $\bar{t}_1$ be the earliest and latest year, respectively, in which the “initial” reporting standard was mostly applied. Further, let T_0 be the set containing these two years as well as those in between. We can then estimate

$$x_{ij,k}^{\bar{t}_0} = \sum_{z=\bar{t}_0}^{\bar{t}_0-1} \rho_{k,(\bar{t}_0-z)} x_{ij,k}^z + \nu_{ij,k}^{\bar{t}_0},$$

where $\rho_{k,l}$ is the coefficient on the l^{th} lag, and $\nu_{ij,k}^{\bar{t}_0}$ is an error term. Correspondingly, if, e.g., $t_1 - \bar{t}_0 = 1$, the adjusted objective function looks as follows

$$\sum_s \sum_{i,j} \left(x_{ij,s}^{t_1} - \sum_k \left(\sum_{z \in T_0 \setminus \{\bar{t}_0\}} x_{ij,k}^z \hat{\rho}_{k,(t_1-z)} \right) \tilde{\gamma}_{k,s} \right)^2.$$

When applying the adjusted objective function, our results remain largely the same (cf. Figure 4.A.B.7 of Appendix 4.A.B). Overall, we see few weights receiving substantially different values. The majority of conversion weights remain very similar, as the correlation coefficient for not a priori determined weights is 0.97.

Multiple years. A further limitation might be the use of only one initial and one target year. We therefore re-estimate the conversion weights by using average product-level bilateral flows over the “initial” and “target” periods, each scaled by their respective total group flows. This averaging may be particularly important for products ordered irregularly such as airplanes or trains, as well as for goods subject to multi-annual seasonal patterns, especially in agriculture. Nonetheless, when applying the algorithm on conversion weights for the vintages HS 2007 to HS 2002, we obtain very similar results between averaged and single-year trade data. As Figure 4.A.B.8 shows for averaged (2007–2009 and 2004–2006) and single-year (2007 and 2006) data, most weights are close to the 45-degree line, with a correlation coefficient of 0.96.

Sampling. To further verify the robustness of our weights, we re-estimate them using subsamples of our data. More precisely, we randomly divide the data into 20 equi-sized bins and look at the distribution of weights after running the regressions twenty times, where each regression leaves out one block (or, five percent of the observations). The conversion weights appear to be fairly stable. For instance, when we convert HS 2007 to HS 2002, the difference between the maximum and minimum estimated coefficient for a given initial and target product pair rarely surpasses 0.1, as illustrated in Figure 4.A.B.9.¹⁹ Moreover, in order to see whether large fluctuations happen for products which are relatively important, in Figure 4.A.B.10 we plot this difference against the

¹⁹This figure excludes pre-specified weights (i.e., conversions known to be equal to zero or one), as these would, by definition, always yield a difference of zero. Also, keeping in mind that about 85% of the groups in the conversion from HS 2007 to HS 2002 have less than five weights which are not pre-specified, a difference of less than ten percentage points can be considered small.

multiplication of the relative size of the initial and target products in their groups. The northeastern area of the graph depicts relatively important links which feature rather unstable conversion weights. As can be seen in Figure 4.A.B.10, that area is almost empty. Results for other vintages look very similar.

Overall, this exercise suggests that our weights are generally not driven by only a few country pairs, and that if they are, it mostly concerns relatively unimportant products. Also, note that taking the difference between the maximum and minimum coefficients is the most conservative test we may choose.

Aggregation level. As discussed in Section 4.3, bilateral product-level flows tend to be persistent over time, although not perfectly so. Product-level imports, on the other hand, tend to fluctuate less over time for a given country, probably because a country's demand structure changes more sluggishly than bilateral transactions at the product level. Hence, one may favor using importer-level rather than country-pair-level data—i.e., aggregating our input variables $x_{ij,k}$ across exporters $i \in I$ —in our optimization routine, despite reducing the number of observations substantially. In a similar vein as in the subsection above, we thus look at how conversion weights estimated using importer-level data compare with our baseline results. In Figure 4.A.B.11 we do this for the conversion from 2007 to 2006 data (or, from HS 2007 to HS 2002).²⁰ Most absolute differences between the coefficients are again below 0.1. Furthermore, in order to assess whether large fluctuations appear for products which are relatively important, in Figure 4.A.B.12 we multiply the coefficients by the relative size of the initial and target products in their groups, and plot them against each other. The outcome reveals that if there are large differences, we find them only for relatively small products. Results for other vintages look very similar.

Given that, although the results are similar, the importer-level data tends to estimate more coefficients (very) close to zero, which is likely due to the lower number of observations, we stick to the country-pair-level data for our

²⁰As above, this figure excludes pre-specified weights (i.e., conversions known to be equal to zero or one), as these would, by definition, always yield a difference of zero.

baseline results.

Regions. Our calculations thus far were focused on global trade patterns. As countries differ in their product mixes though, the estimated conversion weights may vary across jurisdictions. In Figure 4.A.B.13 we compare how the conversion weights calculated using the entire sample perform relative to country-group-specific weights. Panel 4.A.B.13a compares the 2007 trade shares converted into HS 2002 using country-group-specific weights with the 2006 shares for both OECD and non-OECD countries. In panel 4.A.B.13b we convert the 2007 figures using our “global” conversion weights (i.e., those we used in Figure 4.3). Intuitively, the country-group-specific conversion provides a better fit for both country groups, although the global weights also work fairly well. It should be expected that the estimates from the entire sample better suit the OECD countries, since most large pairs belong to that country group.

4.7 Conclusion

Product-level trade data often suffer from intertemporal inconsistencies due to classification changes. Existing methods either oversimplify product links across HS vintages (UNSD, 2017) or bulk the changed HS codes together (Pierce and Schott, 2012). This paper presents an approach to overcome the issues related to those methods by estimating conversion weights between linked HS codes.

The calculated conversion weights avoid the intertemporal inconsistencies encountered with the Comtrade approach (UNSD, 2017), and also do not significantly change the distribution of trade among HS codes (unlike the Pierce and Schott (2012) method). Robustness checks indicate that the weights are not significantly affected by the set of year and country pair observations we include. Similarly, the weights remain relatively stable when using importer-specific rather than country-pair-specific data in the optimization.

The weights may be readily applied to existing trade data and combined

for multiple vintage changes (e.g., multiplying HS 2007 to HS 2002 and HS 2002 to HS 1996 weights when converting HS 2007 to HS 1996). If required, the algorithm can be applied to particular regions or country groups (e.g., OECD vs. non-OECD countries), which might improve the accuracy of the conversion.

