NBER WORKING PAPER SERIES

TRADE WITH CORRELATION

Nelson Lind
Natalia Ramondo

Working Paper 24380
http://www.nber.org/papers/w24380

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2018

We thank Costas Arkolakis for a very helpful discussion on an early draft. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Trade with Correlation
Nelson Lind and Natalia Ramondo
NBER Working Paper No. 24380
March 2018
JEL No. F1

ABSTRACT

We develop a trade model in which productivity—the result of a country's ability to adopt global technologies—presents an arbitrary pattern of spatial correlation. The model generates the full class of import demand systems consistent with Ricardian theory, and, hence, captures its full macroeconomic implications. In particular, our framework formalizes Ricardo's insight—absent from the canonical Ricardian model—that countries gain more from trade partners with relatively dissimilar technology. Incorporating this insight into the calculations of macro counterfactuals entails a simple correction to self-trade shares. Our framework enables general aggregation results which tie micro optimization to macro demand systems and guide counterfactual analysis based on micro estimates. Our quantitative application to a multi-sector trade model suggests that countries specialized in low correlation sectors have 40 percent higher gains from trade relative to countries specialized in high correlation sectors. After accounting for correlation, the model predicts that lower trade costs for imports from China, rather than Canada, have the largest impact on real wages in the United States.

Nelson Lind
Department of Economics
Emory University
1602 Fishburne Drive
Rich Bldg., 3rd Floor
Atlanta, GA 30322-2240
nlind@emory.edu

Natalia Ramondo
School of Int'l Relations and Pacific Studies
University of California at San Diego
9500 Gilman Drive, MC 0519
La Jolla, CA 92093
and NBER
nramondo@ucsd.edu

1 Introduction

Two hundred years ago, David Ricardo (1818) put forward the idea that cross-country differences in production technologies can lead to gains from trade. Ricardo's work, which extended Adam Smith (1776)'s idea on comparative advantage to international trade, had two basic insights. First, if two countries have production possibility frontiers with different slopes (i.e., the ratio of productivity across sectors), then there is scope for engaging in mutually profitable trade. Second, the gains are higher when trade occurs between more dissimilar partners.

Eaton and Kortum (2002, henceforth, EK) capture the first insight of Ricardo's idea by treating technologies as random variables, giving rise to a rich theoretical and quantitative literature. Their results are based on a key property of Fréchet distributions known as max stability: The maximum of Fréchet-distributed random variables is also Fréchet. In particular, they assume independence and a common shape to get

$$\mathbb{P}[\max\{A_1, \ldots, A_N\} \leq a] = \exp\Big(- \sum_{o=1}^{N} T_o a^{-\theta} \Big).$$

The scale of the maximum across origin countries equals the sum of the underlying scale parameters, T_o. The max-stability property gives tractability to the EK model of trade, but independence imposes symmetry and implies that all trading partners are indistinguishable in generating gains from trade. As a result, the EK model of trade does not capture the second aspect of the old Ricardian idea—which may be important to understand why countries choose certain trading partners and not others.[1]

In this paper, we develop a theory of trade that allows for arbitrary patterns of correlation in technology between countries. Relaxing the independence assumption implies that the distribution of the maximum of Fréchet random variables is

$$\mathbb{P}[\max\{A_1, \ldots, A_N\} \leq a] = \exp\Big(- G(T_1, \ldots, T_N) a^{-\theta} \Big),$$

for some correlation function G. Countries can now have different weight on the scale of the maximum due to the correlation structure of technology across countries. In this

[1] Costinot et al. (2015) analyze the relation between the patterns of comparative advantage and optimal trade taxes schemes. They perform the analysis using the canonical Ricardian model of trade, but notice that their analysis carries to more general environments with non-CES utility and arbitrary neoclassical production functions.

way, our framework generalizes EK—while maintaining its tractability—and allows us to extend the results of Arkolakis et al. (2012) (henceforth, ACR) to capture the second aspect of the Ricardian idea.

We start by relating the technology structure of countries to a multivariate Fréchet distribution for productivity with an unrestricted correlation structure. Our main theoretical result, stated in Theorem 1, derives from what we call the *global innovation representation* of productivity. For each good, there is an (unbounded) collection of production technologies available. Each technology has a global productivity component and a bilateral-specific applicability component. The former component is common to all countries and captures the fundamental efficiency of the technology. The later component is unique to each country pair and captures idiosyncratic factors—such as trade frictions, domestic inefficiencies, and country endowments—that lead to differences in productivity. While for each good the global component follows a Poisson process, the applicability component is correlated across countries and i.i.d. across goods and technologies. We apply the spectral representation theorem for max-stable processes (De Haan, 1984; Penrose, 1992; Schlather, 2002) and establish that this global innovation representation of technology is *necessary and sufficient* for productivity to follow a multivariate Fréchet distribution.[2] The effective productivity that a given origin presents to a given destination country is distributed Fréchet across the continuum of goods, while dependence across countries is unrestricted.

Our global innovation representation of productivity, equivalent to assuming a multivariate Fréchet distribution, implies expenditure shares that match choice probabilities in generalized extreme value (GEV) discrete choice models (McFadden, 1978). In other words, there is an equivalence between the class of GEV import demand systems and the Ricardian model when productivity has a global innovation representation. The result is even stronger: The GEV class approximates *any* Ricardian model—without the need to restrict to Fréchet productivity distributions. The key implication is that *any* import demand system generated by the Ricardian trade model can be approximated using a global innovation representation for productivity. Put simply, our framework captures the full macroeconomic implications of Ricardian trade theory.

The generalization represented by GEV import demand systems allows us to calculate the gains from trade as a simple adjustment to the case of a CES demand system. We

[2] The spectral representation theorem for max-stable processes has previously been used in the decision theory context by Dagsvik (1994) to propose behavioral assumptions that justify max-stable random utility processes.

show that the results of ACR generalize, after a simple correction, to the class of models whose demand systems fit into the GEV form. In the Ricardian context, this correction adjusts a country's self-trade share to account for correlation in technology with the rest of the world, formalizing Ricardo's insight that more dissimilar countries have higher gains from trade.[3] Even in this more general framework, for any given pattern of correlation across countries, the adjusted self-trade share is calculated using only data on expenditure shares across countries, preserving the simplicity of ACR.

Perhaps not surprisingly, our general correlation structure also entails a different estimate of the trade elasticity. In fact, our model provides guidance on how to relate the macro substitution patterns implied by the observable trade flows to a given underlying micro-structure. In this way, we get guidance on how to use the micro data to discipline the estimates of macro parameters that capture the aggregate correlation structure. This is possible because our GEV structure is able to accommodate richer models based on EK, such as sectoral models (Costinot et al., 2012; Costinot and Rodrìguez-Clare, 2014; Caliendo and Parro, 2015; Ossa, 2015; Levchenko and Zhang, 2016), multinational production models (Ramondo and Rodríguez-Clare, 2013), global value chains models (Antràs and de Gortari, 2017), and models of trade with domestic geography (Fajgelbaum and Redding, 2014; Ramondo et al., 2016; Redding, 2016).[4]

To quantitatively evaluate the relevance of the correlation adjustment for the gains from trade, we use a multi-sector model of trade and assume that the macro correlation function takes a cross-nested CES form. We show that, in this case, key elasticities can be estimated by a simple two-step Ordinary Least Square (OLS) procedure using gravity and data on sectoral bilateral trade and bilateral trade costs.

Accounting for correlation in productivity can have substantial effects on estimates of the gains from trade, with gains almost doubling for some small countries. Crucially, the correlation correction delivers gains from trade that are much more heterogenous across

[3] Our framework is able to give an answer to the question posed by ACR (page 109): "A natural question at this point is whether there are many other Ricardian models, beyond Eaton and Kortum (2002), that satisfy R3 [the import demand system is CES]? The short answer is: "Probably not." [...] it is hard to imagine R3 holding in a Ricardian economy in the absence of very specific functional forms on the distribution of unit labor requirements." With our generalization to GEV demand systems, our answer is: "Yes, it is possible, once we adjust trade shares for correlation". Moreover, it is worth noting that correlation-adjusted trade shares constitute a gravity system.

[4] It is worth noting that multi-sector extensions of the EK model of trade (e.g., Caliendo and Parro, 2015) assume that the shape parameter θ is sector specific. The combination of independence with heterogenous shape parameters implies that the distribution of productivity across traded goods is not Fréchet. As a result, these sectoral models do not aggregate to an equivalent macro model with Fréchet distributed productivity and do not deliver a gravity system for aggregate bilateral trade flows.

countries with the same self-trade share—and within countries across time. We calculate that countries specialized in low correlations sectors have around 40 percent higher gains from trade relative to countries specialized in high correlation sectors. This result captures the second aspect of Ricardo's idea and is a reflection of the ever-evolving pattern of comparative advantage of countries, documented by Hanson et al. (2015). Reinforcing these results, our trade-liberalization exercises—which account for correlation and hence capture Ricardo's second insight—suggest that imports from China, rather than Canada, have the largest impact on real wages in the United States.

Our benchmark model is based on technology determining the patterns of trade across countries. From this supply-side point of view, substitution patterns come from the degree of technological (dis)similarity—i.e., correlation—between countries. More broadly, comparative advantage may come from demand-side factors as in the Armington model of trade (Anderson, 1979), and from entry of heterogenous firms (Krugman, 1980; Melitz, 2003). In the Appendix, we examine extensions of our framework that encompass these trade models. Similarly to ACR, these results make clear which assumptions on economic fundamentals lead to equivalence within a large and useful class of models.

Our work relates to several strands of the literature. First, we naturally make contact with the large trade literature using the Ricardian-EK framework in its various forms (see Eaton and Kortum, 2012, for a survey). More generally, our approach can be applied to any environment that requires Fréchet tools, with the potential of changing some of their quantitative conclusions. In particular, it can be applied to selection models used in the growth literature (such as Hsieh et al., 2013), and the macro development literature (such as Lagakos and Waugh, 2013), as well as to recent trade models used in the urban literature (such as Ahlfeldt et al., 2015; Monte et al., 2015; Caliendo et al., 2017), reviewed in Redding and Rossi-Hansberg (2017).

Second, we relate to papers in the international trade literature that use non-CES demand systems.[5] In particular, a recent paper by Adao et al. (2017) shows how to calculate macro counterfactual exercises in neoclassical trade models with invertible factor demand systems. In our paper, we provide necessary and sufficient conditions for a model's fundamentals to be distributed Fréchet, and, therefore, to fit into the GEV class—a subclass of models with the invertibility property. Our aggregation results allow us to relate various microstructures to the kind of macro counterfactuals they study.

[5] A related trade literature departs from CES with the goal of analyzing endogenous mark-ups and their effects on the gains from trade. See DeLoecker et al. (2016), Feenstra and Weinstein (2017), Bertoletti et al. (2017), and Arkolakis et al. (2017), among others.

Papers such as Caron et al. (2014), Brooks and Pujolas (2017), Lashkari and Mestieri (2016), Feenstra et al. (2017), and Bas et al. (2017), among others, also depart from CES demand systems. They aim, as we do, at showing the consequences of abandoning the assumptions that lead to linear gravity systems, and at incorporating more detailed micro data (i.e., sectoral trade) to estimate key model's elasticities. They all notice the failure of aggregate theories to incorporate the richness of the micro data (e.g., heterogeneous price and income elasticities across trade goods), and "fix it" by assuming non-CES demand systems.[6] By linking various micro structures to common primitives of technology, our general framework provides guidance—given by our aggregation results—on how to incorporate the micro estimates in this literature into macro counterfactual exercises.[7]

2 Ricardian Model

Consider a global economy consisting of N countries. Countries produce and trade in a continuum of product varieties $v \in [0,1]$. Consumers in all countries have identical constant elasticity of substitution (CES) preferences with (expenditure) elasticity of substitution $\sigma > -1$: $C_d = \left(\int_0^1 C_d(v)^{\frac{\sigma}{\sigma+1}} \mathrm{d}v\right)^{\frac{\sigma+1}{\sigma}}$. Given total expenditure of X_d by destination country d, their expenditure on variety v is $X_d(v) \equiv P_d(v)C_d(v) = (P_d(v)/P_d)^{-\sigma}X_d$ where $P_d(v)$ is the cost of the variety in terms of numeraire and $P_d = \left(\int_0^1 P_d(v)^{-\sigma}\mathrm{d}v\right)^{-\frac{1}{\sigma}}$ is the price level in country d.

