Quantifying the Extensive Margins of Trade and Production*

David Torun
Harvard Kennedy School
dtorun@hks.harvard.edu

This version: September 2024

First version: April 2024

preliminary

Link to the most recent version

Abstract

This paper explores the impact of changing trade barriers on the stability and formation of international trade relationships. It builds a general equilibrium model capturing that most countries have only a few active trading partners within disaggregated industries. In this framework, shocks impact trade flows not only at the intensive margin—as in standard gravity models—but also at the extensive margin; i.e., the set of trading partners responds endogenously to shocks. In turn, this allows for measuring a country pair's resilience to shocks and identifying alternatives to key trading partners. I develop a novel calibration strategy to fit data on industry-level bilateral trade flows and aggregate production. Counterfactual exercises suggest that accounting for the sparsity of trade flows magnifies welfare changes upon trade-cost shocks, particularly for lower-income countries, which usually have a small number of active trading partners. For instance, when simulating a 10% reduction in global trade barriers, incorporating extensive-margin variation increases the average welfare gain by 15% across the entire sample and by 30% for the bottom quartile of the income distribution.

Keywords: extensive margin of trade · trade liberalization and protectionism · alternative trading partners · non-homothetic production function · welfare

JEL Classification: F10 · F11 · F14 · F17

*I am thankful to Rodrigo Adão, Costas Arkolakis, Enea Baselgia, Martin Beraja, Timo Boppart, Jonathan Dingel, Peter Egger, Alessandro Ferrari, Reto Foellmi, Michael Gasiorek, Robert Gold, Charles Gottlieb, Alejandro Graziano, Kyle Handley, Christian Hepenstrick, Gaurav Khanna, Munseob Lee, Stefan Legge, Severin Lenhard, James Markusen, Phillip McCalman, Isabelle Mejean, Marc Muendler, Dennis Novy, Ralph Ossa, Mathieu Parenti, Michael Pfaffermayr, Frank Pisch, Leah Platt Boustan, Ariell Reshef, Ulrich Schetter, Peter Schott, Fernando Stipanicic, Silvana Tenreyro, Mathias Thoenig, Fabian Trottner, Arnau Valladares-Esteban, Eric Verhoogen, Gerald Willmann, Yuan Zi, Josef Zweimüller, and the participants to numerous conferences and seminars for helpful comments and discussions. I gratefully acknowledge financial support from the Swiss National Science Foundation under Grant 214326.

1 Introduction

After decades of continued international trade liberalization, recent global events like the US-China trade war, Brexit, the Covid-19 pandemic, and the Russian invasion of Ukraine keep testing the resilience of trade networks. Such disruptions and backlashes against globalization have sparked renewed interest in understanding the formation and stability of trade connections. While established trade models help explain the volume of existing trade relationships—the intensive margin of trade—, they offer limited insight into the formation of these relations—the extensive margin of trade. This paper seeks to contribute to our understanding of the latter while preserving the former.

The global events above have highlighted that countries engage in a limited number of active trade relations within a typical industry. Reflecting this, Table 1 documents substantial variation along the extensive margin: the median country exports in 72% of disaggregated industries and serves only nine out of 149 destinations (conditional on exporting).[1] About 89% of industry-level relationships are "trade zeros," indicating no trade flow between two countries. Quantitatively, this raises the question of how we can account for this sparsity of trade flows and what it implies for a country's resilience to shocks. It is crucial to identify a country's best replacement for a key supplier and model the consequences of the implied changes—think of Germany deciding where it will import more petroleum oil from in response to its sanctions on Russia. However, established quantitative trade models with multiple industries predict that each country exports to and imports from all countries in every industry.[2] Consequently, such models neglect the sparsity of trade flows, thereby abstracting from a potential non-linear impact of trade liberalizations or disruptions.

In this paper, I introduce endogenous (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 might respond to changes in trade barriers and technology along both the extensive and intensive margins. The set of active trading partners changes endogenously in response to shocks. In an application, I show that allowing for extensive-margin variation tends to amplify welfare changes upon trade-cost shocks. Therefore, capturing that countries typically have only a few active trade relations within an industry magnifies the implications of global shocks. This is especially

[1] These narrowly defined industries could also be called "product categories." For example, the 4-digit HS codes "0902" and "8701" report, respectively, trade in "tea" and "tractors." For simplicity, I will stick to the label "industries" throughout the paper.

[2] For example, 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, Donaldson, and Komunjer (2012)). The relation to the literature is discussed at the end of this section.

Table 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
No. 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 served by an exporter within an industry. 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 4-digit HS 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.

true for countries with a low to medium (per worker) income level, which tend to have the lowest number of active trade relations. To my knowledge, this is the first study to quantify the trade and welfare consequences of integrating the extensive margin into a trade model with many heterogeneous countries and industries.

Section 2 details this flexible framework. The environment is based on the multi-industry extension of the Eaton and Kortum (2002) model (henceforth, EK). It preserves the baseline model's ability to fit the intensive margin of trade, and much of its tractability. Households obtain utility from consuming a finite set of non-tradable final goods—the industry level—that are composed out of a continuum of intermediate varieties. These intermediate inputs are tradable and produced by perfectly competitive firms. I extend the canonical setup in two ways: with a non-homothetic final-goods-assembly function and an upper-bounded distribution of firms' productivities. 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.[3] The country-industry-specific truncated distribution translates into a (positive) minimum price that an exporting country can offer to a specific destination within an industry. Naturally, if an exporter's minimum price is above the importer'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 does not necessarily produce intermediate varieties

[3]This contrasts with standard CES aggregators, which would necessarily use all inputs in equilibrium.

in every industry.[4] 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 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 expenditure, 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.[5] 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. I call this expression for trade flows a "piecewise gravity equation," because it consists of elements looking like standard gravity equations. 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 3, I develop a novel calibration strategy to fit data on aggregate production and industry-level bilateral trade flows. The algorithm minimizes the relative distance between empirical and model-generated 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 4-digit HS industry level for the top 80 countries in terms of total production. The model is able to replicate several important empirical features. It reproduces the distribution of bilateral zeros: The main specification matches most empirical zeros (roughly 88%), and its intensive-margin fit is 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 industry-level 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 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-

[4] About 34% of exporter-importer dyads do not have a single positive flow across industries in the data used for Table 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 4-digit HS code "2709" records trade in "crude petroleum oil."

[5] The 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.

cost shocks. In the main counterfactual scenarios, I simulate both a drop and a rise in global trade costs of 10%. It is important to note that positive and negative shocks do not need to yield symmetric effects, because the presence of trade zeros yields a non-constant trade elasticity. The main insight from these exercises is twofold. First, the results reveal a significant reaction to changes in trade barriers along the extensive trade margin, from which standard analyses typically abstract. Industry-level bilateral zeros change by approximately two-thirds of the shock size. For instance, a 10% rise in trade barriers increases the number of inactive trade relations by about 6.2%. This is a sizable change, since the share of bilateral industry-level zeros in the baseline equilibrium exceeds 60%. The impact of these dissolved trade links in terms of trade lost along the intensive margin is modest but non-negligible. Moreover, after the shock, the extensive model predicts that in 86 (97) cases, a country starts (stops) producing in a given industry.

Second, I find that the welfare responses are generally larger in the extensive version than in the benchmark model, although the 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. Therefore, capturing that most countries have only a few active trade relations within a typical industry magnifies the implications of global shocks. For instance, the welfare gains from the liberalization exercise are 15% larger on average in the extensive model. However, there is substantial heterogeneity across countries' income levels: While the bottom quartile of the income distribution has, on average, a 30% larger gain, the top quartile's welfare gain is only 2.4% higher when allowing for extensive-margin variation. This is a result of the higher productivity upper bounds in richer economies, which are driven by the fact that they tend to serve more destinations within any given industry. The 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. I introduce a quantifiable GE model of international trade to study the macroeconomic consequences of changing trade barriers, with a novel examination of the extensive margins of trade and production across countries and industries. This contributes to the quantitative trade (or gravity) literature that builds on the seminal contributions by Eaton and Kortum (2002); Anderson and Van Wincoop (2003); Melitz (2003); Chaney (2008); Arkolakis, Costinot, and Rodríguez-Clare (2012). Other studies have highlighted the importance of accounting for industry-specific productivities (e.g., Costinot et al., 2012), input-output linkages (e.g., Caliendo and Parro, 2015), heterogeneous elasticities across industries (e.g., Ossa, 2015; Giri, Yi, and Yilmazkuday, 2021), dynamics (e.g., Eaton, Kortum, Neiman, and Romalis, 2016; Caliendo, Dvorkin, and Parro, 2019), and capital accumulation (e.g., Ravikumar, Santacreu, and Sposi, 2019)—see Costinot and Rodríguez-Clare (2014); Head and Mayer (2014); Caliendo and Parro (2022) for survey articles. As opposed to these frameworks, my model endogenizes the selection of partner countries within an industry both in the cross section and in response to changes in the trade environment.

Closely related in spirit, some modeling techniques to account for the large share of bilateral zeros in the data have been proposed: for instance, by introducing bounded productivity distributions (Helpman, Melitz, and Rubinstein, 2008; Chor, 2010),[6] a finite number of firms (Eaton, Kortum, and Sotelo, 2013), or a stochastic probability of not observing small shipments (Eaton and Fieler, 2019).[7] While these approaches showed innovative ways to explain zeros in trade flows, applying them to multi-industry frameworks and bringing them to the data encounters two major obstacles. First, typically, they are computationally expensive as they do not yield a closed-form solution for trade flows. Second, a full calibration requires industry-level production data, unavailable at narrow industry levels for most countries. The framework I present overcomes both of these challenges, providing an expression for the level of trade flows allowing for zeros, which is well-suited for counterfactual analysis, and replicating salient empirical features despite employing only readily available data.

The non-homothetic assembly function in my framework, which yields choke prices necessary to create the trade zeros, is theoretically equivalent to a lower-tier utility function. While homothetic preferences are the predominant assumption in the field, non-homothetic preferences help explain consumption patterns across countries (e.g., Matsuyama, 2000; Caron, Fally, and Markusen, 2014; Matsuyama, 2019), the distributional effects of trade (e.g., Matsuyama, 2000; Fajgelbaum and Khandelwal, 2016), the relation between trade and per capita income (e.g., Fieler, 2011; Hepenstrick and Tarasov, 2015; Matsuyama, 2019), and international price setting (e.g., Simonovska, 2015; Foellmi, Hepenstrick, and Zweimüller, 2018).

Moreover, the presence of trade zeros implies that the so-called trade elasticity—the elasticity of trade flows with respect to trade costs—is not constant, which contradicts

[6] In line with Helpman et al. (2008), Adão, Arkolakis, and Ganapati (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. Moreover, taking an entirely different approach, Armenter and Koren (2014) explain trade zeros using a statistical model that captures the sparsity of trade data.

[7] 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, which would explain why a country does not export or produce in that industry. In one specification, Fally and Sayre (2018) include a choke price above which a commodity is not used; this relates to my approach.

the predictions of standard frameworks. This relates to the literature highlighting heterogeneous effects of trade-cost changes and proposing different drivers of variable trade elasticities such as non-homothetic preferences (Novy, 2013; Carrère, Mrázová, and Neary, 2020; Chen and Novy, 2022), endogenous firm selection (Melitz and Redding, 2015; Feenstra, 2018; Adão et al., 2020), and incumbency effects (Egger, Foellmi, Schetter, and Torun, 2023). Varying trade elasticities can have sizable implications for the gains from trade liberalization and the losses from protectionism (e.g., Melitz and Redding, 2015; Egger et al., 2023), as emphasized by the quantification exercises in this paper.