Our proposed conversion method also allows for a more sensible descriptive investigation of trade patterns. When looking at imported product varieties as in [Broda and Weinstein \(2006\)](#), our method largely smooths the time series and is in line with general trade developments. The same holds for the analysis of exporter entry and exit rates, which relates to a growing literature investigating firm-level exporter dynamics (cf. [Alessandria et al. \(2020\)](#) for an overview), in the sense that we should be able to compare the trends in micro and macro data.

4.A Appendices to Chapter 4

4.A.A The Harmonized System Codification

The classification of products into Harmonized System (HS) codes is administered by the World Customs Organisation through the HS Convention signed in 1983. Each of the 150+ signatory countries is obliged to align its tariff schedules with the HS code classification. As the HS code system at its most disaggregated level reaches 6 digits, individual countries can expand their tariff schedule to a more detailed classification—at the 8-, 10-, or even 12-digit level—as long as those codes adhere at the 6-digit level to the internationally agreed standard (art.3(3) of the HS Convention).

Amendments to HS codes are proposed by the Harmonized System Committee, consisting of representatives from each contracting country. This Committee shall “propose amendments [to the current HS classification system] as may be considered desirable, having regard, in particular, to the needs of users and to changes in technology or in patterns of international trade” (art.7(1)a of the HS Convention). The proposed changes are then recommended to all contracting parties of the HS Convention and—provided no objections are received within 6 months—enter into force on the 1st of January of the following year (art.16 of the HS Convention).

Changes to the HS code classification may also be due to other reasons than those explicitly listed in the HS Convention. According to the Compilers Manual of the International Merchandise Trade Statistics (2004), some of the HS code changes in 2002 reflected “user needs but many of which were made at the request of other international organizations seeking to be able to identify trade in sensitive goods, including other hazardous materials” (p.32).

The latest HS code vintage to be adopted (HS 2022) will *inter alia* provide new classifications to address environmental issues (better tracking of e-waste), health concerns (separate codes for fentanyl), public security (clearer classification of components for explosives), or to facilitate international medical research (by reclassifying placebo drugs without having to differentiate by their ingredients). However, the 2022 amendments also take into account tech-

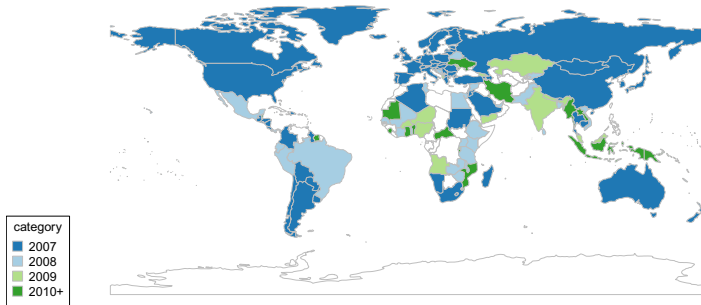
nological progress through separate HS codes for smartphones, drones, and a reclassification of the glass fiber products.

The algorithm in this paper is based on the correlation tables provided by the World Customs Organisation (and published by UN Comtrade²¹). Despite these tables being “examined by the Harmonized System Committee, they are not to be regarded as constituting classification decisions taken by that Committee; they constitute a guide published by the Secretariat and whose sole purpose is to facilitate implementation of the 2017 edition of the Harmonized System” (WCO, 2015). The correlation tables are drawn up in a way that is “as comprehensive as possible” (WCO, 2015)—it is therefore not required for a majority of country members of the HS Committee to agree on every link between a given old and new code. The final correspondence tables published by the Secretariat may therefore result in being too broad, i.e., creating too many links. This characteristic further supports our approach of letting trade data “speak for itself” in terms of which HS codes have a quantitative correspondence.

²¹The tables can be obtained here: <https://unstats.un.org/unsd/trade/classifications/correspondence-tables.asp>, accessed on February 1, 2022.

4.A.B Additional Figures

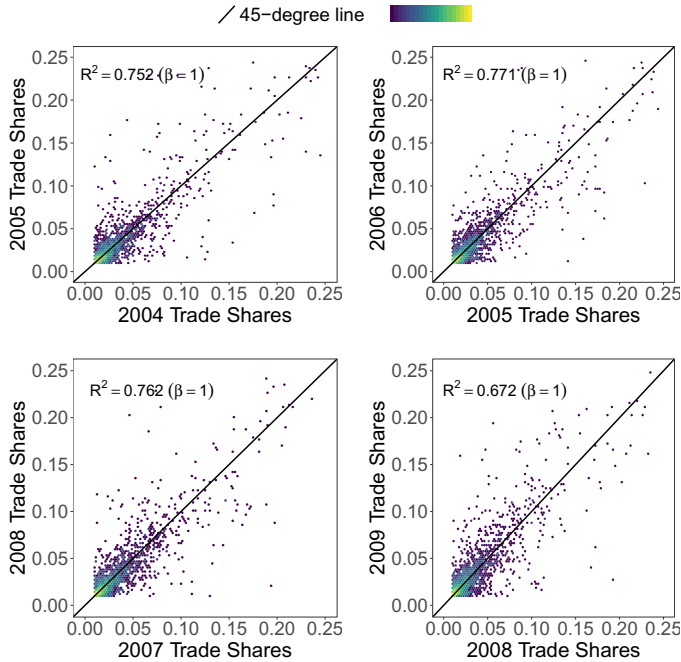
Figure 4.A.B.1: Year of HS 2007 Vintage Adoption by Country



Notes: This figure plots for each country the year in which it first adopted the HS 2007 code vintage. Note that some countries never adopted the HS 2007 classification (marked by white filling).

Data source: [UN Comtrade Database](#).

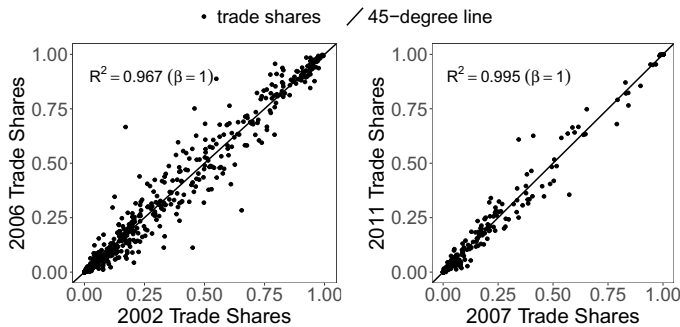
Figure 4.A.B.2: Bilateral Product-Level Trade Shares within HS Product Groups over Time



Notes: This figure plots for each bilateral product-level flow its trade share within its specific group across several years. We include only groups which contain at least one 1:n or m:n assignment from HS 2007 to HS 2002. The figure uses Comtrade data provided in the original HS code version and is restricted to 70 importers reporting in the latest available HS code version in a given year. Due to the volume of observations, the data is plotted in 100 bins (per axis), and a lighter color is applied to areas with more observations. Product groups are explained in the main text. The provided R-squared assumes a fit along the 45-degree line. For illustrative purposes, the axes are capped at 0.25, and observations are excluded for which the trade share was below 0.01 in both years.