We assume that the production function for varieties presents constant returns to scale in labor and depends on both the origin country o where the good gets produced and the destination market d where it gets delivered. For each $v \in [0,1]$, output $Y_{od}(v)$ satisfies

$$Y_{od}(v) = A_{od}(v)L_{od}(v). \tag{1}$$

[6] Caron et al. (2014) use a constant-relative-elasticity-of-income utility functions to link characteristics of goods in production to their characteristics in preferences, and in this way explain some "puzzles" observed in the data on trade patterns. Lashkari and Mestieri (2016) uses constant-relative-elasticity-of-income-and-substitution (CREIS) utility functions that allows for general patterns of correlations between income and price elasticity. Brooks and Pujolas (2017) analyze the expression for gains from trade arising from models with unrestricted utility functions (typically non-homothetic) that generate a non-constant trade elasticity. Feenstra et al. (2017) use a nested CES utility function to estimate micro and macro elasticities of substitution in a multi-sector model. Finally, Bas et al. (2017) break the Pareto assumption in the Melitz model of trade to get country-pair specific aggregate elasticities, which they estimate using sectoral-level trade data.

[7] In this regard, our paper shares a common theme with Redding and Weinstein (2017) as they develop a framework for aggregating from micro trade transactions to macro trade and prices using the whole class of invertible demand system.

$L_{od}(v)$ is the amount of labor used to produce variety v at origin o for delivery to d and $A_{od}(v)$ is the marginal product of labor—and referred as *productivity*.

As in EK, we capture heterogeneity in production possibilities by modeling productivity as a random draw. We focus on multivariate random variables which satisfy a property known as max stability. The EK model—built on independent Fréchet random variables—gets its tractability from this property. By relaxing their independence assumption, we get a flexible, yet tractable, model of trade.

2.1 Productivity as a Multivariate Fréchet Distribution

This section provides an introduction to multivariate extreme value type 2 (Fréchet) random variables with arbitrary dependence structure. The flexibility of this class of random variables allows us to capture Ricardo's second insight that the degree of technological similarity determines the gains from trade.

We begin by defining a multivariate Fréchet random vector.

Definition 1 (Multivariate Fréchet). *A random vector, $(A_1, \ldots, A_K)$, has a multivariate θ-Fréchet distribution if for any $\alpha_k \geq 0$ with $k = 1, \ldots, K$ the random variable $\max_{k=1,\ldots,K} \alpha_k A_k$ has a Fréchet distribution with shape parameter θ. In this case, the marginal distributions are Fréchet with (common) shape parameter θ and, for each $k = 1, \ldots, K$, satisfy*

$$\mathbb{P}\left[A_k \leq a\right] = \exp\left[-T_k a^{-\theta}\right], \tag{2}$$

for some scale parameter T_k.

This definition implies that a multivariate θ-Fréchet distribution is *max stable*—the maximum has the same marginal distribution up to scaling. This property holds for the distribution of productivity in the EK model and gives the model its tractability. Our setup includes as special cases EK, where productivity is independent across countries, and the symmetric multivariate Fréchet distribution used in Ramondo and Rodríguez-Clare (2013).

By working with the class of multivariate θ-Fréchet random variables, we can put minimal restrictions on dependence and maintain the key property of max-stability.[8] To make

[8] It is worth noting that the restriction to a common shape is necessary for max stability. General multivariate Fréchet distributions may have marginal distributions with different shape parameters, in which case the maximum—even with independence—is not distributed Fréchet.

headway without the independence assumption, we characterize the joint distribution of a multivariate Fréchet random variable by first defining a function that summarizes its correlation structure.

Definition 2 (Correlation Function). *A function $G : \mathbb{R}_+^K \to \mathbb{R}_+$ is a correlation function if it satisfies four restrictions:*

1. ***(Normalization).*** $G(0, \ldots, 0, 1, 0, \ldots, 0) = 1$*;*
2. ***(Homogeneity).*** *G is homogeneous of degree one, for any $\lambda \geq 0$ $G(\lambda x_1, \ldots, \lambda x_K) = \lambda G(x_1, \ldots, x_K)$;*
3. ***(Unboundedness).*** *$G(x_1, \ldots, x_K) \to \infty$ as $x_k \to \infty$ for any $k = 1, \ldots, K$; and*
4. ***(Differentiability).*** *The mixed partial derivatives of G exist and are continuous up to order K. The k'th partial derivative of G with respect to k distinct arguments is non-negative if k is odd and non-positive if k is even.*

This definition adds a normalization restriction to the definition of a social surplus function in GEV discrete choice models (McFadden, 1978).[9] This normalization restriction provides us with notation to distinguish between absolute advantage—captured by scale parameters—and comparative advantage—captured by a correlation function. Correlation functions reflect comparative advantage because they measure relative productivity levels across varieties and across origin countries within the same destination market.

Next, we characterize the joint distribution of any multivariate θ-Fréchet random variable in terms of the location parameters of its marginal distributions and a correlation function.

Lemma 1 (Correlation Function Representation). *The random vector $(A_1, \ldots, A_K)$ is multivariate θ-Fréchet if and only if there exists scale parameters T_k for $k = 1, \ldots, K$ and a correlation function G such that its joint distribution satisfies*

$$\mathbb{P}\left[A_k \leq a_k, \;\; k = 1, \ldots, K\right] = \exp\left[-G\left(T_1 a_1^{-\theta}, \ldots, T_K a_K^{-\theta}\right)\right]. \tag{3}$$

Proof. The result follows closely Theorem 3.1 of Smith (1984). See Appendix B. □

This standard result from probability theory allows us to parameterize joint distributions using scale parameters and correlation functions. The restrictions defining a correlation

[9] Correlation functions are often referred to as *tail dependence functions* or a *extremal index functions* in probability and statistics.

function ensure that (3) characterizes a valid multivariate extreme value type 2 (Fréchet) distribution.

Importantly, using the characterization in Lemma 1 and the homogeneity property of the correlation function, we get the max stability property. The maximum of a multivariate θ-Fréchet random variable is θ-Fréchet,

$$\mathbb{P}\left[\max_{k=1,\ldots,K} A_k \leq a\right] = \exp\left[-G\left(T_1,\ldots,T_K\right)a^{-\theta}\right]. \tag{4}$$

When evaluated at the scale parameters of the marginal distributions, the correlation function acts as an aggregator that returns the scale parameter of the maximum.

Moreover, max-stability, as in EK, entails that the conditional distribution of the maximum is identical to the unconditional distribution of the maximum,

$$\mathbb{P}\left[\max_{k'=1,\ldots,K} A_{k'} \leq a \mid A_k = \max_{k'=1,\ldots,N} A_{k'}\right] = \mathbb{P}\left[\max_{k'=1,\ldots,K} A_{k'} \leq a\right]. \tag{5}$$

This result is crucial for tractability in EK because it ensures that expenditure shares simply reflect the probability of importing from an origin country. Because this property holds for general multivariate Fréchet random variables, we inherit this tractability.[10]

To fix ideas, suppose that the vector of productivity across origins for delivering to destination d is independent across origins with Fréchet marginals—as in the EK model. Independence implies that the correlation function is additive,

$$\mathbb{P}[A_{1d}(v) \leq a_1,\ldots,A_{od}(v) \leq a_N] = \prod_{o=1,\ldots,N} \mathbb{P}[A_{od}(v) \leq a_o] = \exp\left(-\sum_{o=1}^{N} T_{od}a_o^{-\theta}\right).$$

The max-stability property also holds since

$$\mathbb{P}\left[\max_{o=1,\ldots,N} A_{od}(v) \leq a\right] = \exp\left[-\left(\sum_{o=1}^{N} T_{od}\right)a^{-\theta}\right].$$

An additive correlation function imposes a strong assumption, namely, that comparative advantages across countries are symmetric. Our framework relaxes the independence assumption of EK while maintaining the key max-stability property. By breaking this symmetry and allowing for heterogeneity in correlation, our model captures heterogeneity in

[10]Appendix A formally presents this and other useful properties of Fréchet random variables which we use throughout the paper.

comparative advantage, and, as we show in Section 4.1, allows us to formalize Ricardo's second insight that the degree of technological similarity matters for the gains from trade.

2.2 Global Innovation Representation of Productivity

We next present a structure for technology that is *necessary and sufficient* for productivity to be distributed multivariate Fréchet. This structure can be interpreted as resulting from adopting technologies, which are a product of global innovations, based on a country's ability to apply each innovation. We abstract from micro details on how this adoption occurs. However—in light of our aggregation results in Section 5—we can interpret this macro model as an aggregated version of some underlying micro model.

Our main result in Theorem 1 characterizes multivariate Fréchet distributions as precisely those productivity distributions arising from global innovation. To present this result, we consider a technology structure that satisfies three assumptions.

Assumption 1 (Innovation Decomposition). *For each v, there exists a countable set of global innovations $\{Z_i(v)\}_{i=1,2,\dots}$ and bilateral applicability $\{\{A_{iod}(v)\}_{o=1}^N\}_{i=1,2,\dots}$ such that*

$$A_{od}(v) = \max_{i=1,2,\dots} Z_i(v)A_{iod}(v). \tag{6}$$

Assumption 2 (Independence). $\{\{A_{iod}(v)\}_{o=1}^N\}_{i=1,2,\dots}$ *is independent of* $\{Z_i(v)\}_{i=1,2,\dots}$ *and i.i.d. over* (i, v).

Assumption 3 (Poisson Innovations). *There exists $\theta > 0$ such that the collection $\{Z_i(v)^\theta\}_{i=1,2,\dots}$ consists of the points of a Poisson process with intensity measure $z^{-2}dz$, and is i.i.d. over v.*

Assumption 1 defines a structure for technology that can be interpreted as arising from global innovation and technology adoption. For each good v, there is a countable collection of technological innovations $i = 1, 2, \dots$ influencing the marginal product of labor. These innovations represent physical techniques (i.e. blueprints) for producing a good. Each innovation i has global productivity $Z_i(v)$ and an origin-destination specific *applicability* component $A_{iod}(v)$. Its global productivity measures the fundamental efficiency of the production technique and is identical across all origins and destinations. In turn, applicability captures origin-destination specific factors that determine the efficiency of the technique when adopted at origin o to deliver goods to destination d.

The key aspect of Assumption 2 is that it does not impose independence of applicability across origin countries; instead, it allows for arbitrary patterns of spatial correlation.

Assumption 3 implies that for any interval $(a, b]$, the number of innovations whose global productivity satisfies $a < Z_i(v) \leq b$ is a Poisson random variable. It also implies that the distribution of $Z_i(v)$ is Pareto with shape parameter $\theta \times i$.[11] One can interpret this assumption as arising from a random discovery process as in Eaton and Kortum (1999, 2010). Over time, innovations arrive at a Poisson rate and each innovation represents a new production technique. The fundamental efficiency of each new technique—its global productivity—is distributed Pareto. The key distinction here is that rather than assuming innovations are country specific, we capture the idea that innovations represent physical methods to produce a good and are globally applicable. Origin countries differentially load on global productivity via their individual draw of applicability, $A_{iod}(v)$, and adopt whichever innovation is most efficient for them.

Theorem 1 (Global Innovation Representation). *Productivity, $\{A_{od}(v)\}_{o=1}^{N}$, is multivariate θ-Fréchet if and only if it satisfies Assumptions 1, 2, and 3. In this case, we say that productivity has a global innovation representation.*

Let $\{A_{iod}(v)\}_{o=1}^{N}$ denote underlying applicability. Then the joint productivity distribution satisfies

$$\mathbb{P}\left[A_{1d}(v) \leq a_1, \ldots, A_{Nd}(v) \leq a_N\right] = \exp\left[-G^d\left(T_{1d}a_1^{-\theta}, \ldots, T_{Nd}a_N^{-\theta}\right)\right], \tag{7}$$

where $T_{od} \equiv \mathbb{E}A_{iod}(v)^{\theta}$ for $o = 1, \ldots, N$ are the scale parameters of the marginal distributions and the correlation function is

$$G^d(x_1, \ldots, x_N) \equiv \mathbb{E} \max_{o=1,\ldots,N} \frac{A_{iod}(v)^{\theta}}{T_{od}} x_o. \tag{8}$$

Proof. The equivalence is a consequence of the spectral representation theorem for max-stable processes (De Haan, 1984; Penrose, 1992; Schlather, 2002). See Appendix C. □

This characterization of productivity establishes primitive assumptions on global technology that are *necessary and sufficient* for Fréchet-distributed productivity across origin countries. Fréchet-distributed productivity can always be interpreted as arising from the applicability of global technologies. Intuitively, both absolute advantage (the scale parameters) and comparative advantage (the correlation functions) arise from the adoption of technology and patterns of adoption depend on the ability of exporters to apply inno-

[11] We have $\mathbb{P}[Z_i(v) > z] = \mathbb{P}[Z_i(v)^{-\theta} < t]$ for $t = z^{-\theta}$. Then, since $\{Z_i(v)^{\theta}\}_{i=1,2,\ldots}$ are the points of a Poisson process with intensity measure $z^{-2}\mathrm{d}z$, $\{Z_i(v)^{-\theta}\}_{i=1,2,\ldots}$ are the points of a homogenous poisson process with intensity of 1, and $\mathbb{P}[Z_i(v)^{-\theta} < t] = \sum_{j=i}^{\infty} \frac{t^j}{j!}e^{-t} = \frac{t^i}{i!} = \frac{z^{-\theta i}}{i!}$. Therefore, $\mathbb{P}[Z_i(v) \leq z] = 1 - \frac{z^{-\theta i}}{i!}$.

vations.[12] The result also provides a method to compute scale parameters and correlation functions of multivariate Fréchet random variables: They are simply moments of bilateral applicability.