Empirically, the extensive margin of trade and the frequent occurrence of trade zeros have been studied extensively. A seminal contribution by Santos Silva and Tenreyro (2006) proposed using the now-standard Poisson pseudo-maximum-likelihood (PPML) estimator, enabling the inclusion of zeros in the dependent variable—for a recent discussion on issues related to logs with zeros, see Chen and Roth (2024). Empirical analyses demonstrate that accounting for the extensive margin of trade helps explain the size of trade flows (Hummels and Klenow, 2005), trade growth (Kehoe and Ruhl, 2013; Dutt, Mihov, and Van Zandt, 2013), and the impact of free trade agreements (Baier, Bergstrand, and Feng, 2014; French and Zylkin, 2024). I add to empirical ex-post analyses by introducing a mechanism that generates (bilateral) trade zeros and by simulating counterfactual scenarios.

Finally, a large body of work studies how countries diversify along their development paths in terms of production or exports across products (see, e.g., Hausmann and Rodrik (2003); Imbs and Wacziarg (2003); Hidalgo, Klinger, Barabási, and Hausmann (2007); Cadot, Carrère, and Strauss-Kahn (2011); Atkin, Costinot, and Fukui (2021); Diodato, Hausmann, and Schetter (2022)). I contribute to this literature by providing a framework that predicts in which industries a country is likely to start (or stop) producing or exporting in response to changes in the trade environment or productivity.

2 Modeling the Extensive and Intensive Margins of International Trade and Production

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 canonical multi-industry version of Eaton and Kortum (2002) (henceforth, EK), as in Costinot et al. (2012). The key novelty is an upper bound of the productivity distribution and a non-homothetic final-goods aggregator. Together, these modifications can create zeros along several dimensions. A country 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. As a consequence, a country pair may not trade with each other at all, and a country need not produce varieties in every industry. A country's set of trading partners within an industry will adjust endogenously in response to exogenous shocks to the trade environment. Countries are indexed by $i, j = \{1, 2, \ldots, N\}$, while final goods are indexed by $k = \{1, 2, \ldots, K\}$. I use "industry" to refer to the final-consumption-goods level, k, while "varieties" refers to tradable intermediate inputs that compose the final goods. Households obtain utility from consuming the final goods, which they assemble using industry-specific tradable intermediate inputs.[8] They inelastically supply labor to competitive domestic intermediate producers that draw their productivities from an upper-bounded distribution. Below I explain each of these layers in detail.

2.1 Demand Side: 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.[9]

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 final goods

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

where $\sigma > 0$ is the elasticity of substitution, and ξ_j^k are exogenous demand shifters satisfying $\sum_{k=1}^{K} \xi_j^k = 1$.[10] Each final product k is composed of k-specific tradable varieties according to

$$q_j^k = \int_0^1 q_j^k(\omega^k) \mathrm{d}\omega^k, \tag{1}$$

where $q_j^k(\omega^k) \in \{0, 1\}$ is the quantity of intermediate ω^k used for final product k. Note that the assembly function is theoretically equivalent to a lower-tier (or sub-) utility function.

[8] In Section 2.4, I extend the model to allow for final goods to be composed of varieties from all industries.

[9] A 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.

[10] The derivation of trade flows in Section 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 quantification (cf. Section 3). Appendix A.4 characterizes the equilibrium with a general additively separable utility function.

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.[11] 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 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 to refer exclusively to the final product's quantity.[12]

The household maximizes its utility subject to its budget constraint

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

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

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

[11]Note that the standard CES assembly function is the other limiting case where there is only substitution of varieties along the intensive margin. The function in equation (1) has been employed in other studies to easily incorporate consumption zeros (see, e.g., Murphy, Shleifer, and Vishny, 1989; Matsuyama, 2000; Foellmi and Zweimüller, 2006). In Appendix A.3, I derive an equilibrium with an alternative assembly function that yields a choke price while allowing for intensive-margin adjustments of $q_j^k(\omega^k)$. I prefer equation (1) for the baseline specification because this is a straightforward way to introduce choke prices while maintaining tractability. Moreover, in EK-type models, the equilibrium solely characterizes total expenditures on varieties from a given location rather than expenditures on a specific variety.

[12]For Ricardian models with explicit quality differentiation, see, e.g., Jaimovich and Merella (2015); Eaton and Fieler (2019).

where λ_j is the Lagrange multiplier.[13] 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

$$P_j^k := \frac{1}{q_j^k} \int_0^{\bar{\omega}_j^k} p_j^k(\omega^k) \mathrm{d}\omega^k. \tag{2}$$

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

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

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

$$P_j^k q_j^k = E_j \frac{\xi_j^k \left(\bar{p}_j^k\right)^{-\sigma} P_j^k}{\sum_{k=1}^{K} \xi_j^k \left(\bar{p}_j^k\right)^{-\sigma} P_j^k}. \tag{3}$$

These expenditure shares pin down the choke prices in equilibrium, as outlined in Section 2.3.

Welfare. Welfare changes will be measured as consumption-equivalent variation. Let q_j^k $\left(\tilde{q}_j^k\right)$ denote final goods consumption in the initial (counterfactual) equilibrium. The consumption-equivalent variation, ζ, is calculated as follows

$$\frac{\sigma}{\sigma-1} \sum_{k=1}^{K} \left(\xi_j^k\right)^{\frac{1}{\sigma}} \left((1+\zeta) q_j^k\right)^{\frac{\sigma-1}{\sigma}} = \frac{\sigma}{\sigma-1} \sum_{k=1}^{K} \left(\xi_j^k\right)^{\frac{1}{\sigma}} \left(\tilde{q}_j^k\right)^{\frac{\sigma-1}{\sigma}}$$

$$\Rightarrow \zeta = \left(\frac{\sum_{k=1}^{K} \left(\xi_j^k\right)^{\frac{1}{\sigma}} \left(\tilde{q}_j^k\right)^{\frac{\sigma-1}{\sigma}}}{\sum_{k=1}^{K} \left(\xi_j^k\right)^{\frac{1}{\sigma}} \left(q_j^k\right)^{\frac{\sigma-1}{\sigma}}}\right)^{\frac{\sigma}{\sigma-1}} - 1,$$

i.e., the relative change in consumption that would make the household indifferent between the initial equilibrium and a counterfactual scenario.

[13] Specifically, the Lagrange function is given by

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

Since in equilibrium there will be an increasing continuum of intermediate prices (cf. Section 2.2), there is a single crossing point between marginal utility and "marginal" prices.

The next section shows that the direct mapping above between average prices and total expenditures helps 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, as discussed at the end of Section 2.2.

2.2 Supply Side: Production of Tradable Intermediate Inputs

Perfectly competitive firms produce and ship tradable intermediate varieties using labor as the only input in production.

Trade costs. Throughout the paper, i indexes exporters, while j indexes importers (i.e., a subscript ij can be read as "from country i to j"). 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. I follow the literature normalizing $\tau_{jj}^k = 1$ and assuming that the triangle inequality holds, such that $\tau_{ij}^k \leq \tau_{ih}^k \tau_{hj}^k \; \forall i, h, j$.

Technology. Labor is 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 $\phi_i^k(\omega^k)$. Hence, the cost of producing one unit of ω^k and shipping it to country j is equal to the marginal cost

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

where w_i is the wage rate in country i.

Following the literature, I assume that $\phi_i^k(\omega^k)$ is independently drawn from a Fréchet distribution.[14] However, the distribution is truncated from above, that is,

$$F_i^k(\phi) = \frac{\exp\left(-T_i^k \phi^{-\theta}\right)}{\exp\left(-T_i^k (\bar{\phi})^{-\theta}\right)}, \qquad \text{for } \phi \in (0, \bar{\phi}], \tag{4}$$

where $\bar{\phi}$ is the productivity upper bound, T_i^k captures the general state of technology in country i to produce the varieties of industry k (a higher T_i^k increases the average

[14] I 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).

productivity in an i-k cell), and θ is inversely related to the dispersion of productivities within a country-industry cell, thereby governing the comparative advantage across countries. The truncation, together with the assembly function in equation (1), is necessary to create (bilateral) industry-level zeros, as will become clear below. The productivity upper bound can be interpreted as the technological frontier. The entire derivation below would be analogous with country-industry-specific technological frontiers, $\bar{\phi}_i^k$, and industry-specific shape parameters, θ^k.

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

$$G_{ij}^k(p) = \begin{cases} 1 - \dfrac{\exp\left[-T_i^k\left(w_i\tau_{ij}^k\right)^{-\theta}p^\theta\right]}{\exp\left(-T_i^k(\bar{\phi})^{-\theta}\right)} & \text{for } p \in \left[\dfrac{w_i\tau_{ij}^k}{\bar{\phi}}, \infty\right) \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

The corresponding probability density function, $g_{ij}^k(p)$, is given by $\mathrm{d}G_{ij}^k(p)/\mathrm{d}p$.[15]

With the distributions above, we can derive the overall price distribution in country j's industry k. Let $\underline{p}_{ij}^k := w_i\tau_{ij}^k/\bar{\phi}$, 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

$$\underline{p}_{1j}^k \leq \underline{p}_{2j}^k \leq \cdots \leq \underline{p}_{Nj}^k, \quad (6)$$

which comes in handy in the derivations below and is without loss of generality. To economize on notation, the rest of this section suppresses the industry superscript 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 inputs from the cheapest available producer. Recall that there is no product differentiation across countries in this model. 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 (following expres-

[15] There is a crucial difference between this and the well-known benchmark case where $\bar{\phi} \to \infty$. 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 (in expectation) offer prices that less competitive source countries are unable to provide.

sion (6)), we obtain

$$G_j(p) = \begin{cases} 1 - y_j^{1:n} \exp\left(-p^\theta \Phi_j^{1:n}\right) & \text{for } p \in \left[\underline{p}_{nj}, \underline{p}_{(n+1)j}\right) \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

where

$$y_j^{1:n} := \exp\left(\sum_{i=1}^{n} T_i \phi^{-\theta}\right)$$

$$\Phi_j^{1:n} := \sum_{i=1}^{n} T_i \left(w_i \tau_{ij}\right)^{-\theta}$$

This piecewise definition of $G_j(p)$ arises from the exporter-specific lower bounds of the respective bilateral price distributions.[16] Note that $\Phi_j^{1:n}$ captures the overall state of technology and trade barriers across producers i for a given destination j. The higher $\Phi_j^{1:n}$ (i.e., the better connected j is to its suppliers, and the more productive these producers are), the lower the average prices in country j. However, within a specific price interval $p \in \left[\underline{p}_{nj}, \underline{p}_{(n+1)j}\right)$, 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 segment. As I show next, this property of the price distribution, coupled with the choke prices from Section 2.1, can result in bilateral industry-level trade zeros. The mechanism is as intuitive as it is simple: If the minimum price that country i can offer to j, $\underline{p}_{ij}$, is above country j's maximum willingness to pay for varieties, $\bar{p}_j$, there will be no exports from i to j (in a given industry).