Data source: [UN Comtrade Database](#).

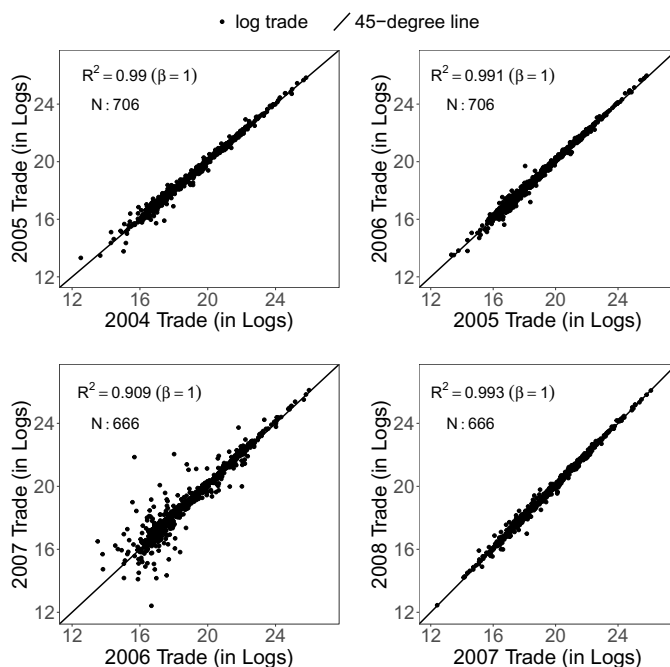
Figure 4.A.B.3: Product Trade Shares within HS Product Groups (HS Codes as Reported)



Notes: This figure plots for each product its trade share within its specific group across several years. We include only groups which contain at least one 1:n or m:n assignment from HS 2007 to HS 2002, and trade data for 70 countries which used the latest available HS code classification in each of the years used in the plot. The figure uses Comtrade data provided in the latest available HS vintage. Product groups are explained in the main text. The provided R-squared assumes a fit along the 45-degree line.

Data source: [UN Comtrade Database](#).

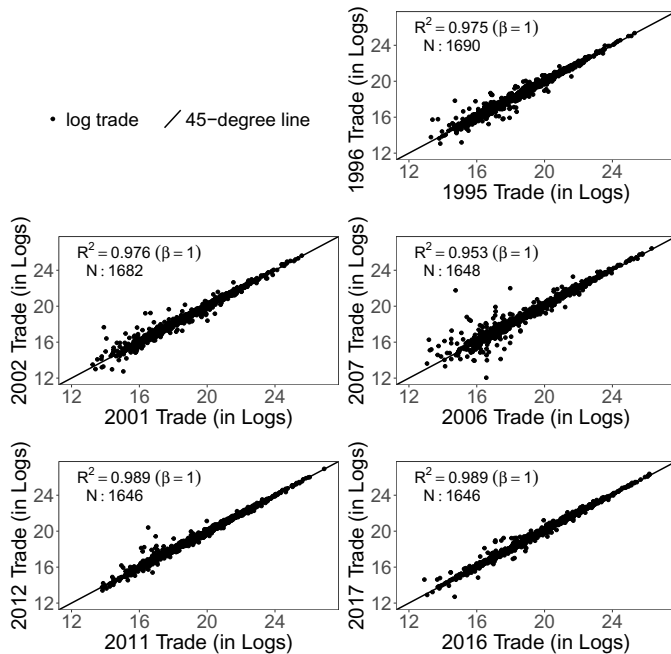
Figure 4.A.B.4: Log of Product-Level Trade over Time (Our Conversion Weights)



Notes: This figure plots for each product its log of overall trade across several years. We include only products in groups which contain at least one 1:n or m:n assignment from HS 2007 to HS 2002, and trade data for 70 countries which used the latest available HS code classification in each year between 2004 and 2008. The figure uses Comtrade data provided in the latest available HS vintage, and transforms 2007 and 2008 data using our weights back into HS 2002. Product groups and the algorithm to obtain our conversion weights are explained in the main text. The provided R-squared assumes a fit along the 45-degree line.

Data source: [UN Comtrade Database](#).

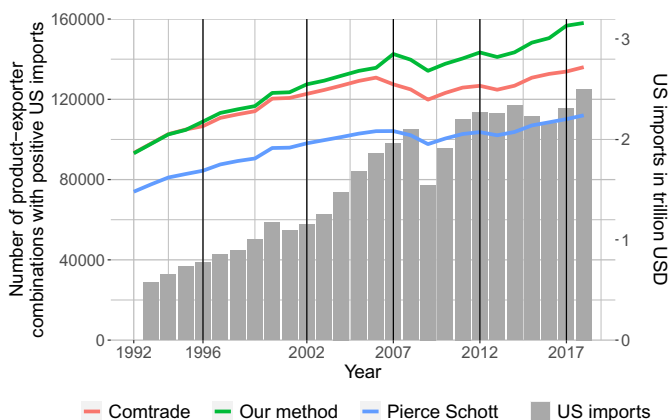
Figure 4.A.B.5: Log of Product-Level Trade over Time (Our Conversion Weights – HS 1992)



Notes: This figure plots for each product its log of overall trade across several years. We include only products which had at least one non-1:1 assignment over the period 1995–2017, and trade data for 38 countries which used the latest available HS code classification in each year between 1995 and 2017. The figure uses Comtrade data provided in the latest available HS vintage, and transforms all years after 1995 using our weights back into HS 1992. Product groups and the algorithm to obtain our conversion weights are explained in the main text. The provided R-squared assumes a fit along the 45-degree line.

Data source: [UN Comtrade Database](#).

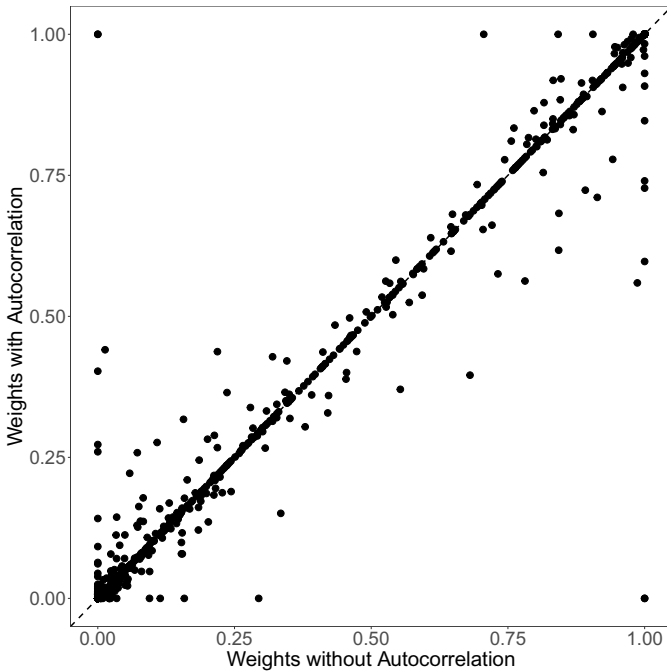
Figure 4.A.B.6: Number of Varieties Imported by the U.S. according to Different Conversion Methods



Notes: This figure plots for each year the number of varieties imported by the United States. A variety is defined as an exporter-product combination (in the case of Pierce Schott: synthetic product codes). The three lines depict the different number of products based on the Comtrade, the [Pierce and Schott \(2012\)](#), and our conversion method; with the first and the last method used to convert all codes back to the HS 1992 vintage. The bars show total U.S. imports in trillion current USD.