Concretely, assume that the applicability of individual technologies follows a multivariate Fréchet distribution across origin countries with symmetric correlation, as in Ramondo and Rodríguez-Clare (2013). Then, the joint distribution of applicability is

$$\mathbb{P}[A_{i1d}(v) \leq a_1, \ldots, A_{iNd}(v) \leq a_N] = \exp\left[- \left(\sum_{o=1}^{N} (T_{od} a_o^{-\theta})^{1/(1-\rho)} \right)^{1-\rho} \right],$$

where each T_{od} is an origin-destination specific scale parameter and $0 \leq \rho < 1$ parameterizes correlation in applicability across origin countries. We can compute the correlation function for this example using the max-stability property of multivariate Fréchet distributions in (4). Specifically, for any given vector $(a_1, \ldots, a_N)$, the random variable $\max_{o=1,\ldots,N} A_{iod}(v)/a_o$ must be θ-Fréchet with scale $[\sum_{o=1}^{N} (T_{od} a_o^{-\theta})^{1/(1-\rho)}]^{1-\rho}$. Since the θ-moment of a θ-Fréchet random variable is given by its scale parameter, the correlation function is

$$G^d(x_1, \ldots, x_N) = \left[\sum_{o=1}^{N} \left(T_{od} \frac{x_o}{T_{od}} \right)^{\frac{1}{1-\rho}} \right]^{1-\rho} = \left(\sum_{o=1}^{N} x_o^{\frac{1}{1-\rho}} \right)^{1-\rho}.$$

This correlation function takes the form of a CES aggregator with the correlation parameter ρ determining the elasticity of substitution. Notice that, in this particular example, the joint distribution of productivity is identical to the joint distribution of applicability. Also, when $\rho = 0$, we get an additive correlation function, which corresponds to the case of independence as in EK; and, as $\rho \to 1$, productivity draws become identical across countries.

This example, which nests the EK model as a special case, is a useful building block for generating correlation functions associated with richer Fréchet distributions. We can do so by simply using the max-stability property and Theorem 1.

Consider a latent (within country) technology adoption decision. In particular, suppose that applicability of technology i comes from a choice of how to apply innovation i. Let this choice amount to selecting an application m across M alternatives so that $A_{iod}(v) = \max_{m=1,\ldots,M} A_{imod}(v)$. Assume that each application is independent and, under the m'th application, applicability follows a symmetric multi-variate Fréchet distribution across

[12] This global innovation representation connects our static framework to dynamic models of innovation and knowledge diffusion, such as Buera and Oberfield (2016).

origins,

$$\mathbb{P}[A_{im1d}(v) \leq a_1, \ldots, A_{imNd}(v) \leq a_N] = \exp\left[-\left(\sum_{o=1}^{N} (T_{mod} a_o^{-\theta})^{1/(1-\rho_m)}\right)^{1-\rho_m}\right],$$

where the parameter $0 \leq \rho_m < 1$ captures the similarity in applicability across countries under application m. To calculate the correlation function, we need two intermediate results. First, due to independence across m, the marginal distributions of $A_{iod}(v)$ are θ-Fréchet with scale $T_{od} = \sum_{m=1}^{M} T_{mod}$. Second, for a fixed m, the random variable $\max_{o=1,\ldots,N} A_{imod}(v)$ is θ-Fréchet with scale $\left(\sum_{o=1}^{N} T_{mod}^{1/(1-\rho_m)}\right)^{1-\rho_m}$. Independence across m then implies that the correlation function is

$$G^d(x_1, \ldots, x_N) = \mathbb{E} \max_{m=1,\ldots,M} \max_{o=1,\ldots,N} A_{imod}(v)^\theta \frac{x_o}{T_{od}} = \sum_{m=1}^{M} \left(\sum_{o=1}^{N} (\omega_{mod} x_o)^{1/(1-\rho_m)}\right)^{1-\rho_m}, \quad (9)$$

where $\omega_{mod} \equiv T_{mod}/T_{od} = T_{mod}/\sum_{m'=1}^{M} T_{m'od}$. The first equality follows from interchanging the max over o and m in the definition of the correlation function, while the second equality uses the max-stability property in (4). This example shows how to build tractable correlation functions from specific assumptions on the structure of technological applicability. In this specific case, we get a cross-nested CES form by incorporating a latent choice over the application of innovation i.[13] In Section 5, we present our general aggregation results.

Summing up, Theorem 1 allows us to relate alternative parametric specifications for the correlation function to underlying primitive assumptions on the nature of technological applicability. When building models based on Fréchet-distributed productivity, one can either use a particular specification for applicability—possibly arising from a model of innovation—and derive the implied G^d using Theorem 1, or, alternatively, directly specify a correlation function satisfying the restrictions in Definition 2.

2.3 Prices and Trade Shares

We now proceed to characterize import price distributions and expenditure shares under the assumption that productivity has a global innovation representation.

[13] In a recent paper, Lashkaripour and Lugovskyy (2017) use a nested Fréchet, with latent factors representing firm-specific quality, to generalize EK. Their example fits into the productivity representation in our Theorem 1.

The minimum marginal cost to deliver a particular variety v to destination d from origin o is

$$c_{od}(v) = \frac{W_o}{A_{od}(v)}, \tag{10}$$

where W_o is the nominal wage in country o. To map our setup into standard variables in the trade literature, we further set notation to separate out different components of productivity. Define origin country o's *productivity index* as

$$A_o \equiv T_{oo}^{1/\theta}, \tag{11}$$

and the *iceberg trade cost* from country o to d as

$$\tau_{od} \equiv \left(\frac{T_{oo}}{T_{od}}\right)^{1/\theta}. \tag{12}$$

The index A_o measure a country's ability to produce goods in their domestic market, while τ_{od} measures efficiency losses associated with delivering goods to market d.

With perfect competition, potential import prices for variety v produced in o and delivered to d equal its marginal cost, $P_{od}(v) = c_{od}(v)$. It is notationally convenient to work with an *import price index* defined as

$$P_{od} \equiv \frac{\tau_{od} W_o}{A_o}. \tag{13}$$

Then the joint distribution of potential import prices follows from Theorem 1.

Proposition 1 (Potential Import Price Distribution). *Suppose productivity has a global innovation representation, and markets are perfectly competitive. Then the joint distribution of prices presented to destination market d is a multivariate Weibull distribution satisfying*[14]

$$\mathbb{P}[P_{1d}(v) \geq p_1, \ldots, P_{Nd}(v) \geq p_N] = \exp\left[-G^d\left(P_{1d}^{-\theta} p_1^{\theta}, \ldots, P_{Nd}^{-\theta} p_N^{\theta}\right)\right].$$

Proof. This result follows directly from Theorem 1. See Appendix D. □

The joint distribution of productivity determines the joint distribution of potential import prices. For each origin o, the marginal distribution of prices, $\mathbb{P}[P_{od}(v) \leq p] = 1 -$

[14] A *multivariate Weibull random variable* is a random vector $(B_1, \ldots, B_K)$ whose marginal distributions are Weibull: $\mathbb{P}[B_k \leq b] = 1 - e^{-S_k b^{\theta_k}}$ for some scale $S_k > 0$ and shape $\theta_k > 0$ across $k = 1, \ldots, K$. Note that if $(A_1, \ldots, A_k)$ is θ-Fréchet, then the vector $(B_1^{-1}, \ldots, B_K^{-1})$ is multivariate Weibull and its marginals have common shape $\theta_k = \theta$ for each $k = 1, \ldots, K$.

$\exp\left[-P_{od}^{-\theta}p^{\theta}\right]$, is a Weibull distribution with scale parameter determined by the import price index, P_{od}, and shape parameter of θ. The dependence structure of prices is summarized by the correlation function G^d.

Given the distribution of potential import prices, a country imports each variety from the cheapest source. The previous characterization of the potential import price distribution leads to the following closed-form results, which generalize Eaton and Kortum (2002).

Proposition 2 (Generalized EK). *Suppose productivity has a global innovation representation with $\theta > \sigma$, and markets are perfectly competitive. Then:*

1. *The share of varieties that destination d imports from o is*

$$\pi_{od} = \frac{P_{od}^{-\theta} G_{od}}{G^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})} \tag{14}$$

where

$$G_{od} \equiv G_o^d\left(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta}\right) \equiv \frac{\partial G^d\left(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta}\right)}{\partial P_{od}^{-\theta}}; \tag{15}$$

2. *The distribution of prices among goods imported into country d from o is*

$$\mathbb{P}[P_{od}(v) \geq p \mid P_{od}(v) \leq P_{o'd}(v) \quad \forall o' \neq o] = \exp\left[-G^d\left(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta}\right)p^{\theta}\right]; \tag{16}$$

3. *Total expenditure by country d on goods from country o is*

$$X_{od} = \pi_{od} X_d; \text{ and} \tag{17}$$

4. *The price index in country d is*

$$P_d = \gamma G^d\left(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta}\right)^{-\frac{1}{\theta}}, \tag{18}$$

where $\gamma = \Gamma\left(\frac{\theta-\sigma}{\theta}\right)^{-\frac{1}{\sigma}}$ and $\Gamma(\cdot)$ is the gamma function.

Proof. Follows from proposition 1 and the general properties of Fréchet random variables established in Appendix A. See Appendix E. □

First, the formula for the expenditure share, π_{od}, takes the same form as choice probabilities in GEV discrete choice models (McFadden, 1978), with the import price index taking the place of choice-specific utility. As a result, this model is equivalent at the aggregate level to any trade model whose implied import demand system belongs to the GEV class. We define this class of models and formalize this equivalence in Section 3.

Second, using (14), correlation-adjusted expenditure shares, defined as $\pi^*_{od} \equiv \pi_{od}/G_{od}$, constitute a gravity system (as defined by ACR). Specifically, correlation-adjusted expenditure is CES:

$$\pi^*_{od} = \left(\gamma \frac{P_{od}}{P_d}\right)^{-\theta} \quad \Longrightarrow \quad \ln \pi^*_{od} = S_o - D_d - \theta \ln \tau_{od}, \tag{19}$$

where $S_o \equiv \theta \ln(A_o/W_o)$ and $D_d \equiv \ln(P_d/\gamma)$.

Third, as in EK, the distribution of prices among goods actually imported into market d is identical to the distribution of potential import prices. As a result, we get that expenditures shares are equal to the share of varieties imported into d from o. This result follows from the property that the conditional distribution of the max of a multivariate Fréchet random vector is identical to its unconditional distribution (see Lemma A.3).

Finally, the price level in each destination market is determined by the correlation function, G^d. In the trade context, this function can be interpreted as an aggregator that defines the welfare-relevant price index. In analogy to the discrete choice literature, welfare calculations depend crucially on the specification of this function.[15]

3 GEV Import Demand Systems

Which macro substitution patterns can be rationalized by this theory? To answer this question, we first establish, in Corollary 1, that the Ricardian model with multivariate Fréchet productivity implies expenditure shares that match choice probabilities in GEV discrete choice models (McFadden, 1978). Dagsvik (1995) shows that GEV random utility models are dense in the space of all random utility models. We adapt this result to our context—which differs because variety-level demand is CES—and establish that the demand systems generated by the Ricardian model with Fréchet distributed productivity can approximate any demand system generated by heterogenous (i.e. stochastic) productivity. That is, our framework allows us to describe *all* import demand systems consistent with Ricardian trade, under the assumptions of constant returns to scale in production,

[15] Note that, unlike the discrete choice literature, the central distributional restriction captured in Assumption 3 is on a cardinal rather than an ordinal quantity. Unlike latent utility, productivity is, in principle, a measurable quantity.

competitive markets, and a single productive factor in each country.

First, we define an *import demand system* for destination d as a collection of expenditure share functions $\{\pi_{od}\}_{o=1}^{N}$ such that for each $o = 1, \ldots, N$ the function $\pi_{od} : \mathbb{R}_+^N \times \mathbb{R}_+ \to [0,1]$ is homogenous of degree zero and for any vector of import prices $\mathbf{P}_d \equiv (P_{1d}, \ldots, P_{Nd}) \in \mathbb{R}_+^N$ and level of expenditure $X_d \geq 0$, $\sum_{o=1}^{N} \pi_{od}(\mathbf{P}_d, X_d) = 1$.