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, according to Section 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 A.1, the mass of varieties per capita that i exports to j equals

$$m_{ij} = T_i \left(w_i \tau_{ij}\right)^{-\theta} \sum_{n=i}^{N} \frac{y_j^{1:n}}{\Phi_j^{1:n}} \left[\exp\left(-\min\{\underline{p}_{nj}, \bar{p}_j\}^\theta \Phi_j^{1:n}\right) - \exp\left(-\min\{\underline{p}_{(n+1)j}, \bar{p}_j\}^\theta \Phi_j^{1:n}\right)\right], \tag{8}$$

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

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

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

where $y_j^{1:n}$ and $\Phi_j^{1:n}$ are as defined in equation (7), and exporters are sorted according to inequalities (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 intermediate inputs 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 and distribution cost. Hence, $\bar{p}_j < \infty$ together with $\phi < \infty$ can lead to (i) country i not selling to destination j, and (ii) country i not exporting or importing any variety within an 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,

$$m_j = G_j(\bar{p}_j). \tag{9}$$

Note that $\bar{p}_j \to \infty$ 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 (2), to obtain expenditures we need to integrate over the prices of employed varieties. In Appendix A.2, I show that within each industry this results in per capita trade flows from exporter i to importer j equal to

$$x_{ij} = T_i \left(w_i \tau_{ij}\right)^{-\theta} \sum_{n=i}^{N} \frac{y_j^{1:n}}{\Phi_j^{1:n}} \left(\Phi_j^{1:n}\right)^{-1/\theta} \bar{\Gamma}\left(\frac{\theta+1}{\theta}, \min\{\underline{p}_{nj}^{\theta}, \bar{p}_j^{\theta}\}\Phi_j^{1:n}, \min\{\underline{p}_{(n+1)j}^{\theta}, \bar{p}_j^{\theta}\}\Phi_j^{1:n}\right), \tag{10}$$

where $\bar{\Gamma}(s, \underline{x}, \bar{x}) := \int_{\underline{x}}^{\bar{x}} x^{s-1} \exp(-x) \mathrm{d}x$.[17] By symmetry of the households, total expenditures on intermediates from exporter i are given by $X_{ij} = L_j x_{ij}$. Note that $\bar{\Gamma}\left(\frac{\theta+1}{\theta}, p, p\right) = 0$, showing that equation (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 (10) highlights that, loosely speaking, ϕ predominantly affects the *extensive* margin of trade (via

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

$\underline{p}_{ij} = w_i\tau_{ij}/\bar{\varphi}$), while T_i governs the *intensive* margin.[18] 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 for narrow industry definitions. Third, the model generates something akin to a gravity equation with a crucial difference. Although the terms within the sum are the same across exporters, 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 3. Equation (10) represents a "piecewise" gravity equation—more on that below. Finally, country j's (industry-level) per capita expenditures on intermediate varieties are given by $x_j = \sum_{i=1}^{N} x_{ij}$, yielding

$$x_j = \sum_{n=1}^{N} \gamma_j^{1:n} \left(\Phi_j^{1:n}\right)^{-1/\theta} \Gamma\left(\frac{\theta+1}{\theta}, \min\{\underline{p}_{nj}^{\theta}, \bar{p}_j^{\theta}\}\Phi_j^{1:n}, \min\{\underline{p}_{(n+1)j}^{\theta}, \bar{p}_j^{\theta}\}\Phi_j^{1:n}\right), \tag{11}$$

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

Comparison to standard gravity. It is instructive to compare the expressions above to the results of EK. Denote the components of X_j by

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

such that $X_j = \sum_{n=1}^{N} X_j^{(n)}$. Then we can rewrite equation (10) as

$$X_{ij} = \sum_{n=i}^{N} \frac{T_i \left(w_i\tau_{ij}\right)^{-\theta}}{\sum_{i'=1}^{n} T_{i'} \left(w_{i'}\tau_{i'j}\right)^{-\theta}} X_j^{(n)}, \tag{12}$$

i.e., within each expenditure segment $X_j^{(n)}$ the expenditure share on intermediates from country i is determined by its relative value of $T_i\left(w_i\tau_{ij}\right)^{-\theta}$ (for exporters $i \leq n$). Therefore, we obtain a "piecewise" gravity equation. To see the connection to the standard gravity equation, suppose the productivity distribution was unbounded (i.e., $\bar{\varphi} \to \infty$). Equation (10) then reduces to

$$X_{ij}\Big|_{\bar{\varphi}\to\infty} = \frac{T_i \left(w_i\tau_{ij}\right)^{-\theta}}{\sum_{i'=1}^{N} T_{i'} \left(w_{i'}\tau_{i'j}\right)^{-\theta}} X_j,$$

[18] For 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,

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

which, after dividing by total expenditure X_j, is equal to EK's well-known expression for trade shares. This can be seen directly from the definition of $X_j^{(n)}$ above and the fact that $\underline{p}_{nj} \to 0 \ \forall n$ as $\phi \to \infty$.

Therefore, in terms of trade flows, my model nests the canonical EK version as a limiting 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 also holds if we have a finite choke price $\bar{p}_j < \infty$.[19] Accordingly, with $\phi \to \infty$, I obtain the same formula for trade shares when replacing the "0-1" assembly function in equation (1) with a standard CES aggregator (see also the derivations in Appendix A.3).

What happens if instead every variety had to be used in the final-goods-assembly process, i.e., if we let $\bar{p}_j \to \infty$? Equation (10) then becomes

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

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.

2.3 Equilibrium

In equilibrium, trade is balanced, that is,

$$\sum_{i \neq j} \sum_{k=1}^{K} X_{ij}^k - X_{ji}^k = 0, \tag{13}$$

which pins down the set of wages.[20]

Choke prices. To characterize the equilibrium in this economy, it is key to identify the choke prices of intermediate inputs, $\bar{p}_j^k$ (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 $Q_j^k = M_j^k = G_j^k(\bar{p}_j^k) L_j$ (equation (9)). I use $P_j^k = X_j^k / Q_j^k$ to substitute for the final good's average price—or,

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

[20]Therefore, I do not allow for trade deficits—a prevalent feature of the data—, as is frequently done 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, Eaton, and Kortum, 2007). 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.

the average 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 (3), we obtain

$$\frac{w_j L_j}{G_j^k\left(\bar{p}_j^k\right)} = \left(\xi_j^k\right)^{-1} \left(\bar{p}_j^k\right)^{\sigma} \left[\sum_{k=1}^{K} \xi_j^k \left(\bar{p}_j^k\right)^{-\sigma} \frac{X_j^k}{G_j^k\left(\bar{p}_j^k\right)}\right]. \tag{14}$$

For a given set of wages and parameters, equation (14) 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 (11). Note that $w_j L_j = \sum_{k=1}^{K} X_j^k$.

Set of equations. To sum up, for given parameters θ, σ, $\{L_i\}_i$, $\{\xi_i^k\}_{i,k}$, $\bar{\phi}$, $\{T_i^k\}_{i,k}$, $\{\tau_{ij}^k\}_{i,j,k}$, the equilibrium in this economy is determined by a set of wages and choke prices, $\{w_i\}_i$, $\{\bar{p}_i^k\}_{i,k}$, which satisfy the conditions in equations (9)–(11), (13), and (14). Section 3 details the calibration strategy.

2.4 Extensions

Before turning to the calibration strategy, I discuss possible modifications and extensions of the model. First, I alter the final-goods-assembly function. Second, I outline how to introduce consumption zeros. Third, I allow for final products to be composed out of varieties from all industries.

Alternative assembly function. To emphasize that the main theoretical results above do not hinge on the specific functional form in equation (1), Appendix A.3 derives an equilibrium with an alternative assembly function. This alternative specification yields a choke price while allowing for intensive-margin adjustments in the quantity used of a variety ($q_j^k(\omega^k)$). I show that this also results in a piecewise gravity equation like equation (12), with only the expression for total expenditures changing. Accordingly, the gravity equation reduces to the standard EK expression once the productivity distribution is unbounded (i.e., $\bar{\phi} \to \infty$), analogous to the above discussion.

Consumption zeros. In Appendix A.4, I derive the equilibrium with a general additively separable utility function. The major difference from the specification above is that preferences may be such that the marginal utility at zero is finite. This, in turn, implies that a household does not necessarily consume every final product k 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 A.4.

Sectoral linkages. In Appendix A.5, I outline an easy way to allow for sectoral 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 in the main text.

3 Bringing the Model to the Data

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

3.1 Data and Solution Algorithm

The calibration strategy follows naturally from the equilibrium conditions in Section 2.3. To capture the extensive margin in the data, I allow for country-industry-specific productivity upper bounds, $\bar{\phi}_i^k$. 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 set the technology and trade-cost parameters such that the model mimics the data, by reducing the relative differences between empirical and model-implied trade flows.[21]

Data. To calibrate the model, I use bilateral goods trade data at the 4-digit HS level from the Atlas of Economic Complexity, which is a cleaned version of the widely-used UN Comtrade data. 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

[21]The 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.

code, 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 $i \neq j$.[22]

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, Larch, Shikher, and Yotov (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 4.2. I set $\theta = 4.14$, the well-known preferred estimate from Simonovska and Waugh (2014), and $\sigma = 2$, a common value in macroeconomic models.[23] The remaining parameters and endogenous variables are calibrated applying the algorithm below.

Solving the equilibrium. For a given set of parameters, θ, σ, $\{L_i\}_i$, $\{\xi_i^k\}_{i,k}$, $\{\bar{\phi}_i^k\}_{i,k}$, $\{T_i^k\}_{i,k}$, $\{\tau_{ij}^k\}_{i,j,k}$, 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., $\boldsymbol{T} := \{T_i^k\}_{i,k}$. Regarding trade costs, let $\boldsymbol{\tau} := \{\tau_{ij}^k\}_{i,j,k}$, $\boldsymbol{\tau}^{\mathbf{1}} := \{\tau_{ij}\}_{i,j}$, and $\boldsymbol{\tau}^{\mathbf{2}} := \{\tau_j^k\}_{j,k}$. A crucial benefit of the conditions in Section 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.

[22] Recall that on the left-hand side of equation (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 my framework, 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. Their approach is unfeasible in my model due to the extensive-margin variation.

[23] There is a vast literature that delivers estimates of related elasticities (see, e.g., Broda and Weinstein, 2006; Simonovska and Waugh, 2014; Caliendo and Parro, 2015; Giri et al., 2021). 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 4.2. I, therefore, refrain from separately identifying these elasticities.

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

A.2 Use equation (14) to pin down $\boldsymbol{p}$.

A.3 Verify equation (13). If this condition is satisfied, an equilibrium is found. Otherwise, update $\boldsymbol{w}$ and return to step A.2.

Step A.1 requires no further explanation.

Step A.2 pins down intermediate choke prices, $\boldsymbol{p}$, using equation (14). More precisely, I compute

$$\frac{w_j L_j}{G_j^k\left(\bar{p}_j^k\right)} - \left(\xi_j^k\right)^{-1}\left(\bar{p}_j^k\right)^{\sigma}\left[\sum_{k=1}^{K} \xi_j^k \left(\bar{p}_j^k\right)^{-\sigma} \frac{X_j^k}{G_j^k\left(\bar{p}_j^k\right)}\right], \qquad (14')$$

where trade flows are calculated according to equation (10). The first element of equation (14′) 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 being capped at $\bar{p}_j^k$. Therefore, the right-hand side of equation (14) is weakly increasing in $\bar{p}_j^k$, ceteris paribus. Accordingly, to satisfy equation (14), I increase (decrease) $\bar{p}_j^k$ if equation (14′) is positive (negative). The relative size of this change is determined by the expression in equation (14′) divided by $w_j L_j / G_j^k(\bar{p}_j^k)$. Hence, larger deviations from equation (14) lead to larger changes in $\bar{p}_j^k$. This procedure mirrors the tâtonnement algorithm by Alvarez and Lucas (2007).

Step A.3 pins down the set of wages, $\boldsymbol{w}$, using equation (13). 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, $\boldsymbol{w}$ 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. The relative size of this change is determined by the expression above. Next, I outline how the technology parameters, ′ and $\boldsymbol{T}$, and trade costs, , are calibrated.