Data source: [UN Comtrade Database](#).

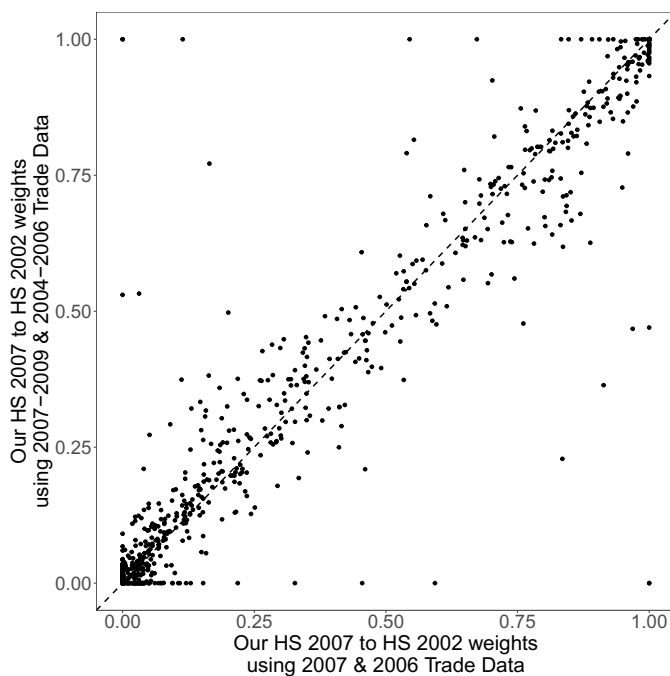
Figure 4.A.B.7: Conversion Weights for HS 2007 to HS 2002 with and without Autocorrelation Adjustment



Notes: This figure plots every not a priori determined weight for converting HS 2007 to HS 2002, once estimated while adjusting for product-specific trends (cf. Equation (4.2)) and once without trend adjustments (cf. Equation (4.1)).

Data source: [UN Comtrade Database](#).

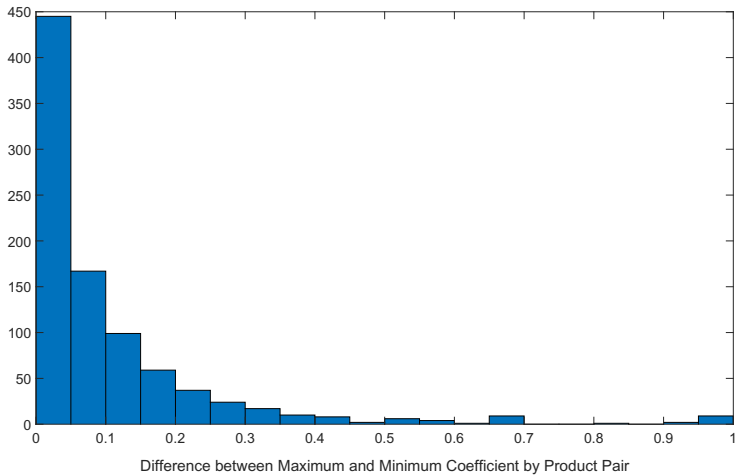
Figure 4.A.B.8: Conversion Weights for HS 2007 to HS 2002 using Averaged vs. Single-Year Trade Data



Notes: This figure plots every not a priori determined weight converting HS 2007 to HS 2002 when using trade data for only 2006 and 2007 against using averaged trade for the years 2004–2006 and 2007–2009.

Data source: [UN Comtrade Database](#).

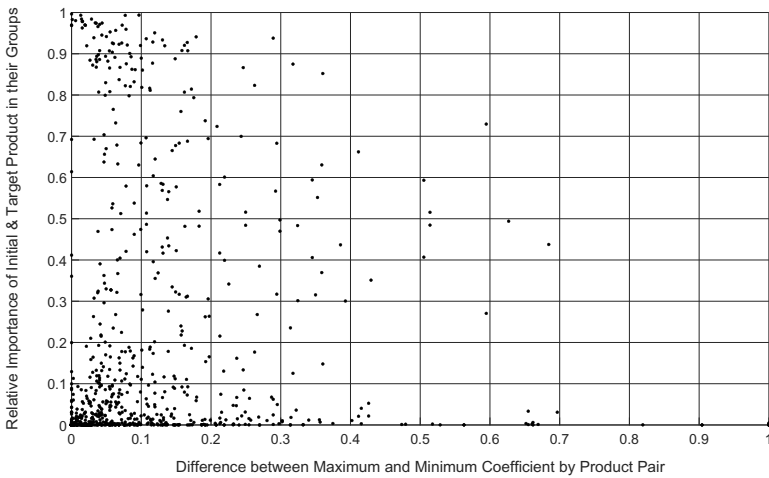
Figure 4.A.B.9: Sampling: Maximum Fluctuations of Conversion Weights



Notes: This figure plots the distribution of maximum differences in estimated conversion weights across bins (cf. Section 4.6). The initial year is 2007 and the target year is 2006 (i.e., converting HS 2007 to HS 2002). The data are randomly split into 20 blocks, and twenty regressions are computed, each leaving out one block. The difference we plot is the difference between the largest and smallest coefficient we get for each product pair. This figure excludes pre-specified weights (i.e., conversions known to be equal to zero or one), as these would, by definition, always yield a difference of zero.

Data source: [UN Comtrade Database](#).