Next, we define the class of GEV import demand systems.

Definition 3 (GEV Import Demand System). *A generalized extreme value (GEV) import demand system for destination d is an import demand system such that there exists a shape parameter $\theta > 0$, and a correlation function G^d satisfying*

$$\pi_{od}(\mathbf{P}_d, X_d) = \frac{P_{od}^{-\theta} G_o^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})}{G^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})}, \tag{20}$$

for all $o = 1, \ldots, N$.

As mentioned above, this specification for expenditure shares is closely related to the functional form for choice probabilities in GEV discrete choice models. It differs slightly in that our correlation function is a restricted version of the social surplus function in GEV models due to our normalization restriction in Definition 2. Note that the GEV class of import demand systems is homothetic since expenditure shares do not depend on overall expenditure.

An important class of models within the GEV class are CES import demand systems, as in EK and ACR. These models come from specifying an additive correlation function, $G^d(x_1, \ldots, x_N) = \sum_{o=1}^{N} x_o$, and have import demand shares of the form

$$\pi_{od}(\mathbf{P}_d, X_d) = \frac{P_{od}^{-\theta}}{\sum_{o'=1}^{N} P_{o'd}^{-\theta}}.$$

These models lead to a gravity system at the macro level and include workhorse models of trade, such as Armington, Melitz, and EK (see Arkolakis et al., 2012).

The GEV class, however, is much larger than the CES class. For example, consider the cross-nested CES specification for the correlation function that we derived in Section 2.2:

$$G^d(x_1, \ldots, x_N) = \sum_{m=1}^{M} \left(\sum_{o=1}^{N} (\omega_{mod} x_o)^{1/(1-\rho_m)} \right)^{1-\rho_m}. \tag{21}$$

The parameter $0 \leq \rho_m < 1$ measures correlation across origin countries in technological applicability under application m, while ω_{mod} measures the efficiency of application m (relative to alternative applications) for origin o when delivering to destination d. The import demand system implied by this correlation function is

$$\pi_{od}(\mathbf{P}_d, X_d) = \sum_{m=1}^{M} \left(\frac{P_{mod}}{P_{md}}\right)^{-\frac{\theta}{1-\rho_m}} \frac{P_{md}^{-\theta}}{\sum_{m'=1}^{M} P_{m'd}^{-\theta}}, \tag{22}$$

where $P_{mod} \equiv \omega_{mod}^{-\frac{1}{\theta}} P_{od}$, and (with a slight abuse of notation)

$$P_{md} \equiv \left[\sum_{o=1}^{N} (P_{mod})^{-\frac{\theta}{1-\rho_m}}\right]^{-\frac{1-\rho_m}{\theta}}. \tag{23}$$

The first fraction on the right-hand side represents the probability of importing from country o given that it chose application m. Due to correlation in applicability across countries under m, this conditional probability has an elasticity of substitution of $-\theta/(1-\rho_m)$ in origin o's price for m-application goods relative to the destination market price index P_{md}.[16] The second fraction on the right-hand side represents the probability of importing m-goods from any country. Due to independence in applicability across m's, the elasticity of substitution between them is simply $-\theta$. This example further shows that we can build macro-level import demand systems from underlying micro-foundations for technological applicability, a result that we formalize in Section 5.

Returning to the general GEV formulation in (20), if we take import price indices as being $P_{od} = \tau_{od} W_o / A_o$, the import demand system implies expenditure shares as in (14) in Proposition 2; that is, they match the Ricardian model in which productivity has a global innovation representation. Together with Theorem 1, we get the following result.

Corollary 1 (GEV Equivalence). *The class of GEV import demand systems is identical to the set of import demand systems generated by assuming that productivity has a global innovation representation.*

The GEV class is generated by correlation functions, and correlation functions can always be constructed using the global innovation representation in Theorem 1. In other words, the GEV class is equivalent to the Ricardian model when productivity has a global innovation representation.

Importantly, a direct implication is that the Ricardian model can rationalize many existing

[16] As correlation in applicability under m becomes perfect ($\rho_m \to 1$), countries become perfectly substitutable at delivering goods produced under m.

trade models that fall into the GEV class of import demand systems. For example, sectoral models (Costinot et al., 2012; Costinot and Rodrìguez-Clare, 2014; Caliendo and Parro, 2015; Ossa, 2015; Levchenko and Zhang, 2016), multinational production models (Ramondo and Rodríguez-Clare, 2013; Tintelnot, 2017), global value chains models (Antràs and de Gortari, 2017), and models of trade with domestic geography (Fajgelbaum and Redding, 2014; Ramondo et al., 2016; Redding, 2016) are all cases of a cross-nested CES import demand system.[17] In our empirical application in Section 6 we also use a cross-nested CES specification and interpret the latent factor m as a sector.

In fact, we can push the result in Corollary 1 one step further by adapting results from the discrete choice literature. We know that GEV random utility models are dense in the space of all random utility models (Dagsvik, 1995). This result for choice probabilities does not directly apply since our model features CES demand at the variety level. However, an analogous result holds. The set of import demand systems generated by *any* Ricardian model—without restricting to Frèchet productivity distributions—can be approximated arbitrarily well by the class of GEV import demand systems.

Proposition 3 (GEV Approximation). *Let $\{A_{od}(v)\}_{o=1}^{N}$ be a multivariate random variable for productivity whose marginals have finite moment of order σ. Then for price indices defined by $P_{od} \equiv \mathbb{E}[(W_o/A_{od}(v))^{-\sigma}]^{-\frac{1}{\sigma}}$, the import demand system generated by the Ricardian model is*

$$\pi_{od}(\mathbf{P}_d, X_d) \equiv \frac{\mathbb{E}\left[(P_{od}/U_{od}(v))^{-\sigma}\mathbf{1}\left\{P_{od}/U_{od}(v) = \min_{o'=1,\ldots,N} P_{o'd}/U_{o'd}(v)\right\}\right]}{\mathbb{E}\left[\left(\min_{o'=1,\ldots,N} P_{o'd}/U_{o'd}(v)\right)^{-\sigma}\right]}$$

for $U_{od}(v) \equiv A_{od}(v)/\mathbb{E}[A_{od}(v)^{\sigma}]^{1/\sigma}$.

Also, for any compact $K \subset \mathbb{R}_{+}^{N}$ and any $\epsilon > 0$, there exists a shape parameter $\theta > \sigma$, and

[17] Consider the model of multinational production in Ramondo and Rodríguez-Clare (2013). Relabel the latent factor dimension m of the cross-nested CES in (21) to correspond to the home country i of the multinational producing a given good. Define productivity and multinational production inefficiency indices T_i and h_{io} such that $\omega_{iod} \equiv T_i^{1-\rho_i}(h_{io}A_o)^{-\theta}$. Then,

$$G^d(x_1,\ldots,x_N) = \sum_{i=1}^{N} T_i \left(\sum_{o=1}^{N} (h_{io}A_o)^{-\theta/(1-\rho_i)} x_o^{1/(1-\rho_i)}\right)^{1-\rho_i}.$$

Define the cost index for multinationals from home country i as $c_{id} = \left(\sum_{o=1}^{N}(h_{io}W_o\tau_{od})^{-\theta/(1-\rho_i)}\right)^{-(1-\rho_i)/\theta}$. The expenditure share on goods produces in o for d is

$$\pi_{od} = \sum_{i=1}^{N} \pi_{iod} \equiv \sum_{i=1}^{N} \frac{T_i c_{id}^{-\theta}}{\sum_{i'=1}^{N} T_{i'} c_{i'd}^{-\theta}} \frac{(h_{io}W_o\tau_{od})^{-\theta/(1-\rho_i)}}{\sum_{o'=1}^{N}(h_{io'}W_{o'}\tau_{o'd})^{-\theta/(1-\rho_i)}}.$$

This import demand system matches the one of Ramondo and Rodríguez-Clare (2013) for $\rho_i = \rho$, for all i.

correlation function G^d such that

$$\sup_{\mathbf{P}_d \in K} \left| \pi_{od}(\mathbf{P}_d, X_d) - \frac{P_{od}^{-\theta} G_o^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})}{G^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})} \right| < \epsilon.$$

Proof. The proof is constructive and based on choosing a correlation function for some $\theta > \sigma$ of the form: $G^d(x_1, \ldots, x_N) = \left[\mathbb{E}\left(\sum_o (U_{od}(v)^\theta x_o)\right)^{\frac{\sigma}{\theta}}\right]^{\frac{\theta}{\sigma}}$. Note that the price level implied by this correlation function approximates the true price level because

$$P_d = \Gamma((\theta - \sigma)/\theta) G^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})^{-\frac{1}{\theta}} \overset{\theta \to \infty}{\to} \left[\mathbb{E} \min_o \left(P_{od}/U_{od}(v)\right)^{-\sigma}\right]^{-\frac{1}{\sigma}}$$

point wise. The result follows from establishing that the import demand system associated with this correlation function converges uniformly to the true demand system. See Appendix F. □

The key implication is that *any* import demand system generated by the Ricardian trade model can be approximated by a Ricardian trade model where productivity has a global innovation representation. In fact, the class of cross-nested CES models—which we use in our empirical application—can approximate any GEV model (Fosgerau et al., 2013) and by extension any Ricardian model.[18] Put simply, our framework encompasses the full macroeconomic implications of Ricardian trade theory.

4 Macro Counterfactuals

We show that heterogeneity in correlation leads to heterogeneity in the gains from trade and how it affects the calculation of any (counterfactual) departure from the current equilibrium. It turns out that calculations for a GEV demand system are virtually identical, after a correction for correlation, to the calculations in ACR for trade models with CES import demand systems. The correlation correction only requires data on expenditure shares across countries, preserving the simplicity of the ACR gains from trade calculation.

Using the expression for the price index in (18), we can write the real wage in each country

[18] In that regard, the cross-nested CES specification can arbitrarily approximate the mixed-CES import demand system used in Adao et al. (2017). Their empirical application can be interpreted as arising from some underlying Ricardian model.

as

$$\frac{W_d}{P_d} = \gamma^{-1} W_d G^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})^{\frac{1}{\theta}}. \tag{24}$$

From (14), the self-trade share is

$$\pi_{dd} = \frac{(W_d/A_d)^{-\theta} G_{dd}}{G^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})}, \tag{25}$$

from which we can write the real wage in country d as

$$\frac{W_d}{P_d} = \gamma^{-1} A_d \left(\frac{\pi_{dd}}{G_{dd}}\right)^{-\frac{1}{\theta}}. \tag{26}$$

Let $\hat{x} \equiv x'/x$ denote the change from x to x' in an equilibrium outcome due to some change in fundamentals. Using (26), it is straightforward to show that the change in real wages between two equilibria is given by

$$\frac{\hat{W}_d}{\hat{P}_d} \equiv \frac{W_d'/P_d'}{W_d/P_d} = (\hat{\pi}_{dd}^*)^{-\frac{1}{\theta}}, \tag{27}$$

where $\pi_{dd}^* \equiv \pi_{dd}/G_{dd}$ is the correlation-adjusted trade share. That is, for any trade model that implies a GEV import demand system, a (log) change in equilibrium real wages—triggered by some shock to the model's parameters—is proportional to the (log) change in the correlation-adjusted self-trade share, with the factor of proportionally given by the trade elasticity.

4.1 Gains From Trade: Autarky

What are the consequences of correlation in technology for the gains from trade relative to autarky? Intuitively, if two countries have identical idiosyncratic productivity draws across varieties, with their average productivity determining the cost of labor, they will offer each other identical prices across varieties, and there is no scope for trade between them. This intuition captures the second part of Ricardo's insight —-countries with similar production possibilities gain less from trading with each other.

In autarky, country d purchases only their own goods and so $\pi_{dd} = 1$. Moreover, as $\tau_{od} \to \infty$, $P_{od}^{-\theta} \to 0$ for $o \neq d$. As a result, $G_{dd} = 1$—intuitively, correlation with other countries is irrelevant in autarky. Therefore, real wages in autarky are

$$\left(\frac{W_d}{P_d}\right)^{\text{Autarky}} = \gamma^{-1} A_d. \tag{28}$$

Comparing the real wage in (26) to this counterfactual autarky real wage gives the following result for the gains from trade.

Proposition 4 (**Gains From Trade**)**.** *Consider a trade model that implies a GEV import demand system with import price indices satisfying $P_{od} = \tau_{od} W_o / A_o$. Then the gains from trade relative to autarky are*

$$GT_d \equiv \frac{W_d/P_d}{(W_d/P_d)^{Autarky}} = \left(\frac{\pi_{dd}}{G_{dd}}\right)^{-\frac{1}{\theta}}. \tag{29}$$

Proof. Simply divide (26) by (28). □

This proposition generalizes the results of ACR to the class of models with GEV demand systems. Yet, the simplicity of ACR is preserved after adjusting self-trade shares for correlation in technology.