How to fit the data. The main parameters of the model are calibrated via an algorithm consisting of three steps.[24] The steps are first listed for ease of reference and then explained in more detail below.

[24]Recall that θ, σ, $\boldsymbol{L}$, , and $\boldsymbol{w}$ 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 (13).

B.1 Guess a set of productivity upper bounds $'$, technology shifters $\boldsymbol{T}$, and trade cost components $^{\boldsymbol{1}}$ and $^{\boldsymbol{2}}$.

B.2 Use equation (14) to pin down $\boldsymbol{p}$.

B.3 If empirical trade flows are matched well, stop. Otherwise, update $'$, $\boldsymbol{T}$, $^{\boldsymbol{1}}$, $^{\boldsymbol{2}}$, and return to step B.2.

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 $'$, $\boldsymbol{T}$, $^{\boldsymbol{1}}$, and $^{\boldsymbol{2}}$, while holding the remaining parameters, as well as wages and choke prices constant. Bilateral industry-level trade flows are calculated using equation (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 $^{\boldsymbol{1}}$, $^{\boldsymbol{2}}$, and $\boldsymbol{T}$, the extensive margin determines $'$. To update ϕ_i^k, I use the extensive margin of an exporter i in industry k across importers j. An increase in ϕ_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 (10)). Therefore, I increase (decrease) ϕ_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. Note the resemblance to an exporter-industry fixed effect in a regression, which would capture exporter i's relative size in this industry. The calibration of $^{\boldsymbol{2}}$ 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 $^{\boldsymbol{1}}$, 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.[25]

[25] Recall 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 $^{\boldsymbol{1}}$, $^{\boldsymbol{2}}$, $\boldsymbol{T}$, and *, 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.

3.2 The Model's Performance in Matching the Data

Before targeting real data, I have tested the algorithm using simulated data. First, I randomly generated the model parameters. Then, I 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 ^{1}, ^{2}, T, and $'$ 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 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. 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. While the industry-level export zeros are somewhat underestimated, the import zeros are slightly overestimated by the calibration. There is substantial variation at the bilateral extensive margin, which is matched very well. Recall that the algorithm 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 and the model is around 63%. The model matches roughly 88% of the empirical zeros, which means that around 12% of the model's zeros are "false negatives." It is worth emphasizing that the number of calibrated parameters is around 300,000, which corresponds to less than 4% of the number of observations.[26]

To create a reference point, I re-calibrated the model using steps B.1–B.3 but imposing $\phi_i^k \to \infty \; \forall i, k$. 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 2.2). I label the version of the model with flexible productivity upper bounds—and, thus, extensive-margin variation—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 fall below a fixed threshold.[27] Therefore, in Table C.1 of Appendix C, I reproduce Table 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 42% of the benchmark model's trade flows are equal to zero when imposing the cutoff, of which 94% are empirically correct. For completeness, Table C.2 uses this cutoff in the extensive model, which produces results

[26] Alternatively, one can perfectly fit the data by letting absorb the residual variation. I opted for the specification above because it is more transparent. In Section 4.2, I perform a robustness check that incorporates the residuals into the trade-cost function.

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

Table 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
No. of Observations	7,836,800				
Panel B: Model	Min.	Q1	Median	Q3	Max.
Sh. Exported Industries	0.72	0.95	0.98	0.99	1
Sh. Served Destinations[(i)]	0.01	0.09	0.28	0.66	1
Sh. Imported Industries	0.88	0.98	0.98	1	1
Sh. Source Countries[(ii)]	0.01	0.16	0.34	0.56	1
No. 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 across countries of industries with positive exports, while the second row reports the share of destinations served by an exporter within an industry. 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 4-digit HS codes. The sample consists of 1,240 industries and 80 countries. See Section 3.2 for details. Section 3.1 describes the data.
[(i)] Conditional on exporting in a given industry.
[(ii)] Conditional on importing in a given industry.
Data source. See Section 3.1.

similar to those presented in Table 2.

In terms of the intensive margin, both versions of the model perform in a similar way. To compare the two, I computed the sum of absolute deviations between the empirical trade flows and those generated by the models. There is a slight increase in this measure of roughly 1.8% when I move from the extensive to the benchmark model. A better fit of the extensive version is unsurprising, given that it entails a larger number of parameters being calibrated.[28] Moreover, the benchmark model yields higher trade costs on average. This is reflected in Table C.4 of Appendix C, where standard "gravity variables" have a larger impact on log trade costs from the benchmark calibration (columns (3) and (4)). For instance, the presence of a regional trade agreement between two countries is associated with about 10% lower trade barriers in the extensive model but 25% lower ones in the benchmark version. Therefore, the model pushes more of the variation across

[28]As 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 C.3 reproduces Table 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, among other things, does not impose a specific structure on the destination-specific component or restrictions on , the extensive framework appears to perform reasonably well.

countries into trade costs when it cannot explicitly capture the sparsity of trade flows.

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.

4 Counterfactual Responses to Global Shocks

This section studies changes in trade flows, production, and welfare that arise from both 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 3.1.[29]

4.1 Main Specification

The two main counterfactual exercises increase and decrease, respectively, global bilateral trade costs by 10%. Universal trade-cost shocks are motivated by the long wave of globalization that preceded the recent surge in global disruptions and protectionist measures. The shock size corresponds to the main simulation in Eaton et al. (2013). I leave domestic trade barriers unchanged and do not allow trade costs to fall below one. To put the welfare changes into perspective, I compare them with those predicted by the benchmark model following shocks of the same magnitude as in the extensive model. Additionally, the extensive model allows me to analyze how bilateral zeros and, hence, the composition of trading partners change in response to trade-cost shocks.

Global liberalization. A drop in global trade barriers should analogous to the workhorse models positively impact trade flows and welfare. The latter is confirmed by Figure 1a, which depicts relative welfare changes after a 10% drop in trade costs. Welfare increases relative to the baseline equilibrium for all countries, with growth rates ranging from 0.5% to 7.2% and averaging 2.5% in the extensive model.[30]

The welfare responses implied by the two versions of the model are highly correlated, while, more interestingly, the extensive model predicts larger gains than the benchmark version for most countries. This is intuitive, since low productivity upper bounds broadly

[29] Before computing the counterfactual equilibria, I first solve for the equilibrium using the calibrated parameters. This creates a starting point that features balanced trade, which removes potentially misleading conclusions from closing trade imbalances. Moreover, to ensure that the initial equilibria (and welfare changes) in the extensive and benchmark models are comparable, I re-calibrated the benchmark model (i.e., the version without extensive-margin variation) targeting the trade flows from the extensive model. This is necessary because when each version is calibrated separately using the empirical trade flows, they predict different domestic trade flows and, consequently, different welfare levels.

[30] Recall that welfare changes are not computed as real income changes, but measured as consumption-equivalent variation (cf. Section 2.1). Importantly, preferences are exactly the same in the extensive and the benchmark model. Moreover, the 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: Welfare Changes after a 10% Drop in Global Trade Costs

(a) Welfare Changes

(b) Welfare Gains vs. Income

Notes. This gure depicts the welfare changes after a drop in global trade costs of 10%. The productivity upper bound is exible in the extensive model, and in nite in the benchmark model (eliminating extensive-margin variation). The solid line in Panel 1a is a 45-degree line. Panel 1b relates the (log) initial production level, $\log(X_j)$, to the di erence between the two models' welfare gains (i.e., the vertical minus the horizontal axis in Panel 1a). The results are discussed in Section 4.
Data source: See Section 3.1.

capture industry-level constraints in production, which translates into a larger bene t from trade liberalization. These limits in productivity matter especially for lower-income countries. For instance, regressing countries' log production on the log of their across-

Table 3: Bilateral Zeros after Changes in Global Trade Costs

%-Change in	%-Change in Bil. Zeros Rel. to Baseline	Contribution of Zeros to Trade Change
5%	3:5%	1.05%
10%	7%	1.9%
15%	10:5%	2.6%
+5%	3:3%	0.84%
+10%	6:2%	1.56%
+15%	8:8%	2.13%

Notes. This table reports how bilateral industry-level trade zeros change upon a shock to trade barriers. The rst column shows the percentage change in the number of zeros relative to the baseline equilibrium across di erent shock sizes. For a trade-cost decrease, the second column reports trade among newly established relations relative to total trade growth (i.e., how much of the growth in trade is driven by newly established pairs?). For a trade-cost increase, the second column reports pre-shock trade among relations that stopped trading in the counterfactual equilibrium, scaled by total trade growth. Industries are identi ed by 4-digit HS codes. The results are discussed in Section 4.1.
Data source. See Section 3.1.

industry average productivity upper bound yields an elasticity of 0.87 (cf. Figure B.1 in Appendix B). In line with this, Figure 1b shows that additional welfare gains tend to be largest for countries with low to medium levels of production, with a few exceptions. By contrast, the di erences are minor for most of the largest producers. This re ects 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 are close to the technological frontier (by having high productivity upper bounds), the two models tend to predict similar changes. Figure B.2a in Appendix B highlights that, particularly for lower-income countries, the di erences are often sizable relative to the benchmark gains. The welfare gains are 15% larger on average in the extensive version. However, the bottom quartile of the income distribution has, on average, a 30% larger gain, while the top quartile's welfare gain is only 2.4% higher when allowing for extensive-margin variation. Figure B.2b shows that a similar picture arises when I relate the di erences 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.

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 con rmed: welfare gains tend to be higher in the extensive model, especially for countries with a low- to medium-sized aggregate production level (cf. Figure B.3 in Appendix B), although the welfare implications are highly correlated between the two models. The ranking of countries in terms of their welfare di erences 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 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 7%. 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 289 new positive trade links per industry. However, this hides considerable heterogeneity: the number of newly established relations ranges across industries from 20 to 521 (SD=71). In 2,640 cases, a country starts exporting in a new industry or, there is a 2.9% growth in active exporter-industry cells. In terms of trade creation along the intensive margin, the contribution from these new flows is modest but non-negligible (second column of Table 3), which aligns with the findings of Eaton et al. (2013). Finally, in terms of production, there is also some movement: The extensive model predicts that in 103 (96) cases, a country starts (stops) producing in a given industry.

Rising trade barriers. In recent years, several protectionist measures have been adopted by countries that impede the free flow of goods across borders (see, e.g., Fajgelbaum, Goldberg, Kennedy, and Khandelwal (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 costs. I thus repeat the exercises above in the opposite direction.

Figure 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 2b reveals that in the extensive version of the model, the welfare losses are larger for most countries. This is primarily true for small- to medium-sized economies, which is intuitive in light of the above discussion on the implications of production constraints. 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.[31] Analogous to the liberalization experiment from above, Figure B.4a in Appendix B reveals that the relative additional losses (or gains) from accounting for the extensive margin can be meaningful. In the extensive model, four out of five countries experience larger losses compared to the benchmark model. On average, these additional losses amount to 8%

[31] The 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 (10). Higher productivity upper bounds result in larger destination-size terms (the summation). Therefore, the productivity limits imposed by ' 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 (see Section 2.2 for a detailed discussion). For some large countries, this cushioning appears to slightly outweigh the negative effect of productivity caps.