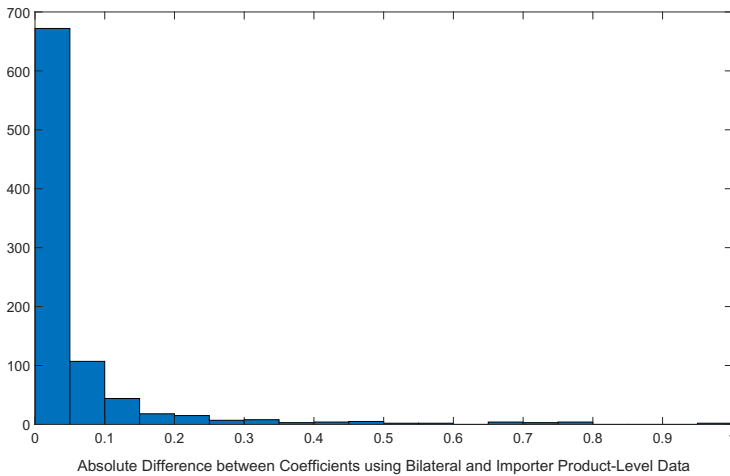
Figure 4.A.B.10: Sampling: Maximum Fluctuations of Weights and Relative Importance of Products



Notes: This figure plots the maximum difference in estimated conversion weights across bins against the relative importance of initial and target products in their groups (cf. Section 4.6). The initial year is 2007 and the target year is 2006 (i.e., converting HS 2007 to HS 2002). The data are randomly split into 20 blocks, and twenty regressions are computed, each leaving out one block. The difference we plot is the difference between the largest and smallest coefficient we get for each product pair. The relative importance is measured as the multiplication of the initial and target product trade shares within their groups. This figure excludes pre-specified weights (i.e., conversions known to be equal to zero or one), as these would, by definition, always yield a difference of zero.

Data source: [UN Comtrade Database](#).

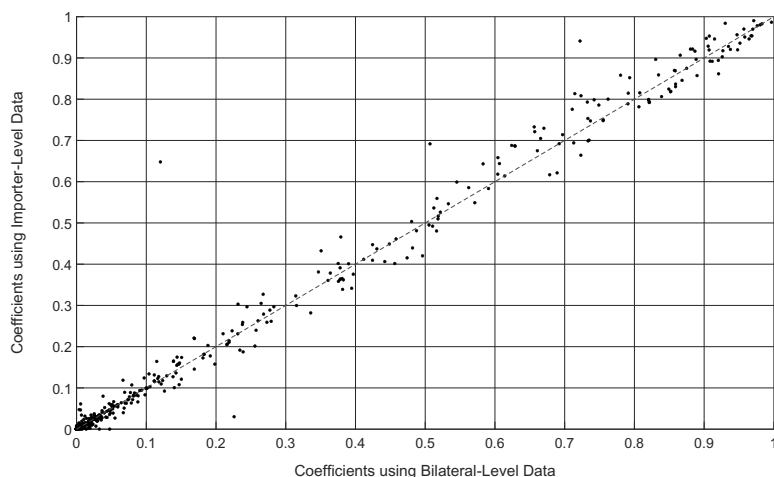
Figure 4.A.B.11: Aggregation: Absolute Differences in Coefficients Estimated Using Bilateral vs. Importer Product-Level Data



Notes: This figure plots the distribution of absolute differences between conversion weights estimated using bilateral and importer product-level data, respectively (cf. Section 4.6). The initial year is 2007 and the target year is 2006 (i.e., converting HS 2007 to HS 2002). This figure excludes pre-specified weights (i.e., conversions known to be equal to zero or one), as these would, by definition, always yield a difference of zero.

Data source: [UN Comtrade Database](#).

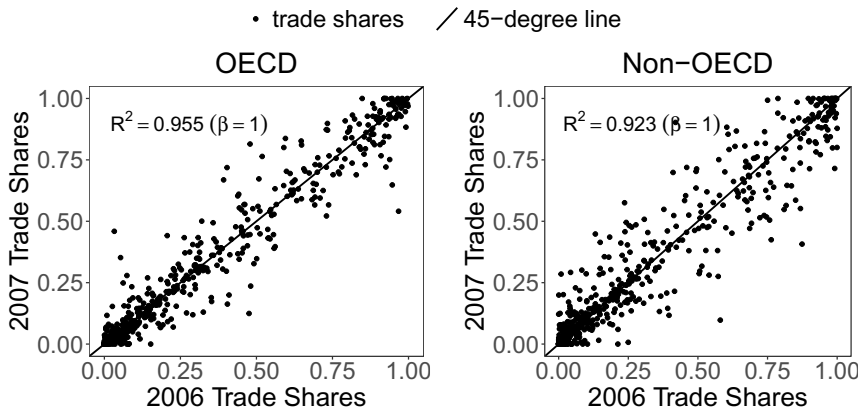
Figure 4.A.B.12: Aggregation: Coefficients Estimated Using Bilateral vs. Importer Product-Level Data Scaled by Relative Importance of Products



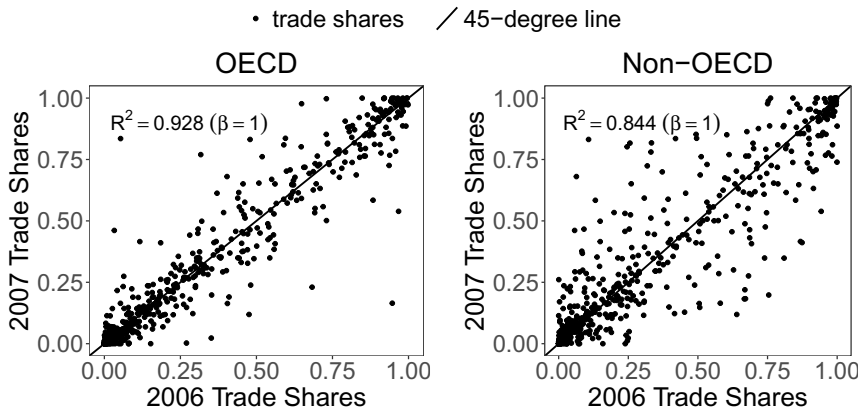
Notes: This figure plots the conversion weights estimated using bilateral and importer product-level data, respectively, scaled by the relative importance of initial and target products in their groups (cf. Section 4.6). The initial year is 2007 and the target year is 2006 (i.e., converting HS 2007 to HS 2002). The relative importance is measured as the multiplication of the initial and target product trade shares within their groups. The dashed line depicts the 45-degree line. This figure excludes pre-specified weights (i.e., conversions known to be equal to zero or one), as these would, by definition, always yield a difference of zero.

Data source: [UN Comtrade Database](#).

Figure 4.A.B.13: Product Trade Shares within HS Product Groups over Time (Our Conversion Weights – Country-Group-Specific vs. Global Weights)



(a) Country-Group-Specific Conversion Weights



(b) Global Conversion Weights

Notes: This figure plots for each product its trade share within its specific group across several years. We include only product groups which contain at least one 1:n or m:n assignment from HS 2007 to HS 2002, and trade data for 36 OECD (left-hand side) and 41 non-OECD (right-hand side) countries which used the latest available HS code classification in 2006 and 2007. The figure uses Comtrade data provided in the latest available HS vintage, and transforms 2007 data using our weights back into HS 2002. In Panel 4.A.B.13a we use OECD- and non-OECD-specific conversion weights, respectively, while Panel 4.A.B.13b uses the weights we calculated with the entire sample. Product groups and the algorithm to obtain our conversion weights are explained in the main text. The provided R-squared assumes a fit along the 45-degree line.