This correlation correction is related to the elasticity of country d's price index to their own wages and related to their real marginal cost of production. Using (14), the self-trade share is just the elasticity of a country's price index to their own wages,

$$\pi_{dd} = \frac{\partial \ln G^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})}{\partial \ln P_{dd}^{-\theta}} = \frac{\partial \ln P_d^{-\theta}}{\partial \ln (W_d/A_d)^{-\theta}} = \frac{\partial \ln P_d}{\partial \ln W_d}$$

which is not surprising given that the correlation function acts as a price aggregator and determines a country's price level, as specified in (18). As a result, this elasticity is linked to the real marginal cost of production and the correlation-correction term via self trade:

$$\frac{\partial \ln P_d}{\partial \ln W_d} = \pi_{dd} = \frac{P_{dd}^{-\theta} G_{dd}}{G^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})} = \left(\gamma \frac{W_d/A_d}{P_d}\right)^{-\theta} G_{dd}.$$

With a CES demand system, such as in the EK model with zero correlation in technology, $G_{dd} = 1$, and the gains from trade in (29) simplify to the ones in ACR: Two countries with the same self-trade share have the same gains from trade relative to autarky. The restriction to uncorrelated technology imposes that a country's real marginal cost of $(W_d/A_d)/P_d$ is log-proportional to the elasticity of prices to wages $\partial \ln P_d / \partial \ln W_d$. It is this restriction—summarized by $G_{dd} = 1$—which leads to the ACR result for the gains from trade. Allowing for correlation in technology breaks this tight link, and G_{dd} is precisely the quantity needed to calculate the gains from trade without imposing this structural relationship between the elasticity of prices to wages and a country's real marginal cost of production.

The expression in (29) suggests that if a country has very similar technology to all other countries—i.e., high correlation—the gains from trade will be small—and smaller than the ones implied by only considering its self-trade share. In contrast, if its technology is very dissimilar to other countries—i.e., low correlation—the gains from trade will be large—and larger than the gains implied by only considering self-trade shares. With a simple correlation correction to self-trade shares, our general framework captures Ricardo's second insight on the heterogeneity of gains from trade across countries.

To gain intuition on how heterogenous correlation in technology implies heterogeneity in the gains from trade, consider a three-country world with correlation function given by

$$G^d(x_1, x_2, x_3) = \left(x_1^{1/(1-\rho)} + x_2^{1/(1-\rho)}\right)^{1-\rho} + x_3,$$

which implies that the joint distribution of productivity across countries is

$$\mathbb{P}[A_{1d}(v) \leq a_1, A_{2d}(v) \leq a_2, A_{3d}(v) \leq a_3] = \exp\left[-\left((T_{1d}a_1^{-\theta})^{\frac{1}{1-\rho}} + (T_{2d}a_2^{-\theta})^{\frac{1}{1-\rho}}\right)^{1-\rho} + T_{3d}a_3^{-\theta}\right].$$

We think of countries 1 and 2 as technological peers, with the parameter ρ measuring the degree of correlation in their technology. Country 3's productivity is uncorrelated with productivity in countries 1 and 2. Then,

$$G_{od} = \left(P_{1d}^{-\theta/(1-\rho)} + P_{2d}^{-\theta/(1-\rho)}\right)^{-\rho} P_{od}^{-\theta\rho/(1-\rho)} \quad \text{for} \quad o = 1, 2 \quad \text{and} \quad G_{3d} = 1.$$

Given that $\pi_{od} = P_{od}^{-\theta} G_{od}/G^d(P_{1d}^{-\theta}, P_{2d}^{-\theta}, P_{3d}^{-\theta})$, we can take the ratio $G_{1d}/G_{2d} = (\pi_{1d}/\pi_{2d})^{\rho}$, which implies that

$$G_{od} = \left(\frac{\pi_{od}}{\pi_{1d} + \pi_{2d}}\right)^{\rho} \quad \text{for} \quad o = 1, 2.$$

As a result, the gains from trade are

$$GT_d = \left[\pi_{dd}^{1-\rho}(\pi_{1d} + \pi_{2d})^{\rho}\right]^{-\frac{1}{\theta}} \quad \text{for} \quad d = 1, 2 \quad \text{and} \quad GT_3 = \pi_{33}^{-\frac{1}{\theta}}.$$

The gains from trade for countries 1 and 2 depend on the degree of correlation in technology, while the gains from trade for country 3 are pinned down by their self-trade share. The corrected self-trade shares for country 1 and 2 end up being a Cobb-Douglas combination—with weight given by ρ—between each country's expenditure share on their own goods and on the combination of their own goods with their peer's goods. The later can be interpreted as the self-trade share if countries 1 and 2 were combined into a single country. When correlation in technology is zero ($\rho = 0$), a correlation correction is unnecessary; for positive correlation, the correction increases effective self-trade and implies lower gains from trade; and for perfect correlation ($\rho = 1$), the two countries are effectively a single country and it is their combined self trade that is relevant for calculating the gains from trade.

4.2 Calculating the Correlation Correction

To make the necessary adjustment for correlated technology, we need to know the correlation structure across countries — which requires estimation of G^d. Given the correlation function, we can then calculate the gains from trade directly from expenditure data. The procedure requires solving a system of equations in the correlation-adjusted expenditures shares, π^*_{od}, given expenditure share data, π_{od}.

For each destination d, we can write the definition of the partial derivative of the correlation function in its o'th argument evaluated at the competitiveness of all origins in destination d as

$$G_{od} = G^d_o(P^{-\theta}_{1d}, \ldots, P^{-\theta}_{Nd}) \quad \text{for} \quad o = 1, \ldots, N.$$

This is a system of N equations, and each equation in this system is the definition of the correlation correction G_{od}. From (14) in Proposition 2, $P^{-\theta}_{od} \propto \pi^*_{od}$ across o for each d. Since $G^d_o(x_1, \ldots, x_N)$ is homogenous of degree zero, we can use this proportionality to re-write the system in terms of the correlation-adjusted expenditure shares as

$$\pi_{od} = \pi^*_{od} G^d_o\left(\pi^*_{1d}, \ldots, \pi^*_{Nd}\right) \quad \text{for} \quad o = 1, \ldots, N. \tag{30}$$

This expression is an identity: When we evaluate the derivative of the correlation function at the correlation-adjusted expenditure shares, we get the correlation correction. Writing the system in this way is useful because, for observed expenditure share data, it gives us a system of N equations in the N unknown correlation-adjusted expenditure shares. Performing the adjustment for correlation amounts to solving this system.[19]

[19] Note that the correlation adjustment is well defined. The mapping from $\mathbb{R}^N_+$ to $\mathbb{R}^N_+$, defined by the right-hand side of the system in (30), satisfies strict gross substitutability and is homogenous of degree one.

Given the correlation function G^d, this result establishes that calculating the gains from trade only requires expenditure share data.[20] We can relax the assumption of independence, depart from CES demand systems, and calculate the gains from trade using the same data as ACR. For the gains from trade relative to autarky, we can directly apply Proposition 4. To calculate the gains from moving from any given equilibrium to any counterfactual equilibrium, we can solve the system in (30) and then apply hat-algebra methods—as explained in Section 4.3.

4.3 Exact Hat-Algebra

We now show how to apply exact hat-algebra methods to solve for a change from the current (observed) equilibrium to any counterfactual equilibrium. First, we describe the model's equilibrium and establish existence and uniqueness. Then, we show that exact hat-algebra methods can be used, as in ACR, to solve for the equilibrium.

Definition 4 (Competitive Equilibrium). *Given endowments $\{L_o\}_{o=1}^N$, productivities $\{A_o\}_{o=1}^N$, trade costs $\{\tau_{od}\}_{o,d=1}^N$, trade imbalances $\{TB_d\}_{d=1}^N$, and correlation functions $\{G^d\}_{d=1}^N$, a competitive equilibrium consists of wages $\{W_o\}_{o=1}^N$, and expenditure shares $\{\pi_{od}\}_{o,d=1}^N$ such that*

1. *Expenditure shares satisfy*

$$\pi_{od} = \frac{P_{od}^{-\theta} G_o^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})}{G^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})} \quad \textit{for} \quad P_{od} \equiv \frac{\tau_{od} W_o}{A_o};$$

As a result, it is injective (see, for instance, Berry et al., 2013) and there is a unique solution for $\{\pi_{od}^*\}_{o=1}^N$, given $\{\pi_{od}\}_{o=1}^N$.

[20] For an example that delivers closed-form solutions for correlation-adjusted trade shares, consider an iso-elastic import demand system. Following Hanoch (1975) and Sato (1977), let $G^d(x_1, \ldots, x_N)$ be specified as an implicit function satisfying

$$1 = \sum_{o=1}^N \left(\frac{x_o}{G^d(x_1, \ldots, x_N)} \right)^{\frac{1}{1-\rho_o}}.$$

The gains from trade, after some algebra, are given by

$$GT_d = \left(\frac{1-\rho_d}{1-\bar{\rho}_d} \pi_{dd} \right)^{-\frac{1-\rho_d}{\theta}},$$

where $\bar{\rho}_d \equiv \sum_{n=1}^N \rho_n \pi_{nd}$ is d's expenditure-weighted exposure to the correlation of its trading partners. This expression for the gains from trade incorporates an elasticity adjustment and a level adjustment relative to ACR. Together, they imply that more correlation in technology reduces the gains from trade—formalizing, again, Ricardo's second insight. The gains from trade, relative to autarky, can now be different for countries with the same self-trade share.

2. *The labor market in o clears*

$$W_o L_o = \sum_{d=1}^{N} \pi_{od} X_d; \text{ and}$$

3. *The resource constraint in each destination d holds*

$$W_d L_d \equiv Y_d = X_d + TB_d.$$

GEV import demand systems satisfy strict gross substitutability. As a result, the existence and uniqueness of equilibrium follows from standard results in general equilibrium theory.

Proposition 5 (Existence and Uniqueness). *Assume that expenditure in each country is always strictly positive, productivity has a global innovation representation, and markets are perfectly competitive. Then, there exists an competitive equilibrium. The equilibrium is unique up to a normalization—i.e., the choice of numeraire—and can be found using a tâtonnement process.*

Proof. The proof follows from establishing that the implied excess demand system also satisfies strict gross substitutability. Since it is homogenous of degree one and satisfies Walras' law, we can apply Proposition 17.F.3 of Mas-Collell et al. (1995) to establish existence and uniqueness. See Appendix G. □

Next, we solve for the equilibrium using exact hat-algebra methods (see Costinot and Rodrìguez-Clare, 2014). First, we use the results of the previous section to solve for correlation-adjusted trade shares, given the structure of G^d and data on bilateral expenditure,

$$\pi_{od} = \pi^*_{od} G^d_o(\pi^*_{1d}, \ldots, \pi^*_{Nd}).$$

Then, for a given counterfactual shock—e.g., $\{\widehat{\tau}_{od}\}_{o,d=1}^{N}$—we solve for $\{\widehat{W}_o\}_{o=1}^{N}$ from

$$\widehat{W}_o Y_o = \sum_{d=1}^{N} \widehat{\pi}_{od} \pi_{od} (\widehat{W}_d Y_d - TB_d) \quad \text{for each} \quad o = 1, \ldots, N,$$

where $\widehat{L}_o = 1$, $\widehat{TB}_d = 1$,

$$\widehat{\pi}_{od} \pi_{od} = \frac{\widehat{P}_{od}^{-\theta} \pi^*_{od} G^d_o(\widehat{P}_{1d}^{-\theta} \pi^*_{1d}, \ldots, \widehat{P}_{Nd}^{-\theta} \pi^*_{Nd})}{G^d(\widehat{P}_{1d}^{-\theta} \pi^*_{1d}, \ldots, \widehat{P}_{Nd}^{-\theta} \pi^*_{Nd})},$$

and $\widehat{P}_{od} \equiv \widehat{\tau}_{od} \widehat{W}_o / \widehat{A}_o$. Proposition 5 ensures that the equilibrium is unique up to a choice of numeraire, and we can use a *tâtonnement* process to solve for the equilibrium, following Alvarez and Lucas (2007).

After solving for the equilibrium change in wages, we can directly compute the equilibrium change in the price level as

$$\hat{P}_d = \frac{\gamma G^d(\widehat{P}_{1d}^{-\theta} P_{1d}^{-\theta}, \ldots, \widehat{P}_{Nd}^{-\theta} P_{Nd}^{-\theta})^{-\frac{1}{\theta}}}{P_d} = G^d(\widehat{P}_{1d}^{-\theta} \pi_{1d}^*, \ldots, \widehat{P}_{Nd}^{-\theta} \pi_{Nd}^*)^{-\frac{1}{\theta}}$$

since $\pi_{od}^* = (\gamma P_{od}/P_d)^{-\theta}$.