Figure 2: Welfare Changes after a 10% Rise in Global Trade Costs

(a) Welfare Changes

(b) Welfare Changes vs. Income

Notes. This gure depicts the welfare changes after a rise in global trade costs of 10%. The productivity upper bound is exible in the extensive model, and in nite in the benchmark model (eliminating extensive-margin variation). The solid line in Panel 2a is a 45-degree line. Panel 2b relates the (log) initial production level, $\log(X_j)$, to the di erence between the two models' welfare losses (i.e., the vertical minus the horizontal axis in Panel 2a), where a negative number implies larger losses in the extensive model. The results are discussed in Section 4.
Data source. See Section 3.1.

of the losses predicted by the benchmark model, but for countries in the bottom quartile of the income distribution, the additional losses increase to about 16% of the benchmark values. Figure B.4b shows that the implications are similar when the di erences are

Figure 3: Relating Welfare Changes to Trade Zeros

(a) 10% Decrease in

(b) 10% Increase in

Notes. This gure depicts the welfare changes after a drop (Panel 3a) and an increase (Panel 3b) in global trade costs of 10%. The productivity upper bound is exible in the extensive model, and in nite in the benchmark model (eliminating extensive-margin variation). Both panels relate a country's pre-shock share of export zeros across industries (i.e., the proportion of destinations that it does not serve) to its welfare change in the extensive relative to the benchmark model. The results are discussed in Section 4. Data source. See Section 3.1.

related to per worker income levels. These insights are con rmed when I instead increase trade costs by 5% or 15% (cf. Figure B.5 in Appendix B).

The number of bilateral industry-level zeros increases by 6.2% (cf. Table 3), which corresponds to an average of 254 dissolved trade relations per industry. Therefore, numerous industry-level connections appear to be relatively weak. The second column in Table 3 shows that these links tend to be small, causing 1.56% of the overall trade decline. In 3,157 cases, a country stops exporting in a given industry (representing 3.4% of active exporter-industry cells). Finally, the extensive model predicts that in 86 (97) cases, a country starts (stops) producing in a given industry.

Note that the effects of increasing or reducing on the extensive margin do not have to be symmetric. To gain intuition, consider a scenario where all countries trade with each other in every industry. Reductions in trade costs would then have no effect along the extensive margin. This is, however, not true for a trade-barrier increase in such an economy.

The key takeaway is twofold. First, industry-level bilateral zeros change by approximately two-thirds of the shock size. Second, welfare responses upon trade-cost shocks are often amplified when accounting for the extensive margin (i.e., when capturing the sparsity of trade flows in the data). 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 than in the benchmark model. Intuitively, since productivity upper bounds are closely related to economic size (cf. Figure B.1 in Appendix B), countries that serve few destinations across industries experience the largest relative welfare changes from modeling the extensive margin, as Figure 3 illustrates.

4.2 Alternative Specifications

In this section, I alter the calibration strategy in several ways to test the robustness of the main insights presented in Section 4.1.

As opposed to a single-industry EK model, the multi-industry version requires a calibration of to assess welfare changes. Therefore, the welfare implications do not only depend on the so-called trade elasticity,, but also on the elasticity of substitution. I re-run the calibration and main counterfactual exercises using $= \{1.5; 3\}$ instead of $= 2$. The welfare changes are comparable throughout, and the main insight that welfare responses are generally amplified through the extensive margin is preserved in both specifications (cf. Figures B.6 and B.7 in Appendix B). Analogous to the main specification, this primarily holds for small- to medium-sized economies.

I further test whether the choice of drives the main conclusions. The well-known baseline value $= 4.14$ 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 $= 6$ in his multi-industry EK model to match an aggregate elasticity of 4.1. Using $= 6$ in the calibration, the main counterfactuals yield similar conclusions as the baseline setting (cf. Figure B.8 in Appendix B). The magnitude of the welfare changes is slightly smaller with $= 6$. An intermediate value, $= 5$, is also commonly used (see, e.g., Costinot and Rodríguez-Clare, 2014). For completeness, Figure B.9 illustrates the results of the main simulation with $= 5$. The outcome is largely consistent with the insights from Section 4.1. A larger trade elasticity seems to predict smaller welfare losses upon shocks in the extensive relative to the benchmark model for more countries than the baseline specification.

The main specification equalizes the preference shifters, ^k_j, across countries due to a lack of consumption data at the 4-digit HS industry level. In an additional specification, I set ^k_j 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, ^k_j is set equal to the minimum positive value of ^k_j across importers within this industry. Otherwise, countries with zero imports would never consume product k. The results in Figure B.10 show that the main findings are robust to varying the preference shifters. The welfare changes tend to be somewhat larger in this alternative specification.[32]

Since the main calibration does not perfectly fit the data, an additional specification adds the remaining residual variation to the trade costs to fit the data as closely as possible. Figure B.11 presents the results, demonstrating that the main insights remain unaffected.

Finally, Table C.5 in Appendix C shows that the implications in terms of trade zeros are similar across all specifications. A higher trade elasticity, , leads to larger changes, which is intuitive. Interestingly, the calibration that perfectly fits the data predicts the largest responses. Thus, adding the difference between empirical and model-implied trade flows to the trade cost function creates many observations that will change along the extensive margin after a trade-cost shock. Therefore, relying on the main specification, which does not perfectly match the data, seems more conservative.[33]

[32] I have also tried omitting ^k_j. In this case, there are challenges in fitting the model to 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, Ganapati, and Muendler (2021) argue that local preference shifters are crucial in explaining variation in firm-level product exports.

[33] Relatedly, Dingel and Tintelnot (2021) demonstrate that in sparse settings, approaches that equate model-implied and empirical flows (in a hat-algebra way) yield worse predictions than those with fewer parameters than observations.

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 calibration does not require data on industry-level production, which is often unavailable for highly disaggregated industries. The framework replicates the empirical 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 two-thirds of the shock size. For instance, a 10% rise in trade costs increases the number of zeros by 6.2%. The welfare changes after trade-cost shocks are typically amplified when allowing for variation along the extensive margin of trade. This is primarily true for low- to medium-income countries, which tend to have the lowest number of active trade relations in most industries. Instead, many high-income countries have similar welfare changes as in a version of the model with infinite productivity upper bounds.

References

Adão, R., C. Arkolakis, and S. Ganapati (2020): Aggregate Implications of Firm Heterogeneity: A Nonparametric Analysis of Monopolistic Competition Trade Models, NBER Working Paper Series, No. 28081.

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, The American Economic Review, 93, 170 192.

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.

Armenter, R. and M. Koren (2014): A Balls-and-Bins Model of Trade, American Economic Review, 104, 2127 51.

Atkin, D., A. Costinot, and M. Fukui (2021): Globalization and the Ladder of Development: Pushed to the Top or Held at the Bottom? NBER Working Paper Series, No. 29500.

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.

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.

Cadot, O., C. Carrère, and V. Strauss-Kahn (2011): Export Diversi cation: What's behind the Hump? Review of Economics and Statistics, 93, 590 605.

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 E ects of NAFTA, The Review of Economic Studies, 82, 1 44.

(2022): Trade Policy, Handbook of International Economics, 5, 219 295.

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.

Carrère, C., M. Mrázová, and J. P. Neary (2020): Gravity without Apology: The Science of Elasticities, Distance and Trade, The Economic Journal, 130, 880 910.

Chaney, T. (2008): Distorted Gravity: The Intensive and Extensive Margins of International Trade, American Economic Review, 98, 1707 21.

Chen, J. and J. Roth (2024): Logs with Zeros? Some Problems and Solutions, The Quarterly Journal of Economics, 139, 891 936.

Chen, N. and D. Novy (2022): Gravity and Heterogeneous Trade Cost Elasticities, The Economic Journal, 132, 1349 1377.

Chor, D. (2010): Unpacking Sources of Comparative Advantage: A Quantitative Approach, Journal of International Economics, 82, 152 167.

Costinot, A., D. Donaldson, and I. Komunjer (2012): What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas, The Review of Economic Studies, 79, 581 608.

Costinot, A. and A. Rodríguez-Clare (2014): Trade Theory with Numbers: Quantifying the Consequences of Globalization, in Handbook of International Economics, Elsevier, vol. 4, 197 261.

Dekle, R., J. Eaton, and S. Kortum (2007): Unbalanced Trade, American Economic Review, 97, 351 355.

Dingel, J. I. and F. Tintelnot (2021): Spatial Economics for Granular Settings, NBER Working Paper Series, No. 27287.

Diodato, D., R. Hausmann, and U. Schetter (2022): A Simple Theory of Economic Development at the Extensive Industry Margin, HKS Working Paper, No. RWP22-016.

Dutt, P., I. Mihov, and T. Van Zandt (2013): The E ect of WTO on the Extensive and the Intensive Margins of Trade, Journal of International Economics, 91, 204 219.

Eaton, J. and A. C. Fieler (2019): The Margins of Trade, NBER Working Paper Series, No. 26124.

Eaton, J. and S. Kortum (2002): Technology, Geography, and Trade, Econometrica, 70, 1741 1779.

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.

Egger, P., R. Foellmi, U. Schetter, and D. Torun (2023): Gravity with History: On Incumbency E ects in International Trade, CEPR Discussion Paper, No. 18421.

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.

Fajgelbaum, P. D. and A. K. Khandelwal (2016): Measuring the Unequal Gains from Trade, The Quarterly Journal of Economics, 131, 1113 1180.

Fally, T. and J. Sayre (2018): Commodity Trade Matters, NBER Working Paper Series, No. 24965.

Feenstra, R. C. (2018): Restoring the Product Variety and Pro-Competitive Gains from Trade with Heterogeneous Firms and Bounded Productivity, Journal of International Economics, 110, 16 27.

Fieler, A. C. (2011): Nonhomotheticity and Bilateral Trade: Evidence and a Quantitative Explanation, Econometrica, 79, 1069 1101.

Foellmi, R., C. Hepenstrick, and J. Zweimüller (2018): International Arbitrage and the Extensive Margin of Trade between Rich and Poor Countries, The Review of Economic Studies, 85, 475 510.

Foellmi, R. and J. Zweimüller (2006): Income Distribution and Demand-Induced Innovations, The Review of Economic Studies, 73, 941 960.

French, S. (2016): The Composition of Trade Flows and the Aggregate E ects of Trade Barriers, Journal of International Economics, 98, 114 137.

French, S. and T. Zylkin (2024): The E ects of Free Trade Agreements on Product-Level Trade, European Economic Review, 162, 104673.

Giri, R., K.-M. Yi, and H. Yilmazkuday (2021): Gains from Trade: Does Sectoral Heterogeneity Matter? Journal of International Economics, 129, 103429.

Hausmann, R. and D. Rodrik (2003): Economic Development as Self-Discovery, Journal of development Economics, 72, 603 633.

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.

Hidalgo, C. A., B. Klinger, A.-L. Barabási, and R. Hausmann (2007): The Product Space Conditions the Development of Nations, Science, 317, 482 487.

Hummels, D. and P. J. Klenow (2005): The Variety and Quality of a Nation's Exports, American Economic Review, 95, 704 723.

Imbs, J. and R. Wacziarg (2003): Stages of Diversi cation, American economic review, 93, 63 86.

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.

Matsuyama, K. (2000): A Ricardian Model with a Continuum of Goods under Nonhomothetic Preferences: Demand Complementarities, Income Distribution, and North-South Trade, Journal of political Economy, 108, 1093 1120.

(2019): Engel's Law in the Global Economy: Demand-Induced Patterns of Structural Change, Innovation, and Trade, Econometrica, 87, 497 528.

Melitz, M. J. (2003): The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity, Econometrica, 71, 1695 1725.

Melitz, M. J. and S. J. Redding (2015): New Trade Models, New Welfare Implications, American Economic Review, 105, 1105 1146.

Murphy, K. M., A. Shleifer, and R. Vishny (1989): Income Distribution, Market Size, and Industrialization, The Quarterly Journal of Economics, 104, 537 564.

Novy, D. (2013): International Trade without CES: Estimating Translog Gravity, Journal of International Economics, 89, 271 282.