References

- ADÃO, R., C. ARKOLAKIS, AND S. GANAPATI (2020): “Aggregate Implications of Firm Heterogeneity: A Nonparametric Analysis of Monopolistic Competition Trade Models,” Working Paper 28081, NBER Working Paper Series.
- ALESSANDRIA, G. AND H. CHOI (2014): “Establishment Heterogeneity, Exporter Dynamics, and the Effects of Trade Liberalization,” *Journal of International Economics*, 94, 207–223.
- ALESSANDRIA, G., H. CHOI, AND K. RUHL (2014): “Trade Adjustment Dynamics and the Welfare Gains from Trade,” Working Paper 20663, NBER Working Paper Series.
- ALESSANDRIA, G. A., C. ARKOLAKIS, AND K. J. RUHL (2020): “Firm Dynamics and Trade,” *National Bureau of Economic Research*, Working Paper 27934.
- ALLEN, T. AND C. ARKOLAKIS (2020): “The Welfare Effects of Transportation Infrastructure Improvements,” NBER Working Paper Series, No. 25487.
- ALLEN, T. AND D. DONALDSON (2020): “Persistence and Path Dependence in the Spatial Economy,” Working Paper 28059, NBER Working Paper Series.
- ALVAREZ, F. AND R. E. J. LUCAS (2007): “General Equilibrium Analysis of the Eaton–Kortum Model of International Trade,” *Journal of Monetary Economics*, 54, 1726–1768.
- ANDERSON, J. E. AND E. VAN WINCOOP (2003): “Gravity with Gravitas: A Solution to the Border Puzzle,” *American Economic Review*, 93, 170–192.
- ARKOLAKIS, C. (2010): “Market Penetration Costs and the New Consumers Margin in International Trade,” *Journal of Political Economy*, 118, 1151–1199.
- ARKOLAKIS, C., A. COSTINOT, AND A. RODRÍGUEZ-CLARE (2012): “New Trade Models, Same Old Gains?” *American Economic Review*, 102, 94–130.
- ARKOLAKIS, C., S. GANAPATI, AND M.-A. MUENDLER (2021): “The Extensive Margin of Exporting Products: A Firm-Level Analysis,” *American Economic Journal: Macroeconomics*, 13, 182–245.
- ARKOLAKIS, C., N. RAMONDO, A. RODRÍGUEZ-CLARE, AND S. YEAPLE (2018): “Innovation and Production in the Global Economy,” *American Economic Review*, 108, 2128–73.
- ARMENTER, R. AND M. KOREN (2014): “A Balls-and-Bins Model of Trade,” *American Economic Review*, 104, 2127–51.
- ARPING, S. AND K. M. DIAW (2008): “Sunk Costs, Entry Deterrence, and Financial Constraints,” *International Journal of Industrial Organization*, 26, 490–501.
- ATKIN, D., A. COSTINOT, AND M. FUKUI (2021): “Globalization and the Ladder of Development: Pushed to the Top or Held at the Bottom?” Working Paper 29500, NBER Working Paper Series.
- BAIER, S. L., J. H. BERGSTRAND, AND M. FENG (2014): “Economic Integration Agreements and the Margins of International Trade,” *Journal of International Economics*, 93, 339–350.
- BALDWIN, R. (1988): “Hysteresis in Import Prices: The Beachhead Effect,” *The American Economic Review*, 78, 773–785, publisher: American Economic Association.
- BALDWIN, R. AND P. KRUGMAN (1989): “Persistent Trade Effects of Large Exchange Rate Shocks,” *The Quarterly Journal of Economics*, 104, 635–654, publisher: Oxford University Press.

- BOEHM, C. E., A. A. LEVCHENKO, AND N. PANDALAI-NAYAR (2020): “The Long and Short (Run) of Trade Elasticities,” Working Paper 27064, NBER Working Paper Series.
- BORCHERT, I., M. LARCH, S. SHIKHER, AND Y. V. YOTOV (2021): “The International Trade and Production Database for Estimation (ITPD-E),” *International Economics*, 166, 140–166.
- BRODA, C. AND D. E. WEINSTEIN (2006): “Globalization and the Gains From Variety,” *The Quarterly Journal of Economics*, 121, 541–585.
- BUERA, F. J. AND E. OBERFIELD (2020): “The Global Diffusion of Ideas,” *Econometrica*, 88, 83–114.
- BURSTEIN, A. AND M. J. MELITZ (2013): “Trade Liberalization and Firm Dynamics,” in *Advances in Economics and Econometrics Tenth World Congress. Applied Economics, Econometric Society Monographs.*, Cambridge, UK: Cambridge University Press, vol. 2, doi: 10.3386/w16960.
- CALIENDO, L., M. DVORKIN, AND F. PARRO (2019): “Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock,” *Econometrica*, 87, 741–835.
- CALIENDO, L. AND F. PARRO (2015): “Estimates of the Trade and Welfare Effects of NAFTA,” *The Review of Economic Studies*, 82, 1–44.
- CALIENDO, L., F. PARRO, L. D. OPROMOLLA, AND A. SFORZA (2021): “Goods and Factor Market Integration: A Quantitative Assessment of the EU Enlargement,” *Journal of Political Economy*, 129, 3491–3545.
- CARON, J., T. FALLY, AND J. R. MARKUSEN (2014): “International Trade Puzzles: A Solution Linking Production and Preferences,” *The Quarterly Journal of Economics*, 129, 1501–1552.
- CEBECI, T. (2012): “A Concordance among Harmonized System 1996, 2002 and 2007 Classifications,” *World Bank Working Papers*, 74576, <http://documents.worldbank.org/curated/en/914591468335500796/A-concordance-among-harmonized-system-1996-2002-and-2007-classifications>.
- CHANEY, T. (2005): “Productivity Overshooting: The Dynamic Impact of Trade Opening with Heterogeneous Firms,” Unpublished manuscript.
- (2008): “Distorted Gravity: The Intensive and Extensive Margins of International Trade,” *American Economic Review*, 98, 1707–1721.
- CHEN, N. (2004): “Intra-National versus International Trade in the European Union: Why Do National Borders Matter?” *Journal of international Economics*, 63, 93–118.
- CHOR, D. (2010): “Unpacking sources of comparative advantage: A quantitative approach,” *Journal of International Economics*, 82, 152–167.
- COSTANTINI, J. A. AND M. J. MELITZ (2007): “The Dynamics of Firm-Level Adjustment to Trade Liberalization,” in *The Organization of Firms in a Global Economy*, Cambridge, Ma: Harvard University Press.
- COSTINOT, A., D. DONALDSON, AND I. KOMUNJER (2012): “What Goods Do Countries Trade? A Quantitative Exploration of Ricardo’s Ideas,” *Review of Economic Studies*, 79, 581–608.
- COSTINOT, A. AND A. RODRÍGUEZ-CLARE (2013): “Online Appendix to Trade Theory with Numbers: Quantifying the Consequences of Globalization,” March. Available at: <https://economics.mit.edu/files/9215>.
- (2014): “Trade Theory with Numbers: Quantifying the Consequences of Globalization,” in *Handbook of International Economics*, Elsevier, vol. 4, 197–261.