In our empirical application in Section 6, we examine counterfactuals for unilateral and bilateral trade liberalizations. To do so, we first solve for counterfactual wages as above. Then, we do welfare analysis by computing the change in the price level and the implied change in real wages.

5 Aggregation

This section provides a set of aggregation results that let us link the macro trade model we have studied so far to underlying micro trade models. Our aggregation results follow directly from Definition 1: max-linear combinations of multivariate Fréchet random variables give multivariate Fréchet random variables.

This property implies that we can aggregate micro models built on optimizing behavior and multivariate Fréchet productivity to get equivalent macro models where productivity is also multivariate Fréchet. Put differently, the macro-level scale parameters and macro-level correlation functions characterizing the multivariate Fréchet are a result of the underlying micro structure. Our example in Section 2 in which we used a correlation function with a cross-nested CES structure illustrated this aggregation result. As also mentioned in Section 3, many existing trade models have this cross-nested CES structure in which the latent factors take the form of sectors, regions, or the home country of firm.

We formalize our aggregation result in the following proposition.

Proposition 6 (Aggregation of Productivity Process). *Consider a model with M_o micro factors within each origin $o = 1, \ldots, N$. Let micro productivity, $\{\{A_{mod}(v)\}_{m=1}^{M_o}\}_{o=1}^{N}$, have a global innovation representation with $\theta > \sigma$ and underlying applicability $\{\{A_{imod}(v)\}_{m=1}^{M_o}\}_{o=1}^{N}$. Denote the associated micro scale parameters by $\{\{T_{mod}\}_{m=1}^{M_o}\}_{o=1}^{N}$ and micro correlation function by $F^d : \mathbb{R}_+^{M_1} \times \cdots \times \mathbb{R}_+^{M_N} \to \mathbb{R}_+$.*

Then the implied macro productivity defined by $A_{od}(v) = \max_{m=1,\dots,M_o} A_{mod}(v)$ *for each* $o = 1, \dots, N$ *also has a global innovation representation. It's underlying applicability,* $\{A_{iod}(v)\}_{o=1}^{N}$*, satisfies* $A_{iod}(v) = \max_{m=1,\dots,M_o} A_{imod}(v)$*. The macro scale parameters are, for each* $o = 1, \dots, N$*,*

$$T_{od} = F^d\left(\mathbf{0}_1, \dots, \mathbf{0}_{o-1}, \mathbf{T}_{od}, \mathbf{0}_{o+1}, \dots, \mathbf{0}_N\right) \equiv F^{od}(\mathbf{T}_{od})$$

where $\mathbf{0}_o$ *is the zero vector of length* M_o *and* $\mathbf{T}_{od} \equiv (T_{1od}, \dots, T_{M_o od})$*. The function* F^{od} *is o's within-country micro correlation function. The macro correlation function is*

$$G^d(x_1, \dots, x_N) = F^d(\mathbf{\Omega}_{1d}x_1, \dots, \mathbf{\Omega}_{Nd}x_N)$$

where, for each $o = 1, \dots, N$*,* $\mathbf{\Omega}_{od} \equiv (\omega_{1od}, \dots, \omega_{M_o od})$ *is an aggregation weight vector with elements* $\omega_{mod} \equiv T_{mod}/T_{od}$ *for each* $m = 1, \dots, M_o$*.*

Proof. This aggregation result is a direct implication of the global innovation representation in Theorem 1 and the max-stability property. See Appendix H. □

Proposition 6 states that we can relate a given macro model with a global innovation representation to an underlying micro model in which productivity also has a global innovation representation. The link between the micro and macro levels comes from maximization of productivity across within-country micro factors.

Two important consequences of the aggregation result are that it gives a model-consistent way to incorporating micro level trade cost data into macro models, and that it produces aggregate expenditure shares belonging to the GEV class, as we show next.

Corollary 2 (Aggregation of Productivity and Trade Cost Indices). *Suppose the hypotheses and notation of Proposition 6 hold. For each* $o = 1, \dots, N$ *and* $m = 1, \dots, M_o$*, we can calculate the scale parameter* T_{mod} *from observed absolute advantage* T_{moo} *and trade costs* τ_{mod}*,*

$$T_{mod} = T_{moo}\tau_{mod}^{\theta}.$$

We can then aggregate the scale parameters using the within-country micro correlation functions

$$T_{od} = F^{od}\left(\mathbf{T}_{od}\right) \quad \textit{with} \quad \mathbf{T}_{od} \equiv (T_{1od}, \dots, T_{M_o od}).$$

Finally, the implied macro productivity and macro trade cost indices are, respectively,

$$A_o = T_{oo}^{1/\theta} \quad \text{and} \quad \tau_{od} \equiv \left(\frac{T_{oo}}{T_{od}}\right)^{1/\theta}.$$

A similar aggregation result can be obtained for micro-level expenditure shares.

Corollary 3 (Aggregation of Expenditure Shares). *Suppose that the hypotheses and notation of Proposition 6 hold. Suppose that markets are perfectly competitive. Micro expenditure shares are*

$$\pi_{mod} \equiv \frac{T_{mod} W_o^{-\theta} F_{mo}^d \left(\mathbf{T}_{1d} W_1^{-\theta}, \ldots, \mathbf{T}_{Nd} W_N^{-\theta}\right)}{F^d \left(\mathbf{T}_{1d} W_1^{-\theta}, \ldots, \mathbf{T}_{Nd} W_N^{-\theta}\right)} \quad \text{for} \quad m = 1, \ldots, M_o \quad \text{and} \quad o = 1, \ldots, N$$

where $\mathbf{T}_{od} \equiv (T_{1od}, \ldots, T_{M_o od})$ *and* $F_{mo}^d(\mathbf{X}_1, \ldots, \mathbf{X}_N) \equiv \frac{\partial}{\partial x_{mo}} F^d(\mathbf{X}_1, \ldots, \mathbf{X}_N)$.

For macro productivity A_o *and trade cost indices* τ_{od} *from Corollary 2, define an import price index* $P_{od} \equiv \tau_{od} W_o / A_o$. *Then, macro expenditure shares are*

$$\pi_{od} \equiv \sum_{m=1}^{M_o} \pi_{mod} = \frac{P_{od}^{-\theta} G_o^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})}{G^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})} \quad \text{for} \quad o = 1, \ldots, N.$$

This result states that the implications of an aggregated micro model are identical to the implications of a macro model. That is, we can pass seamlessly between the micro and macro levels. As a result, we can use micro level estimation to measure the micro correlation function, micro productivity, and micro trade cost indices. Then, we can use these aggregation results to derive the macro correlation function, macro productivity, and macro trade cost indices. From there, we can perform macro counterfactual analysis using the results in Section 4.1. These results also enable us to consider which Ricardian micro-foundations might generate macro substitution patterns associated with particular GEV import demand systems—as in Section 3.

In practice, we can use micro-level expenditure shares to compute aggregation weights. The key is that gravity holds after correlation correction at the micro-level:

$$\pi_{mod}^* = \omega_{mod} \left(\gamma \frac{P_{od}}{P_d}\right)^{-\theta}$$

for $\pi_{mod}^* \equiv \pi_{mod} / F_{mod}$ and $F_{mod} \equiv F_{mo}^d \left(\mathbf{\Omega}_{1d} P_{1d}^{-\theta}, \ldots, \mathbf{\Omega}_{Nd} P_{Nd}^{-\theta}\right)$.

To recover aggregation weights from observed micro-level expenditure, first solve for correlation-adjusted expenditure shares:

$$\pi_{mod} = \pi^*_{mod} F^d_{mo} \left(\mathbf{\Pi}^*_{1d}, \ldots, \mathbf{\Pi}^*_{Nd}\right)$$

for $\mathbf{\Pi}^*_{od} \equiv (\pi^*_{1od}, \ldots, \pi^*_{M_o od})$. Then, by homogeneity of F^{od} and the definition of T_{od}, we have $F^{od}(\mathbf{\Omega}_{od}) = 1$. As a result, we can recover aggregation weights as

$$\omega_{mod} = \frac{\pi^*_{mod}}{F^{od}(\mathbf{\Pi}^*_{od})}.$$

Correlation-adjusted expenditure shares are sufficient statistics for aggregation weights. We apply this result in the next section for the case of a cross-nested CES correlation function.

Appendix I presents a series of extensions that incorporate micro-foundations underlying widely-used trade models. In particular, we consider demand side factors as in Armington (Anderson, 1979); global value chains (Antràs and de Gortari, 2017); and heterogenous firms (Melitz, 2003; Chaney, 2008). These extensions help to clarify what restrictions are necessary to have micro and macro import demand systems in the GEV class.

6 Application: Multi-Sector Model of Trade

We now present—and estimate—an application of our framework with a flexible correlation structure. We work with a cross-nested CES correlation structure in which the micro factors are given by sectors. This application is appealing for several reasons: First, it represents a common extension of the EK model of trade; second, it clearly illustrates how a micro structure aggregates into the Ricardian macro structure of Section 2; and finally, since gravity holds at the sector level, we can estimate the macro correlation function—that rationalizes the aggregated sector-level data—using gravity equations.

Assume that the correlation function takes a a cross-nested CES form,

$$G^d(x_1, \cdots, x_N) = \sum_{s=1}^{S} \left(\sum_{o=1}^{N} (\omega_{sod} x_o)^{1/(1-\rho_s)} \right)^{1-\rho_s}.$$

The outer sum captures the factors—i.e., sectors in this case—that induce correlation across origins. The parameter ρ_s measures the degree of correlation across origin countries in each sector s, while ω_{sod} measures the extent to which sector s matters for trade flows from o to d—i.e., it reflects sectoral trade costs and comparative advantage. To satisfy the normalization property for correlation functions in Definition 2, we impose that $\sum_{s=1}^{N} \omega_{sod} = 1$.

6.1 Sectoral Gravity

We implement a two-step procedure to estimate the correlation function and the parameter θ. This amounts to first estimating a within-sector demand system, and then a between-sector demand system.

Using the results in Section 3, define $\sigma_s \equiv \theta/(1-\rho_s)$, $P_{sod} = \omega_{sod}^{-1/\theta} P_{od}$ and $P_{sd} = (\sum_{o=1}^{N} P_{sod}^{-\sigma_s})^{-\frac{1}{\sigma_s}}$. Expenditure shares at the sector level are

$$\pi_{sod} = \left(\frac{P_{sod}}{P_{sd}}\right)^{-\sigma_s} \left(\gamma \frac{P_{sd}}{P_d}\right)^{-\theta}, \tag{31}$$

with P_d the aggregate price index in country d. Summing over origins, we get that

$$\sum_o \pi_{sod} = \left(\frac{P_{sd}}{P_d}\right)^{-\theta}, \tag{32}$$

so that

$$x_{sod} \equiv \frac{\pi_{sod}}{\sum_o \pi_{sod}} = \left(\frac{P_{sod}}{P_{sd}}\right)^{-\sigma_s}. \tag{33}$$

Taking logs in (33), and using that $P_{sod} = \tau_{sod} W_o / A_{so}$ for $A_{so} \equiv T_{soo}^{1/\theta}$ and $\tau_{sod} \equiv (T_{soo}/T_{sod})^{1/\theta}$, we get a sectoral gravity equation,

$$\ln x_{sod} = -\sigma_s \ln \frac{W_o}{A_{so}} - \sigma_s \ln P_{sd} - \sigma_s \ln \tau_{sod}. \tag{34}$$

We assume that trade costs take the following log-linear form,

$$\ln \tau_{sod} = \ln \bar{\tau}_d + \ln(1 + t_{sod}) + \delta_s' Geo_{od} + u_{sod}^1 + u_{sd}^2, \tag{35}$$

where u_{sod}^1 is orthogonal to u_{sd}^2—i.e. mean zero across origins, o, conditional on a sector destination pair, sd. t_{sod} are bilateral sectoral measures of trade costs, such as freight costs and tariffs, and Geo_{od} are other (bilateral) geography covariates.

In a first step, substituting (35) into (34), we use variation in observed trade costs over origins to estimate the sectoral elasticity of substitution, σ_s, from

$$\ln x_{sod} = \alpha_{so} + \beta_{sd} - \sigma_s \ln(1 + t_{sod}) + \epsilon^1_{sod}, \tag{36}$$

where $\alpha_{so} \equiv -\sigma_s \ln W_o/A_{so}$, $\beta_{sd} \equiv -\sigma_s(\ln P_{sd} - \bar{\tau}_d - u^2_{sd})$ are, respectively, a sector-origin and a sector-destination fixed effect, and $\epsilon^1_{sod} \equiv -\sigma_s u^1_{sod}$. The identification assumption in this regression is that $\ln(1+t_{sod})$ is orthogonal to the residual u^1_{sod} in the origin dimension, conditional on sector-origin effects, sector-destination effects, and geography.