Ossa, R. (2015): Why Trade Matters after All, Journal of International Economics, 97, 266 277.

Ravikumar, B., A. M. Santacreu, and M. Sposi (2019): Capital Accumulation and Dynamic Gains from Trade, Journal of International Economics, 119, 93 110.

Santos Silva, J. and S. Tenreyro (2006): The Log of Gravity, The Review of Economics and Statistics, 88, 641 658.

Simonovska, I. (2015): Income Di erences and Prices of Tradables: Insights from an Online Retailer, The Review of Economic Studies, 82, 1612 1656.

Simonovska, I. and M. E. Waugh (2014): The Elasticity of Trade: Estimates and Evidence, Journal of International Economics, 92, 34 50.

Appendix

A Detailed Derivations

This appendix outlines the steps to obtain the expressions presented in Section 2, as well as extensions of the theory.

A.1 Mass of Input Varieties

To derive equation (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 (5), and $g_{ij}(p) = dG_{ij}(p)/dp$, that is,

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

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

$$\begin{aligned} \pi_{ij} &:= \int_{\underline{p}_{ij}}^{\infty} \prod_{i' \neq i} [1 - G_{i'j}(p)]\, g_{ij}(p) dp \\ &= \sum_{n=i}^{N} \left[\int_{\underline{p}_{nj}}^{\underline{p}_{(n+1)j}} \left(\prod_{i' \le n;\, i' \neq i} \frac{\exp\left(-T_{i'} (w_{i'} \tau_{i'j})^{-\theta} p^{\theta}\right)}{\exp\left(-T_{i'} \varphi^{-\theta}\right)} \right) \right. \\ &\qquad \left. \frac{\exp\left(-T_i (w_i \tau_{ij})^{-\theta} p^{\theta}\right)}{\exp\left(-T_i \varphi^{-\theta}\right)} \theta T_i (w_i \tau_{ij})^{-\theta} p^{\theta-1} dp \right] \\ &= T_i (w_i \tau_{ij})^{-\theta} \sum_{n=i}^{N} \int_{\underline{p}_{nj}}^{\underline{p}_{(n+1)j}} \Lambda_j^{1:n} \exp\left(-p^{\theta} \Phi_j^{1:n}\right) \theta p^{\theta-1} dp \\ &= T_i (w_i \tau_{ij})^{-\theta} \sum_{n=i}^{N} \frac{\Lambda_j^{1:n}}{\Phi_j^{1:n}} \left[\exp\left(-\underline{p}_{nj}^{\theta} \Phi_j^{1:n}\right) - \exp\left(-\underline{p}_{(n+1)j}^{\theta} \Phi_j^{1:n}\right) \right]; \qquad \text{(A.1)} \end{aligned}$$

where $\underline{p}_{(N+1)j} = \infty$, and exporters are sorted according to inequalities (6).[34] The second line splits the integral at the exporters' lower-bound prices to exploit the piecewise

---

[34] Letting $\varphi \to \infty$ yields the canonical EK expression for trade shares

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

since $\lim_{\varphi \to \infty} \Lambda_j^{1:n} = 1\ \forall n \le N$, and $\lim_{\varphi \to \infty} \underline{p}_{nj} = 0\ \forall n \le N$.

definition of $G_j(p)$ in equation (7). The variables $\lambda_j^{1:n}$ and $\Phi_j^{1:n}$ are as defined in equation (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). Put differently, there is not necessarily positive demand for a variety at every price level.

Using the notation from equation (2), only varieties with prices smaller than or equal to $p_j^k(\omega_j^k)$ 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,

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

where $1\{\}$ is an indicator function. The expression for m_{ij} is almost identical to that of π_{ij} in equation (A.1) 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 2.2.

A.2 Trade Flows

To understand the derivation of equation (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 (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

$$
\begin{aligned}
E[p_j(\omega) \mid p_j(\omega) \leq \bar{p}_j]\, G_j(\bar{p}_j) &= \int_{\min\{\underline{p}_{1j}, \bar{p}_j\}}^{\bar{p}_j} p\, g_j(p)\,dp \\
&= \sum_{n=1}^{N} \lambda_j^{1:n} \int_{\min\{\underline{p}_{nj}, \bar{p}_j\}}^{\min\{\underline{p}_{(n+1)j}, \bar{p}_j\}} p\, \theta \exp\left(-p^{\theta} \Phi_j^{1:n}\right) p^{\theta-1} dp
\end{aligned}
$$

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

$$E[p_j(!)|p_j(!) \quad p_j] \, G_j(p_j) = \sum_{n=1}^{N} \quad {}_{j}^{1:n} \quad {}_{j}^{1:n} \quad {}^{1=} \int_{\min\{\underline{p}_{nj};\, p_j\} \ {}_{j}^{1:n}}^{\min\{\underline{p}_{(n+1)j};\, p_j\} \ {}_{j}^{1:n}} x^{1=} \exp(\quad x)dx$$

$$= \sum_{n=1}^{N} \quad {}_{j}^{1:n} \quad {}_{j}^{1:n} \quad {}^{1=} \left[e \quad \frac{+1}{}; \min\{\underline{p}_{nj};\, p_j\} \quad {}_{j}^{1:n}; \min\{\underline{p}_{(n+1)j};\, p_j\} \quad {}_{j}^{1:n} \right) \Bigg] ;$$

where $e(\ s; \underline{x}; x) := \int_{\underline{x}}^{x} x^{s \ 1} \exp(\ x)dx$. The fact that $M_j = G_j(p_j)L_j$ gives us $X_j = E[p_j(!)|p_j(!) \quad p_j] M_j = E[p_j(!)|p_j(!) \quad p_j] G_j(p_j) L_j$. Total expenditures by country j on intermediates (within an industry), X_j, are thus equal to the expression above multiplied by j's population size, that is,

$$X_j = L_j \sum_{n=1}^{N} \quad {}_{j}^{1:n} \quad {}_{j}^{1:n} \quad {}^{1=} \ e \quad \frac{+1}{}; \min\{\underline{p}_{nj};\, p_j\} \quad {}_{j}^{1:n}; \min\{\underline{p}_{(n+1)j};\, p_j\} \quad {}_{j}^{1:n} \Bigg) : \qquad \text{(A.2)}$$

By noting that $e \quad \frac{+1}{}; p; p = 0$, we can see how equation (A.2) allows for import zeros. The derivation below makes this clearer.[35]

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 $p \quad p_j$, the expected price can be calculated as

$$\frac{1}{m_{ij} + 1\{m_{ij} = 0\}} \int_{\min\{\underline{p}_{ij};\, p_j\}}^{p_j} p \prod_{i^0 \neq i} [1 \quad G_{i^0j}(p)] \, g_{ij}(p)dp$$

$$= \frac{1}{m_{ij} + 1\{m_{ij} = 0\}} \sum_{n=i}^{N} \quad {}_{j}^{1:n} \frac{T_i(w_i \quad {}_{ij})}{{}_{j}^{1:n}} \int_{\min\{\underline{p}_{nj};\, p_j\}}^{\min\{\underline{p}_{(n+1)j};\, p_j\}} p \quad {}_{j}^{1:n} \exp(\quad p \quad {}_{j}^{1:n}) \, p \quad {}^{1}dp$$

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

[35] To see the connection to the standard model, let $\quad ! \ 1$ and $p_j \ ! \ 1$. This yields $m_j = 1$, and

$$x_j = \quad {}_{j}^{1:N} \quad {}^{1=} \ e \quad \frac{+1}{}; 0; 1 \quad ;$$

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 $e \quad \frac{(\quad 1)}{}; 0; 1 \quad {}^{1=(\quad 1)}$ for $\quad > \ 0$, which is a common constant across importers. Note that such a constant shifter is irrelevant to any calibration exercise or welfare calculation.

above

$$x_{ij} = T_i (w_i \; {}_{ij}) \; \sum_{n=i}^{N} \frac{{}_j^{1:n}}{{}_j^{1:n}} \; {}_j^{1:n} \; {}^{1=} \; e \; \frac{+1}{} ; \min f\, \underline{p}_{nj} ; p_j \, g \; {}_j^{1:n} ; \min f\, \underline{p}_{(n+1)j} ; p_j \, g \; {}_j^{1:n} \; ! \; ;$$

which corresponds to equation (10). Total expenditures on intermediates from i are given by $X_{ij} = L_j x_{ij}$. For a discussion, see Section 2.2.

A.3 Equilibrium with an Alternative Assembly Function

This part of the appendix derives an equilibrium with an alternative nal-goods-assembly function. The main di erence to the speci cation in Section 2.1 is that intermediate inputs are now substitutable along both the extensive and intensive margins.

Assembly function. For what follows, I replace the assembly function in equation (1) with a Stone-Geary-type speci cation (resembling the utility function in Simonovska (2015))[36]

$$q_j^k = \int_0^1 \log \; \underline{q} + q_j^k(!^{\,k}) \; d!^{\,k} ;$$

where $\underline{q} > 0$ is a constant, common across countries. This yields the following demand functions

$$q_j^k(!^{\,k}) = \max \left(\underline{q} \left(\frac{p_j^k}{p_j^k(!^{\,k})} \; 1 \right) ; 0 \right) ;$$

where the choke price is determined by

$$p_j^k = \frac{{}_j^k \, {}^{\frac{1}{}} \; q_j^k \; {}^{\frac{1}{}}}{{}_j \, \underline{q}} :$$

Therefore, this assembly function also naturally yields a choke price (whereby some inputs may not be demanded in equilibrium) while allowing for intensive-margin adjustments of the quantity used of a variety ($q_j^k(!^{\,k})$). Note that the homothetic utility function in Section 2.1 implies there will be at least one variety where $q_j^k(!^{\,k}) > 0$ for every k. Moreover, recall that the assembly function is theoretically equivalent to a subutility function.

Trade ows. To derive total expenditures following similar steps as in Appendix A.2, we can compute the average expenditure per variety and multiply it by the mass of consumed varieties within a country and industry ($G_j^k(p_j^k)$). To economize on notation, I again drop

[36] One can also follow the steps outlined in this appendix to derive an equilibrium with, e.g., a quadratic assembly function $q_j^k = \int_0^1 \underline{q} \quad q_j^k(!^{\,k})^{\,2} d!^{\,k}$, with $\underline{q} > 0$.

the industry superscript k. This yields

$$
\begin{aligned}
x_j &= E[p_j(\omega) q_j(\omega) | p_j(\omega) \quad \bar{p}_j] G_j(\bar{p}_j) \\
&= \int_{\underline{p}_{1j}}^{\bar{p}_j} \underline{q} \left(\frac{\bar{p}_j}{p} \quad 1 \right) p g_j(p) dp \\
&= \underbrace{\sum_{n=1}^{N} \underline{q} \, {}_j^{1:n} \bar{p}_j \left(\left[\exp \underline{p}_{nj} \quad \exp \underline{p}_{(n+1)j} \right] \quad {}_j^{1:n} \, {}^{\frac{1}{}} e \quad \left(\frac{+1}{} ; \underline{p}_{nj} ; \underline{p}_{(n+1)j} \right) \right)}_{x_j^{(n)}} ;
\end{aligned}
\tag{A.3}
$$

where $\underline{p}_{nj}$:= min $\{\underline{p}_{nj}; \bar{p}_j\}$ ${}_j^{1:n}$, and as before, $(s; \underline{x}; x) := \int_{\underline{x}}^{x} x^{s\ 1} \exp(x)dx$. Recall that aggregate expenditures are simply equal to $X_j = L_j x_j$, while ${}_j^{1:n}$ and ${}_j^{1:n}$ are as defined in equation (7).