- DAS, S., M. J. ROBERTS, AND J. R. TYBOUT (2007): “Market Entry Costs, Producer Heterogeneity, and Export Dynamics,” *Econometrica*, 75, 837–873.
- DEKLE, R., J. EATON, AND S. KORTUM (2007): “Unbalanced Trade,” *American Economic Review*, 97, 351–355.
- DIODATO, D. (2018): “A Network-based Method to Harmonize Data Classifications,” *Papers in Evolutionary Economic Geography*, No 18.43.
- DIXIT, A. (1989): “Hysteresis, Import Penetration, and Exchange Rate Pass-Through,” *The Quarterly Journal of Economics*, 104, 205–228, publisher: Oxford University Press.
- DUTT, P., I. MIHOV, AND T. VAN ZANDT (2013): “The Effect of WTO on the Extensive and the Intensive Margins of Trade,” *Journal of International Economics*, 91, 204–219.
- EATON, J., M. ESLAVA, C. J. KRIZAN, M. KUGLER, AND J. TYBOUT (2014): “A Search and Learning Model of Export Dynamics,” *mimeo*.
- EATON, J., M. ESLAVA, M. KUGLER, AND J. TYBOUT (2007): “Export Dynamics in Colombia: Firm-Level Evidence,” Working Paper 13531, NBER Working Paper Series.
- EATON, J. AND A. C. FIELER (2019): “The Margins of Trade,” Working Paper 26124, NBER Working Paper Series.
- EATON, J. AND S. KORTUM (2002): “Technology, Geography, and Trade,” *Econometrica*, 70, 1741–1779.
- (2005): “Technology in the Global Economy: A Framework for Quantitative Analysis,” *Unpublished manuscript*.
- (2012): “Putting Ricardo to Work,” *Journal of Economic Perspectives*, 26, 65–90.
- EATON, J., S. KORTUM, AND F. KRAMARZ (2011): “An Anatomy of International Trade: Evidence From French Firms,” *Econometrica*, 79, 1453–1498.
- EATON, J., S. KORTUM, B. NEIMAN, AND J. ROMALIS (2016): “Trade and the Global Recession,” *American Economic Review*, 106, 3401–38.
- EATON, J., S. KORTUM, AND S. SOTELO (2013): “International Trade: Linking Micro and Macro,” *Advances in Economics and Econometrics: Tenth World Congress*, 2, 329–370, Cambridge University Press.
- FAJGELBAUM, P. D., P. K. GOLDBERG, P. J. KENNEDY, AND A. K. KHANDELWAL (2020): “The Return to Protectionism,” *The Quarterly Journal of Economics*, 135, 1–55.
- FALLY, T. AND J. SAYRE (2018): “Commodity Trade Matters,” Working Paper 24965, NBER Working Paper Series.
- FELBERMAYR, G., F. TETI, AND E. YALCIN (2019): “Rules of Origin and the Profitability of Trade Deflection,” *Journal of International Economics*, 121, 103248.
- FIELER, A. C. (2011): “Nonhomotheticity and Bilateral Trade: Evidence and a Quantitative Explanation,” *Econometrica*, 79, 1069–1101.
- FITZGERALD, D., S. HALLER, AND Y. YEDID-LEVI (2016): “How Exporters Grow,” Working Paper 21935, NBER Working Paper Series.
- FORT, T. C., J. R. PIERCE, AND P. K. SCHOTT (2018): “New Perspectives on the Decline of U.S. Manufacturing Employment,” *Journal of Economic Perspectives*, 32, 47–72.
- FRENCH, S. (2016): “The Composition of Trade Flows and the Aggregate Effects of Trade Barriers,” *Journal of International Economics*, 98, 114–137.

- GANAPATI, S., W. F. WONG, AND O. ZIV (2020): “Entrepôt: Hubs, Scale, and Trade Costs,” CESifo Working Paper, No. 8199.
- GIRI, R., K.-M. YI, AND H. YILMAZKUDAY (2021): “Gains from Trade: Does Sectoral Heterogeneity Matter?” *Journal of International Economics*, 129, 103429.
- GOOLSBEE, A. AND C. SYVERSON (2008): “How Do Incumbents Respond to the Threat of Entry? Evidence from the Major Airlines,” *The Quarterly Journal of Economics*, 123, 1611–1633.
- HEAD, K. AND T. MAYER (2014): “Gravity Equations: Workhorse, Toolkit, and Cookbook,” in *Handbook of International Economics*, Elsevier, vol. 4, 131–195.
- HEAD, K., T. MAYER, AND J. RIES (2010): “The Erosion of Colonial Trade Linkages after Independence,” *Journal of International Economics*, 81, 1–14.
- HELPMAN, E., M. MELITZ, AND Y. RUBINSTEIN (2008): “Estimating Trade Flows: Trading Partners and Trading Volumes,” *The Quarterly Journal of Economics*, 123, 441–487.
- HEPENSTRICK, C. AND A. TARASOV (2015): “Per Capita Income and the Extensive Margin of Bilateral Trade,” *Canadian Journal of Economics*, 48, 1561–1599.
- HSIEH, C.-T., N. LI, R. OSSA, AND M.-J. YANG (2020): “Accounting for the New Gains from Trade Liberalization,” *Journal of International Economics*, 127, 103370.
- HUMMELS, D. AND P. J. KLENOW (2005): “The Variety and Quality of a Nation’s Exports,” *American Economic Review*, 95, 704–723.
- IMPULLITTI, G., A. A. IRARRAZABAL, AND L. D. OPROMOLLA (2013): “A Theory of Entry Into and Exit from Export Markets,” *Journal of International Economics*, 90, 75–90.
- JAIMOVICH, E. AND V. MERELLA (2015): “Love for Quality, Comparative Advantage, and Trade,” *Journal of International Economics*, 97, 376–391.
- KEHOE, T. J. AND K. J. RUHL (2013): “How Important is the New Goods Margin in International Trade?” *Journal of Political Economy*, 121, 358–392.
- KRISHNA, K., C. SALAMANCA, Y. SUZUKI, AND C. VOLPE MARTINCUS (2021): “Learning to Use Trade Agreements,” Working Paper 29319, NBER Working Paper Series.
- MANCHIN, M. AND A. M. PINNA (2009): “Border Effects in the Enlarged EU Area: Evidence from Imports to Accession Countries,” *Applied Economics*, 41, 1835–1854.
- MCCALLUM, J. (1995): “National Borders Matter: Canada-U.S. Regional Trade Patterns,” *The American Economic Review*, 85, 615–623.
- MELITZ, M. J. (2003): “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 71, 1695–1725.
- MELITZ, M. J. AND G. I. P. OTTAVIANO (2008): “Market Size, Trade, and Productivity,” *The Review of Economic Studies*, 75, 295–316.
- MELITZ, M. J. AND S. J. REDDING (2014): *Handbook of International Economics*, Elsevier, chap. Heterogeneous Firms and Trade, 1–54, 4 ed.
- MIAO, G. AND F. FORTANIER (2017): “Estimating Transport and Insurance Costs of International Trade,” Working Paper 2017/04, OECD Statistics Working Papers.
- MORALES, E., G. SHEU, AND A. ZAHLER (2019): “Extended Gravity,” *The Review of Economic Studies*, 86, 2668–2712.
- MOSSAY, P. AND T. TABUCHI (2015): “Preferential Trade Agreements Harm Third Countries,” *The Economic Journal*, 125, 1964–1985.