In a second step, we estimate θ using (32), our assumption on trade costs, and our first-step estimates. First, using (33), we can write relative prices as

$$\ln \frac{P_{sd}}{P_d} = \ln \frac{P_{sd}}{P_{sod}} - \ln \frac{P_{sod}}{P_d} = \frac{1}{\sigma_s} \ln x_{sod} - \ln \frac{W_o}{A_{so}} - \ln \tau_{sod} + \ln P_d. \tag{37}$$

Replacing (35) in (37) and further substituting in (32), we can identify θ from

$$\ln y_{sd} = -\frac{\theta}{\sigma_s} \ln \hat{x}_{sod} + a_{so} + b_d + \theta \delta_s' Geo_{od} + \theta \ln(1 + t_{sod}) + \epsilon^2_{sod}, \tag{38}$$

where $y_{sd} \equiv \sum_o \pi_{sod}$, $\ln \hat{x}_{sod} = \ln x_{sod} - \hat{\sigma}_s \hat{u}^1_{sod}$ (i.e. the predicted value coming from the first stage gravity regression), $a_{so} \equiv \theta \ln W_o/A_{so}$, $b_d \equiv \theta(\ln \bar{\tau}_d - \ln P_d)$, and $\epsilon^2_{sod} \equiv \theta u^2_{sd}$. This specification is a gravity regression with sector-origin and destination effects and the additional first-stage control $\ln \hat{x}_{sod}$. The estimate of θ is consistent as long as this control is included; the identification assumption is that $\ln(1 + t_{sod})$ is orthogonal to u^2_{sd}, conditional on sector-origin effects, destination-specific effects, geography, and the predicted first-stage left-hand side variable. Notice that the control $\hat{x}_{sod}$ depends on the structural residual u^2_{sd} so that its coefficient would give a biased estimate of θ. That is why we use variation in observed trade costs across sectors to estimate the trade elasticity.

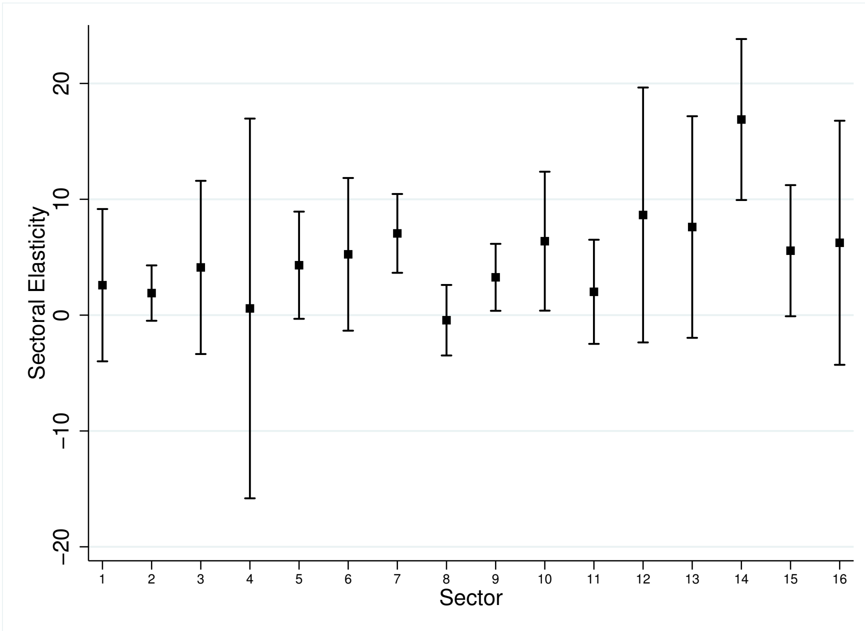
We use data on trade flows and freight costs from Adao et al. (2017). The data contain 36 countries plus the rest of the world, and 16 sectors. Data for freight costs are available for two importers only, Australia and the United States. We restrict the period to 1995-2006. Gravity covariates are from CEPII.

Pooling the data over years, we first estimate, by OLS, the following sectoral gravity equation

$$\Delta_d \ln x_{sodt} = \Delta_d \beta_{sdt} - \sigma_s \Delta_d \ln(1 + t_{sodt}) - \sigma_s \delta_{st}' \Delta_d Geo_{od} + \Delta_d \epsilon^1_{sodt}, \tag{39}$$

where Δ_d denotes the difference between $d = AUS$ and $d = USA$. Figure 1 summarizes

Figure 1: Estimates of Sectoral Elasticities, σ_s, OLS.



Notes: Results from estimating (39) by OLS. Point estimates and 95% confidence intervals shown for standard errors clustered by sector.

the estimates for the sectoral elasticity of substitution, σ_s.

We can also use (39) to estimate the CES model. Under the assumption of zero correlation ($G_{dd} = 1$ and $\rho_s = 0$)—i.e. the ACR case—we have $\sigma_s = \theta$. Pooling the data over sectors, we estimate θ^{ces}—the CES trade elasticity. Columns 1-3 in Table 1 present these estimates. In the following section, we use the estimate in column 3 of $\theta^{ces} = 2.328$ for our counterfactual exercises using the CES model.

To get an estimate of θ for the cross-nested model with heterogenous correlation across sectors, we estimate through OLS the following gravity equation, pooling again the data over years,

$$\Delta_d \ln y_{sdt} = -\frac{\theta}{\sigma_s}\Delta_d \ln \hat{x}_{sodt} + \theta \Delta_d \ln(1 + t_{sodt}) + \Delta_d \beta_{dt} + \delta'_{st}\Delta_d Geo_{od} + \Delta_d \epsilon^2_{sodt}, \tag{40}$$

for differences between $d = AUS$ and $d = USA$. The final three columns of Table 1 show the results. Our preferred estimate is in column 6, $\theta = 0.313$. Sectoral correlations are simply calculated using the definition of σ_s, and the restriction that $\rho_s \geq 0$, $\rho_s = \max(0, 1 - \frac{\theta}{\sigma_s})$, for all s.[21]

[21] The sectoral models in Arkolakis et al. (2012), Costinot and Rodrìguez-Clare (2014), and Caliendo and Parro (2015) have Cobb-Douglas expenditure shares across sectors within each destination country. As a

Table 1: Estimates of the Trade Elasticity θ, OLS.

Dep variable	No correlation ($G_{dd} = 1$) $\Delta \ln x_{sodt}$			Correlation ($G_{dd} \neq 1$) $\Delta \ln y_{sdt}$		
	(1)	(2)	(3)	(4)	(5)	(6)
θ	2.194	2.101	2.328	.055	.150	.313
	(.872)**	(.929)**	(1.063)**	(.030)*	(.046)***	(.040)***
$\Delta \ln \widehat{x}_{sodt}/\sigma_s$				.005	.104	.272
				(.002)**	(.031)***	(.029)***
Share border	2.209	3.311	3.027	-.002	-.123	-.365
	(.599)***	(.640)***	(.976)***	(.002)	(.051)**	(.042)***
Log Distance	-.355	.524	.505	-.002	-.011	-.117
	(.190)*	(.343)	(.422)	(.001)*	(.010)	(.022)***
Year & Sector Effects	✓	✓		✓	✓	✓
Year-Sector Effects			✓			
Year & Sector-Covariate Interactions		✓			✓	
Year-Sector-Covariate Interactions			✓			✓
Obs.	5,880	5,880	5,880	5,920	5,920	5,920
R^2	.39	.44	.46	.93	.93	.97

Notes: Results from estimating (40) by OLS. Standard errors in parenthesis, clustered by exporter country in (1)-(3) and by sector in (4)-(6), with levels of significance denoted by *** p < 0.01, and ** p < 0.05 and * p<0.1.

6.2 Gains from Trade and the Role of Correlation

Armed with estimates for θ and ρ_s, we use the bilateral sectoral data for the 36 countries in the sample, for each year, and construct estimates of the gains from trade,

$$GT_d = \left(\frac{\sum_s \pi_{sdd}}{G_{dd}}\right)^{-\frac{1}{\theta}}, \tag{41}$$

and the correlation-correction term,

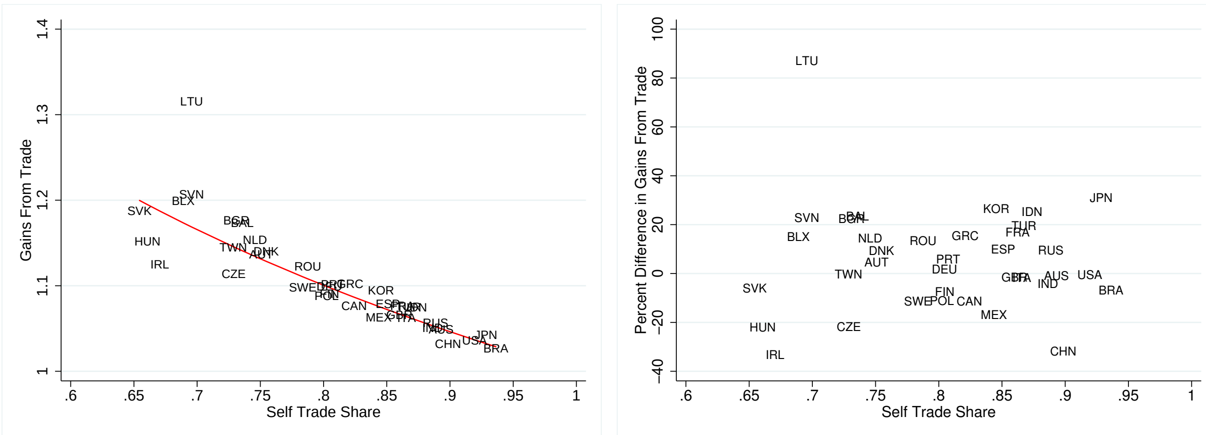
$$G_{dd} = \frac{\sum_s \pi_{sdd}}{\sum_s \pi_{sdd}^{1-\rho_s} \left(\sum_o \pi_{sod}\right)^{\rho_s}}. \tag{42}$$

Figure 2 shows the results, by country, for 2006. Panel (2a) compares the gains from trade calculated under our cross-nested CES specification against the gains from trade calculated using CES. That is, we compare $GT_d^{CN} = (\pi_{dd}/G_{dd})^{-1/\theta}$ and $GT_d^{CES} = \pi_{dd}^{-1/\theta^{ces}}$, with $\theta = 0.313$ and $\theta^{ces} = 2.328$.

result, they are limiting cases of the cross-nested CES model as $\theta \to 0$.

Figure 2: Gains from Trade, 2006.

(a) level (b) percent difference



Notes: (2a) Black data: $GT_d^{CN} = (\sum_s \pi_{sdd}/G_{dd})^{-1/\theta}$. Red line: $GT_d^{CES} = (\sum_s \pi_{sdd})^{-1/\theta^{ces}}$. (2b): Percent difference calculated as $100 \times \frac{GT_d^{CN} - GT_d^{CES}}{GT_d^{CES} - 1}$.

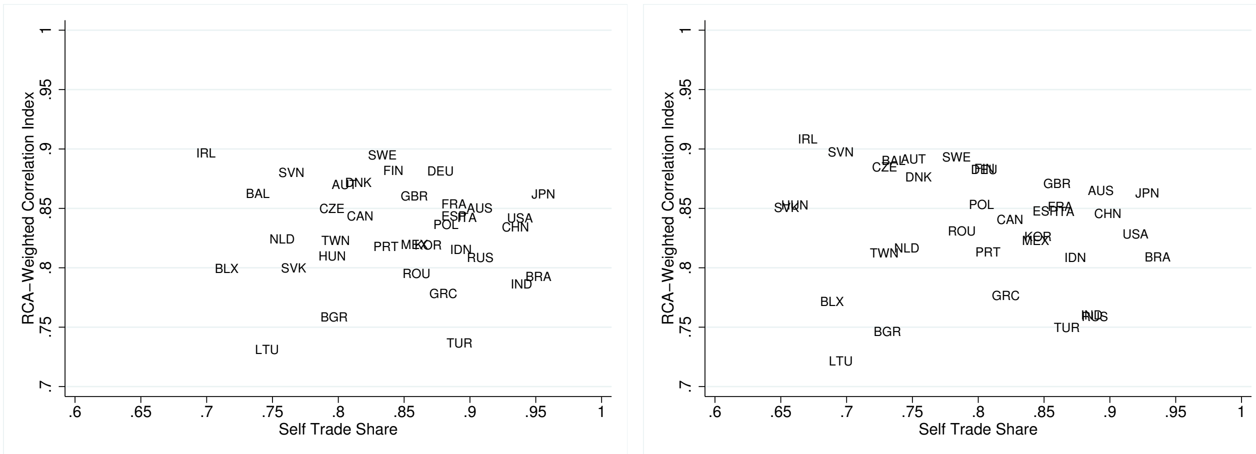
Correcting the gains from trade for correlation implies gains that do not have a one-to-one mapping to self-trade shares; two countries with the same level of openness can now have very different gains from trade. Panel (2b) makes this point even clearer: For instance, while under CES, Korea (KOR) and Mexico (MEX) would have the same gains from trade, correcting those gains for correlation entails gains that are 40 percent higher for Korea. Additionally our calculations suggest that, for individual countries, the corrected gains from trade can be substantially different from the ones without the correction. In the case of Lithuania (LTU), gains from trade under CES are less than 20 percent, but they are above 30 percent when the correlation correction is considered. By contrast, Ireland (IRL) would have gains of around 12 percent with the correlation correction, around eight-percentage points lower than the gains dictated by its self-trade share. Appendix Figures J.1 and J.2 show the results by year, for all years.