Following analogous steps, one can easily show that bilateral trade flows follow the same general expression as shown in Section 2.2 (see equation (12)), that is,

$$
x_{ij} = \sum_{n=i}^{N} \frac{T_i (w_i \quad_{ij})}{\sum_{i'=1}^{n} T_{i'} (w_{i'} \quad_{i'j})} x_j^{(n)} ; \tag{A.4}
$$

where $x_j^{(n)}$ is defined in equation (A.3). In words, $x_j^{(n)}$ is the expenditure segment in which exporters $i \quad n$ are competitive (while $i > n$ are not). Finally, total quantity consumed (in an industry) equals

$$
\begin{aligned}
q_j &= E[q_j(\omega) | p_j(\omega) \quad \bar{p}_j] G_j(\bar{p}_j) \\
&= \int_{\underline{p}_{1j}}^{\bar{p}_j} \underline{q} \left(\frac{\bar{p}_j}{p} \quad 1 \right) g_j(p) dp \\
&= \bar{p}_j \underline{q} \left(\sum_{n=1}^{N} {}_j^{1:n} \, {}_j^{1:n} \, {}^{\frac{1}{}} e \quad \left(\frac{1}{} ; \underline{p}_{nj} ; \underline{p}_{(n+1)j} \right) \right) \quad \underline{q} G_j(\bar{p}_j) :
\end{aligned}
\tag{A.5}
$$

Equilibrium. Let P_j^k denote the average per-unit expenditure, that is,

$$
P_j^k := \frac{1}{q_j^k} \int_0^{\omega_j^k} p_j^k(\omega^k) q_j^k(\omega^k) d\omega^k :
$$

With this definition, we can derive an expression that implicitly pins down the choke prices (as in Section 2.1)

$$
P_j^k q_j^k = E_j \frac{{}_j^k \bar{p}_j^k \quad P_j^k}{\sum_{k'=1}^{K} {}_j^{k'} \bar{p}_j^{k'} \quad P_j^{k'}} :
$$

Using the same market-clearing conditions as in Section 2.3 (i.e. $E_j = w_j$ and $P_j^k = x_j^k / q_j^k$), this expression becomes

$$
\frac{w_j}{q_j^k} = {}_j^k \, {}^{1} \bar{p}_j^k \left[\sum_{k'=1}^{K} {}_j^{k'} \bar{p}_j^{k'} \quad \frac{x_j^{k'}}{q_j^{k'}} \right] ; \tag{A.6}
$$

where x_j^k and q_j^k are determined as described above. The balanced-trade condition (13) pins down the wages.

To sum up, for given parameters $;\ ,\ \{L_i\}_i,\ \{\ _i^k\}_{i;k},\ \varphi;\ \{T_i^k\}_{i;k};\ \{\ _{ij}^k\}_{i;j;k}$, the equilibrium in this economy is determined by a set of wages and choke prices $\{w_i\}_i;\ \{\bar{p}_i^k\}_{i;k}$, which satisfy equations (A.3) (A.6) and (13).

A.4 Equilibrium with a General Utility Function

In this part of the appendix, I characterize the equilibrium derived in Section 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 final 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 $u'(\cdot) > 0; u''(\cdot) < 0; u'(0) \le \infty$; and $q_j^k \ge 0$.[37] The household maximizes its utility subject to a budget constraint $E_j \ge \sum_{k=1}^{K}\left[\int_0^1 p_j^k(\omega^k) q_j^k(\omega^k) d\omega^k\right]$. As in the main text, q_j^k is assembled according to equation (1), 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

$$u'(q_j^k) - \lambda_j \bar{p}_j^k \le 0$$

$$\Rightarrow \quad q_j^k\left[u'(q_j^k) - \lambda_j \bar{p}_j^k\right] = 0;$$

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

$$p_j^k(0) \ge \frac{u'(0)}{\lambda_j};$$

[37] The 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., Fieler (2011); Caron et al. (2014).

where $u''(\cdot) > 0$ implies $\lambda_j > 0$. Hence, whenever $\underline{p}^k_{1j} \geq u'(0)/\lambda_j$, we have $q^k_j = 0$.[38] For $q^k_j > 0$, let average expenditure on intermediates from industry k be given by P^k_j, as defined in equation (2).

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

$$q^k_j = u'^{-1}(\lambda_j p^k_j).$$

Multiplying both sides by P^k_j, and summing over k yields

$$\sum_{k=1}^{\bar{k}} P^k_j q^k_j = \sum_{k=1}^{\bar{k}} u'^{-1}(\lambda_j p^k_j) P^k_j.$$

Note that the left-hand side is equal to the household's total expenditure, and therefore equal to w_j. Using $P^k_j = X^k_j / (G^k_j(p^k_j) L_j)$ when $q^k_j > 0$, we obtain

$$w_j L_j = \sum_{k=1}^{\bar{k}} u'^{-1}(\lambda_j p^k_j) \frac{X^k_j}{G^k_j(p^k_j)}. \tag{A.7}$$

For a given set of wages and choke prices, equation (A.7) pins down λ_j. $G^k_j(p^k_j)$ and X^k_j are still given by equations (9) and (11), respectively, multiplied by L_j. Choke prices are given by

$$p^k_j = \frac{u'(G^k_j(p^k_j))}{\lambda_j} \quad \text{for } k \leq \bar{k}. \tag{A.8}$$

For a given set of wages, Lagrange multipliers, and parameters, equation (A.8) implicitly determines p^k_j for $k \leq \bar{k}$. For $k > \bar{k}$, and thus $q^k_j = 0$, set $p^k_j = \underline{p}^k_{1j}$.

To sum up, for given parameters $; L; \varphi; T;$, the equilibrium is determined by a set of Lagrange multipliers, wages, and choke prices λ, w, p, that satisfy the equilibrium conditions in equations (9)–(13), and (A.7)–(A.8).

A.5 Equilibrium with Sectoral Linkages

In this part of the appendix, I outline a simple way to extend the model presented in Section 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

[38] Recall that $\underline{p}^k_{ij} := w_i \tau^k_{ij} / \varphi$, and that exporters are sorted according to inequalities (6).

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

$$U_j = \frac{}{1} \sum_{k=1}^{K} {}_j^{k\;\frac{1}{}} \; q_j^{k\;\frac{1}{}} :$$

However, final goods are now assembled according to

$$q_j^k = \sum_{s=1}^{K} {}_j^{s;k} \ln \int_0^1 q_j^{s;k}(\omega^s)\, d\omega^s \;;$$

where $q_j^{s;k}(\omega^s) \in \{0;1\}$ is the quantity of variety ω^s used to assemble product k, and ${}_j^{s;k} > 0$ are exogenous s-k-specific productivity shifters.[39] A household first uses the cheapest varieties from s in each industry k. As in Section 2.1, let varieties' indexes ω^s be increasing in prices. Households maximize their utility with respect to $\{\omega_j^{s;k}\}_{s;k}$, where the Lagrange function is given by

$$L\left(\{\omega_j^{s;k}\}_{s;k};\; {}_j\right) = \frac{}{1} \sum_{k=1}^{K} {}_j^{k\;\frac{1}{}} \left[\sum_{s=1}^{K} {}_j^{s;k} \ln \left(\int_0^{\omega_j^{s;k}} d\omega^s \right) \right]^{\frac{1}{}} + {}_j \left[E_j \sum_{k=1}^{K} \sum_{s=1}^{K} \int_0^{\omega_j^{s;k}} p_j^s(\omega^s)\, d\omega^s \right] :$$

The first-order condition for $\omega_j^{s;k}$ is

$${}_j^{k\;\frac{1}{}} \; q_j^{k\;\frac{1}{}} \; \frac{{}_j^{s;k}}{q_j^{s;k}} = {}_j\, p_j^s(\omega_j^{s;k});$$

where $q_j^{s;k} := \int_0^{\omega_j^{s;k}} d\omega^s$ 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 $p_j^{s;k} := p_j^s(\omega_j^{s;k})$. Rearranging we obtain

$$q_j^k = {}_j^k \; {}_j \; {}_j^{s;k} \; p_j^{s;k} q_j^{s;k} \qquad \forall s;k: \tag{A.9}$$

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

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

[39] The derivations would be analogous with a standard Cobb-Douglas function like

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

although the elasticities in equation (A.10) would differ. The log-specification is chosen solely for illustrative purposes.

Multiplying both sides of equation (A.9) by P_j^k, and summing over k, yields an expression for $_j$ to plug back into equation (A.9). Together with the budget constraint, this yields the following expenditure shares

$$P_j^k q_j^k = E_j \frac{{}_j^k \; {}_j^{s;k} \; p_j^{s;k} q_j^{s;k} \; P_j^k}{\sum_{k'=1}^{K} {}_j^{k'} \; {}_j^{s;k'} \; p_j^{s;k'} q_j^{s;k'} \; P_j^{k'}}:$$

These conditions pin down the choke prices in equilibrium.

Equilibrium. To characterize the equilibrium in this economy, we need to determine intermediate choke prices $\{p_j^{s;k}\}_{j;s;k}$. Note that due to constant returns to scale of intermediate producers, the expenditures on industry-s varieties for a nal product k do not a ect the industry-s sales used for product k'. Therefore, for a given choke price $p_j^{s;k}$, $q_j^{s;k}$ is simply given by $G_j^s(p_j^{s;k})$ in equilibrium. Similarly, per capita expenditures on s-varieties to assemble product k, $x_j^{s;k}$, are given by the expression in equation (11) evaluated at $p_j = p_j^{s;k}$, that is,

$$x_j^{s;k} = \sum_{n=1}^{N} {}_j^{s;1:n} \; {}_j^{s;1:n} \; {}^{1=} \; e \; \frac{+1}{} ; \min\{\underline{p}^s_{nj}; p_j^{s;k}\} \; {}_j^{s;1:n} ; \min\{\underline{p}^s_{(n+1)j}; p_j^{s;k}\} \; {}_j^{s;1:n} :$$

Moreover, $q_j^k = \sum_{s=1}^{K} {}_j^{s;k} \ln\left[G_j^s(p_j^{s;k})\right]$, 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

$$q_j^k = w_j \frac{{}_j^k \; {}_j^{s;k} \; p_j^{s;k} G_j^s(p_j^{s;k})}{\sum_{k'=1}^{K} {}_j^{k'} \; {}_j^{s;k'} \; p_j^{s;k'} G_j^s(p_j^{s;k'}) \; \frac{x_j^{k'}}{q_j^{k'}}}: \qquad \text{(A.10)}$$

For a given set of wages and parameters, this equation implicitly determines $p_j^{s;k}$.

Per capita trade ows from source country i in industry s to assemble product k, $x_{ij}^{s;k}$, are given by equation (10) evaluated at $p_j = p_j^{s;k}$. The balanced trade condition that pins down wages becomes

$$\sum_{i \neq j} \sum_{k=1}^{K} \sum_{s=1}^{K} L_j x_{ij}^{s;k} \quad L_i x_{ji}^{s;k} = 0: \qquad \text{(A.11)}$$

To sum up, equations (A.10) and (A.11) pin down equilibrium choke prices and wages for a given set of parameters.

B Additional Figures

Figure B.1: Correlation between Total Production and Productivity Upper Bounds

Notes. This figure plots the log of a country's total production, $\log(X_i)$, against the log of its across-industry average productivity upper bound. The productivity upper bounds, φ_i^k, stem from the calibration using the extensive model (with extensive-margin variation). Section 3.1 describes the calibration strategy.
Data source: See Section 3.1.