- NOH, Y.-H. AND G. MOSCHINI (2006): “Vertical Product Differentiation, Entry-Deterrence Strategies, and Entry Qualities,” *Review of Industrial Organization*, 29, 227–252.
- NOVY, D. (2013): “Gravity Redux: Measuring International Trade Costs with Panel Data,” *Economic Inquiry*, 51, 101–121.
- OSSA, R. (2015): “Why Trade Matters After All,” *Journal of International Economics*, 97, 266–277.
- PAGE, S. E. (2006): “Path Dependence,” *Quarterly Journal of Political Science*, 1, 87–115.
- PARRO, F. (2013): “Capital-Skill Complementarity and the Skill Premium in a Quantitative Model of Trade,” *American Economic Journal: Macroeconomics*, 5, 72–117.
- PIERCE, J. R. AND P. K. SCHOTT (2012): “Concording U.S. Harmonized System Codes over Time,” *Journal of Official Statistics*, 28, 53–68.
- PIVETEAU, P. AND G. SMAGGHUE (2019): “Estimating Firm Product Quality Using Trade Data,” *Journal of International Economics*, 118, 217–232.
- RAVIKUMAR, B., A. M. SANTACREU, AND M. SPOSI (2019): “Capital Accumulation and Dynamic Gains from Trade,” *Journal of International Economics*, 119, 93–110.
- REDDING, S. J. (2011): “Theories of Heterogeneous Firms and Trade,” *Annual Review of Economics*, 3, 77–105.
- ROBERTS, M. J. AND J. R. TYBOUT (1997): “The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs,” *The American Economic Review*, 87, 545–564, publisher: American Economic Association.
- RUHL, K. J. AND J. L. WILLIS (2017): “New Exporter Dynamics,” *International Economic Review*, 58, 703–726.
- SAMPSON, T. (2014): “Selection into Trade and Wage Inequality,” *American Economic Journal: Microeconomics*, 6, 157–202.
- SAMUELSON, P. A. (1954): “The Transfer Problem and Transport Costs, II: Analysis of Effects of Trade Impediments,” *The Economic Journal*, 64, 264–289.
- SANTOS SILVA, J. AND S. TENREYRO (2006): “The Log of Gravity,” *The Review of Economics and Statistics*, 88, 641–658.
- SCHIVARDI, F. AND E. VIVIANO (2011): “Entry Barriers in Retail Trade,” *The Economic Journal*, 121, 145–170.
- SIMONOVSKA, I. AND M. E. WAUGH (2014): “The Elasticity of Trade: Estimates and Evidence,” *Journal of International Economics*, 92, 34–50.
- SOTELO, S. (2020): “Domestic Trade Frictions and Agriculture,” *Journal of Political Economy*, 128, 2690–2738.
- TIMMER, M. P., E. DIETZENBACHER, B. LOS, R. STEHRER, AND G. J. DE VRIES (2015): “An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production,” *Review of International Economics*, 23, 575–605.
- TINBERGEN, J. (1962): *Shaping the World Economy. Suggestions for an International Economic Policy.*, New York, NY: Twentieth Century Fund.
- UNSD (2004): *International Merchandise Trade Statistics Compilers Manual*, New York: United Nations.

- (2017): “Correlation and Conversion Tables Used in UN Comtrade,” <https://unstats.un.org/unsd/trade/classifications/correlation/HS2017%20conversion%20to%20earlier%20HS%20versions%20and%20other%20classifications.pdf>, accessed on February 20, 2022.
- WAGNER, R. AND A. ZAHLER (2015): “New Exports from Emerging Markets: Do Followers Benefit from Pioneers?” *Journal of Development Economics*, 114, 203–223.
- WAUGH, M. E. (2010): “International Trade and Income Differences,” *American Economic Review*, 100, 2093–2124.
- WCO (2015): “Correlation Tables HS 2012–2017,” <http://www.wcoomd.org/en/topics/nomenclature/instrument-and-tools/hs-nomenclature-2017-edition/correlation-tables-hs-2012-to-2017.aspx>, accessed on February 20, 2020.
- WOLF, H. C. (2000): “Intranational Home Bias in Trade,” *Review of Economics and Statistics*, 82, 555–563.
- YI, K.-M. (2010): “Can Multistage Production Explain the Home Bias in Trade?” *American Economic Review*, 100, 364–93.

Curriculum Vitae

David Torun

born November 16, 1991 in Zurich (Switzerland)

EDUCATION

Ph.D. in Economics and Finance, *University of St.Gallen*, 2022

Swiss Program for Beginning Doctoral Students in Economics, *Study Center Gerzensee*, 2018

M.A. in Economics, *University of Zurich*, 2016

B.A. in Economics, *University of Zurich*, 2014

WORK EXPERIENCE

Research Assistant, SIAW-HSG, *University of St.Gallen*, since 2017

Research Assistant, *University of Zurich*, 2016–2017

TEACHING ASSISTANCE

Advanced Macroeconomics I: Growth and Innovation (Master's level), 2018–2021

International Macroeconomics (Master's level), 2019–2022

Makroökonomik I (Bachelor's level, in German), 2020

AWARDS

Best student of the year, *Study Center Gerzensee*, 2018