Figure B.2: Welfare Changes after a 10% Drop in Global Trade Costs   Add-On

(a) Relative Welfare Gains vs. Income

(b) Welfare Gains vs. Income per Worker

Notes. This gure depicts the welfare changes after a drop in global trade costs of 10%. The productivity upper bound is exible in the extensive model, and in nite in the benchmark model. Panel B.2a 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.2b relates the (log) initial income level per worker, $\log(w_j)$, to the di erence between the two models' welfare gains. The results are discussed in Section 4.
Data source. See Section 3.1.

Figure B.3: Welfare Changes after a Drop in Global Trade Costs

(a) 5% Decrease in : Welfare Gains vs. Income

(b) 15% Decrease in : Welfare Gains vs. Income

Notes. These gures depict welfare changes after a drop in global trade costs of 5% (Panel B.3a) and 15% (Panel B.3b). The productivity upper bound is exible in the extensive model, and in nite in the benchmark model. Both panels relate the (log) initial production level, $\log(X_j)$, to the di erence between the two models' welfare gains. The results are discussed in Section 4.
Data source: See Section 3.1.

Figure B.4: Welfare Changes after a 10% Rise in Global Trade Costs   Add-On

(a) Relative Welfare Changes vs. Income

(b)   Welfare Changes vs. Income per Worker

Notes. This gure depicts the welfare changes after a rise in global trade costs of 10%. The productivity upper bound is exible in the extensive model, and in nite in the benchmark model. Panel B.4a 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.4b relates the (log) initial income level per worker, $\log(w_j)$, to the di erence between the two models' welfare losses. The results are discussed in Section 4. Data source: See Section 3.1.

Figure B.5: Welfare Changes after an Increase in Global Trade Costs

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

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

Notes. These gures depict welfare changes after a rise in global trade costs of 5% (Panel B.5a) and 15% (Panel B.5b). The productivity upper bound is exible in the extensive model, and in nite in the benchmark model. Both panels relate the (log) initial production level, $\log(X_j)$, to the di erence between the two models' welfare losses. The results are discussed in Section 4.
Data source. See Section 3.1.

Figure B.6: Welfare Changes after Global Trade-Cost Shocks — Robustness (= 1:5)

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

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

Notes. These plots are the counterparts of Figures 1b (Panel B.6a) and 2b (Panel B.6b) when setting = 1:5 instead of = 2 in the calibration. For details, see Figures 1b and 2b.

Figure B.7: Welfare Changes after Global Trade-Cost Shocks Robustness (= 3)

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

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

Notes. These plots are the counterparts of Figures 1b (Panel B.7a) and 2b (Panel B.7b) when setting = 3 instead of = 2 in the calibration. For details, see Figures 1b and 2b.

Figure B.8: Welfare Changes after Global Trade-Cost Shocks Robustness (= 6)

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

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

Notes. These plots are the counterparts of Figures 1b (Panel B.8a) and 2b (Panel B.8b) when setting = 6 instead of = 4.14 in the calibration. For details, see Figures 1b and 2b.

Figure B.9: Welfare Changes after Global Trade-Cost Shocks   Robustness (= 5)

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

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

Notes. These plots are the counterparts of Figures 1b (Panel B.9a) and 2b (Panel B.9b) when setting $= 5$ instead of $= 4:14$ in the calibration. For details, see Figures 1b and 2b.

Figure B.10: Welfare Changes after Global Trade-Cost Shocks Robustness (^k_j from Imports)

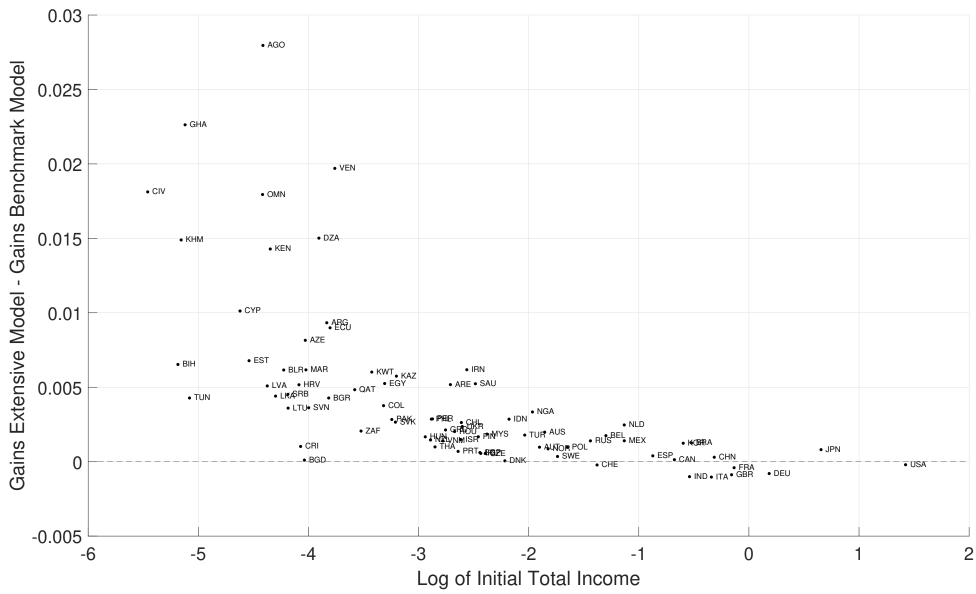
(a) 10% Decrease in : Welfare Changes vs. Income

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

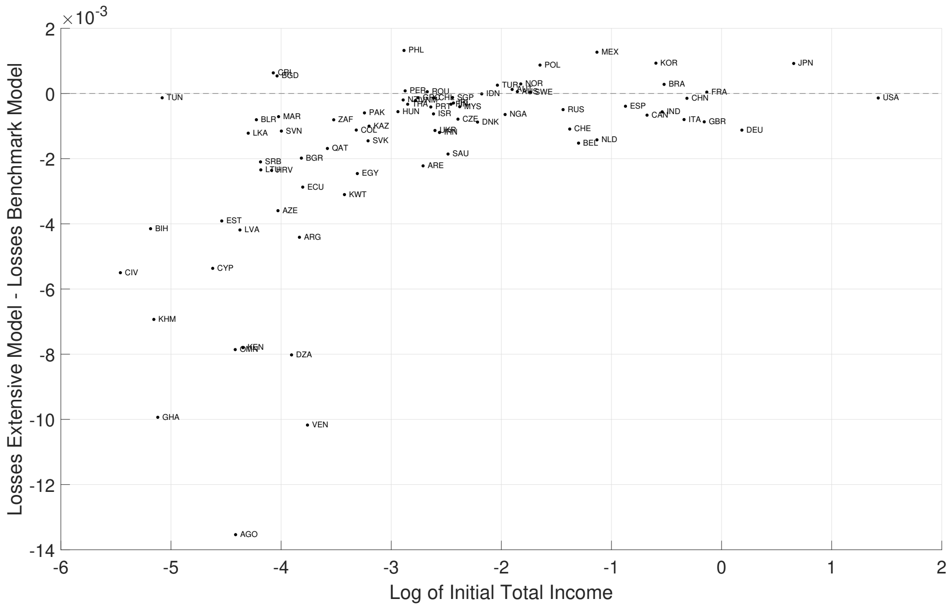
Notes. These plots are the counterparts of Figures 1b (Panel B.10a) and 2b (Panel B.10b) when setting ^k_j equal to industry k's import share in country j . For details, see Figures 1b and 2b.

Figure B.11: Welfare Changes after Global Trade-Cost Shocks – Robustness (Residuals added to Trade Costs)

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



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



Notes. These plots are the counterparts of Figures 1b (Panel B.11a) and 2b (Panel B.11b) when adding the residuals after the calibration to the trade costs to match the data as close as possible. For details, see Figures 1b and 2b.

C Additional Tables

Table 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
No. of Observations	7,836,800				
Panel B: Model	Min.	Q1	Median	Q3	Max.
Sh. Exported Industries	0.83	0.96	0.98	0.99	1
Sh. Served Destinations[(i)]	0.01	0.29	0.65	0.92	1
Sh. Imported Industries	0.98	1	1	1	1
Sh. Source Countries[(ii)]	0.01	0.38	0.65	0.8	1
No. of Observations	7,836,800				

Notes. This table reproduces Table 2 but using the "benchmark model" for Panel B instead of the "extensive model." The "benchmark model" sets $\phi_i^k \to \infty \ \forall i, k$, 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 3.2 and the notes to Table 2.
[(i)] Conditional on exporting in a given industry.
[(ii)] Conditional on importing in a given industry.
Data source. Atlas of Economic Complexity.

Table 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
No. of Observations	7,836,800				
Panel B: Model	Min.	Q1	Median	Q3	Max.
Sh. Exported Industries	0.68	0.95	0.97	0.99	1
Sh. Served Destinations[(i)]	0.01	0.08	0.25	0.62	1
Sh. Imported Industries	0.85	0.98	0.98	1	1
Sh. Source Countries[(ii)]	0.01	0.15	0.33	0.53	1
No. of Observations	7,836,800				

Notes. This table reproduces Table 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 3.2 and the notes to Table 2.
[(i)] Conditional on exporting in a given industry.
[(ii)] Conditional on importing in a given industry.
Data source. Atlas of Economic Complexity.

Table 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
No. 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
No. of Observations	7,836,800				

Notes. This table reproduces Table 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 3.2 and the notes to Table 2.

[(i)] Conditional on exporting in a given industry.

[(ii)] Conditional on importing in a given industry.

Data source. Atlas of Economic Complexity.

Table C.4: Regressions of $\log(\tau)$ on Gravity Variables

	Coefficients (SEs)			
	With Extensive Margin		W/out Extensive Margin	
	(1)	(2)	(3)	(4)
Log of Distance	0.136 (0.014)	0.074 (0.015)	0.151 (0.017)	0.088 (0.018)
Contiguity	-0.338 (0.049)	-0.432 (0.050)	-0.404 (0.059)	-0.455 (0.060)
Colonial Ties	-0.214 (0.065)	-0.124 (0.052)	-0.369 (0.076)	-0.266 (0.062)
Common Language	-0.095 (0.042)	-0.159 (0.039)	-0.108 (0.052)	-0.216 (0.045)
RTA	-0.102 (0.027)	-0.010 (0.027)	-0.292 (0.033)	-0.143 (0.033)
Imp. & Exp. FEs		✓		✓
No. of Observations	6,162	6,162	6,162	6,162

Notes. This table reports OLS results from regressing the log of a country pair's average trade costs across industries on gravity variables. Robust standard errors are reported in parentheses. The first two columns use the trade costs calibrated with the extensive model, while columns (3) and (4) use the benchmark model (without extensive-margin variation). 'RTA' is a dummy variable equal to one if a country pair has an active regional trade agreement in the year 2015 and zero otherwise. Columns (2) and (4) include importer and exporter FEs. The results are discussed in Section 3.2.
Data source. See Section 3.1. Covariates from CEPII/Head, Mayer, and Ries (2010).

Table C.5: Bilateral Zeros after Changes in Global Trade Costs – Robustness

	%-Change in Bilateral Zeros Relative to Baseline					
%-Change in	$\sigma = 1.5$	$\sigma = 3$	$\theta = 5$	$\theta = 6$	ξ_j^k Importer-Specific	Fit Residuals
−10%	−6.9%	−7%	−7.9%	−8.6%	−7.3%	−11.9%
+10%	6%	6.2%	6.9%	7.6%	6.6%	13.3%

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 3 is altered via the elasticity of substitution, σ, the trade elasticity, θ, the preference shifters, , or by adding the residuals after the calibration to the trade costs to match the data as close as possible ('Fit Residuals'). The results are discussed in Section 4.2.