The heterogeneity observed in the gains from trade, and that is not captured by self trade, relates precisely to Ricardo's insight: Countries with higher gains should be countries that are specialized in sectors that present low correlation with—i.e., more dissimilar to—the ones of their trading partners. We explore the source of heterogeneity in the gains from trade by linking it the to specialization patterns of countries. To such end, we use Balassa revealed-comparative-advantage (RCA) indices, constructed at the sector-country-year level,

$$RCA_{sot} = \frac{\sum_d X_{sodt} / \sum_{s'dt} X_{s'odt}}{\sum_{o'd} X_{so'dt} / \sum_{s'o'd} X_{s'o'dt}}, \tag{43}$$

Figure 3: RCA-Weighted Correlation Index, selected years.

(a) 1996 (b) 2006



Notes: Revealed-Comparative-Advantage (RCA)-weighted correlation index: $\rho_{dt}^{RCA} = \sum_s \rho_s \frac{RCA_{sdt}}{\sum_{s'} RCA_{s'dt}}$.

and use them as weights to aggregate, by country and year, our sectoral estimates of the correlation parameters, ρ_s,

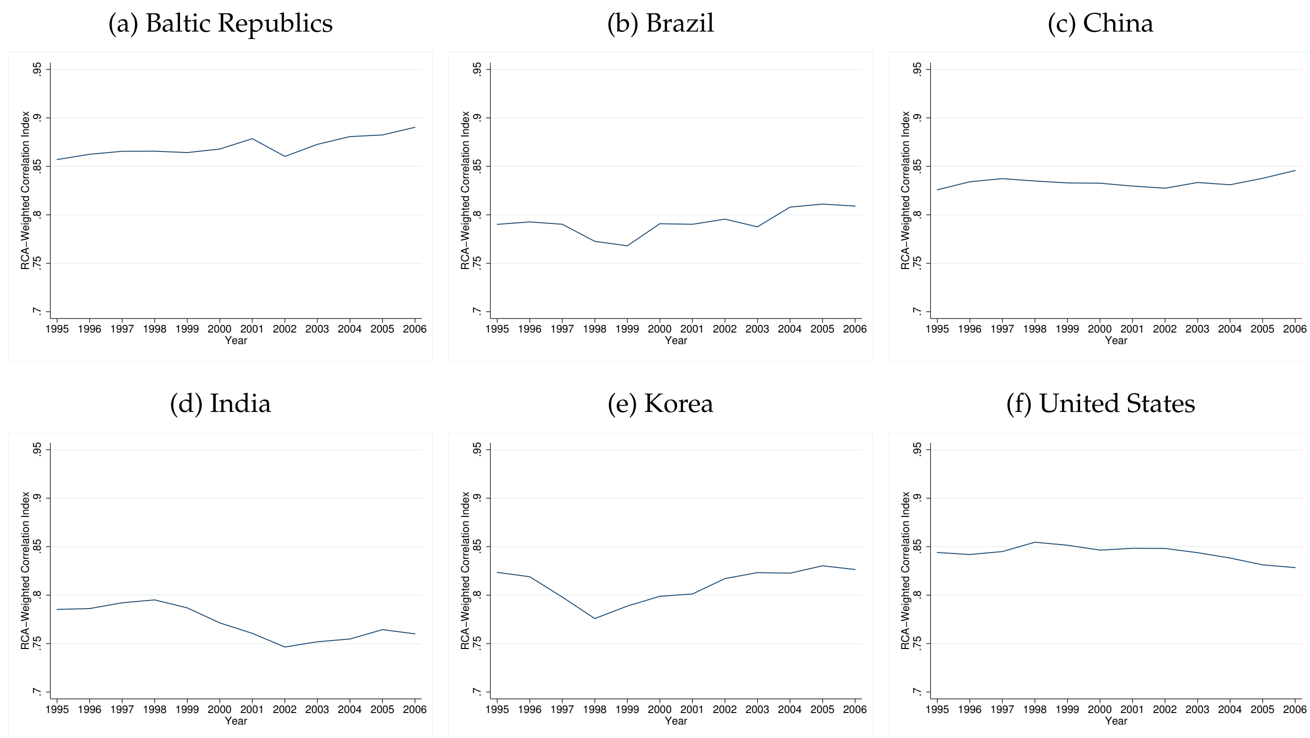
$$\rho_{dt}^{RCA} = \sum_s \rho_s \frac{RCA_{sdt}}{\sum_{s'} RCA_{s'dt}}. \tag{44}$$

We call the expression in (44) the RCA-weighted correlation index. The interpretation is very intuitive: High RCA_{sot} indices in high correlation sectors (i.e., high ρ_s) imply a high ρ_{dt}^{RCA}.

Figure 3 shows our RCA-weighted correlation index against self trade, by country, for 1996 and 2006. First, consider again the case of Ireland (IRL) and Lithuania (LTU), for 2006. These two countries have very similar self-trade shares, which under a CES demand system would imply very similar gains. However, their sectoral trade patterns are very different: While Ireland specializes in sectors which are correlated with those of its trading partners ($\rho_{IRL,2006}^{RCA} = 0.91$), Lithuania specializes in sectors with a lower correlation ($\rho_{LTU,2006}^{RCA} = 0.73$). The lower correlation index for Lithuania materializes in gains from trade that are 120 percent higher than the gains observed for Ireland. Appendix Figure J.3 shows the results by year, for all years.

Differences in specialization patterns imply not only differences in the gains from trade—that go beyond differences captured by self trade—in the cross section of countries, but also differences in the gains from trade across time, for each country. The presence of correlation represents a key channel through which the gains from trade change with a country's comparative advantage patterns. As documented by Hanson et al. (2015),

Figure 4: Evolution of RCA-Weighted Correlation Index, selected countries.

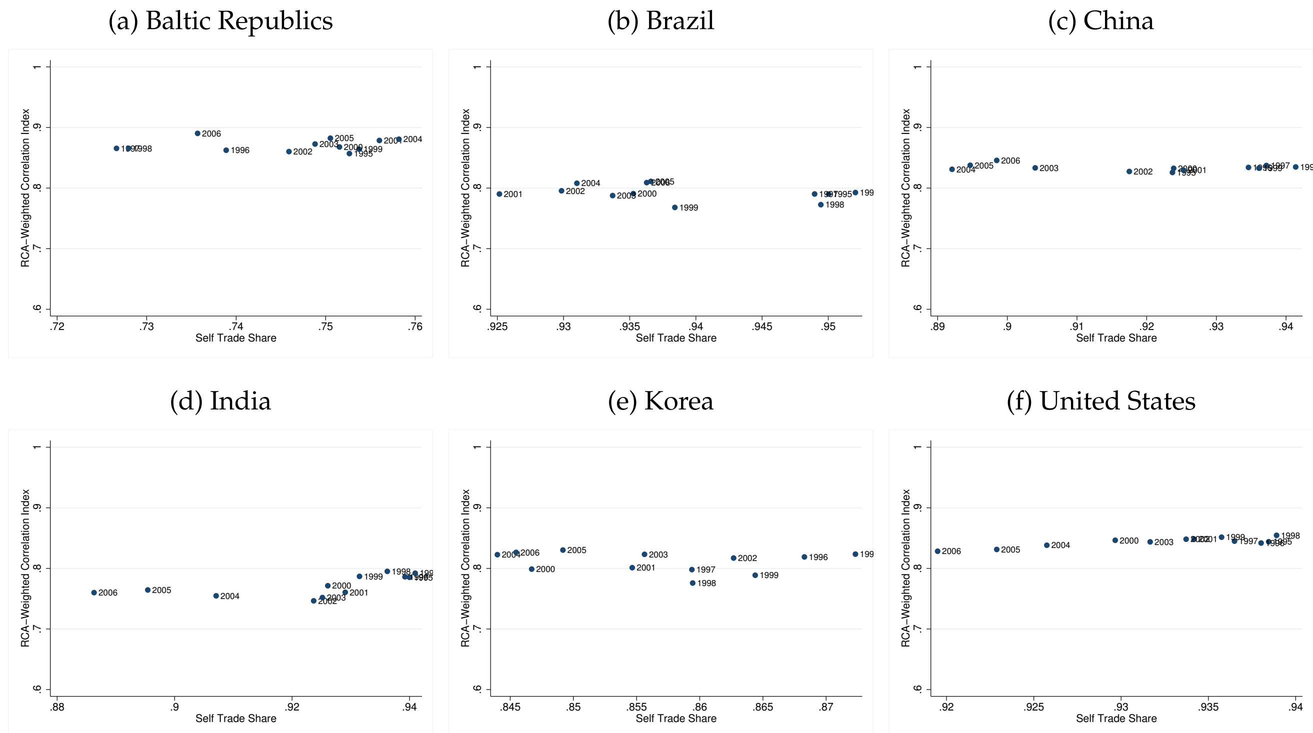


Notes: Revealed-Comparative-Advantage (RCA)-weighted correlation index: $\rho_{dt}^{RCA} = \sum_s \rho_s \frac{RCA_{sdt}}{\sum_{s'} RCA_{s'dt}}$. Baltic Republics are Estonia, Latvia, and Lithuania.

the churning in comparative advantage for individual countries across time is extremely dynamic. In the same vein, Figure 4 links the evolution of specialization patterns, for selected countries, to our RCA-weighted correlation index. Reflecting the ever-evolving pattern of comparative advantage, a country's correlation structure with trading partners can drastically change year by year, and with it, the gains from trade. Appendix Figure J.4 shows the results by country, for all countries.

While Figure 3 shows the link between the RCA-weighted correlation index and self trade for the cross of countries in our sample, Figure 5 shows that relation across time, for selected countries. While for a country like the United States, the changes in the gains from trade are captured fairly well by the changes in self-trade shares, for the Baltic Republics, for instance, changes in their sectoral specialization patterns, which is reflected in changes in the correlation index, can have large impacts in the calculations of their gains from trade in the time series. Heterogeneity in the gains from trade across time, coming from the correlation channel, seems to be important. Appendix Figure J.5 presents the results for each country in our sample.

Figure 5: RCA-Weighted Correlation Index and self-trade shares, selected countries.



Notes: Revealed-Comparative-Advantage (RCA)-weighted correlation index: $\rho_{dt}^{RCA} = \sum_s \rho_s \frac{RCA_{sdt}}{\sum_{s'} RCA_{s'dt}}$. Baltic Republics are Estonia, Latvia, and Lithuania.

The negative correlation between the RCA-weighted correlation index and the gains from trade is confirmed by running an OLS regression of the gains from trade, GT_{dt} and the correlation correlation, G_{ddt}, alternately, on the RCA-weighted correlation index, ρ_{dt}^{RCA} and self-trade shares, π_{ddt}. Results in Table 2 confirm that the correlation index is significantly and negatively related to the correlation-correction term and to the gains from trade, respectively, across countries within a year (columns 2 and 6, respectively). However, the relation is not significant after including country fixed effects, indicating that, once we control by self trade, the RCA-weighted correlation index does not explain variation over time within a given country in either the correlation correction or the gains from trade.

6.3 Implications for Trade Liberalization

Next, we consider the effect of unilateral and bilateral trade liberalization. We show that correlation in technology changes the identity of the trading partner with the most impact on real wages in the United States.

Table 2: Comparative Advantage and the Gains from Trade, OLS.

Dep variable	$\ln\ G_{ddt}$				$\ln\ GT_{dt}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ρ_{dt}^{RCA}	-.094	-.095	.025	.022	-.302	-.303	.081	.070
	(.008)***	(.008)***	(.018)	(.018)	(.026)***	(.026)***	(.056)	(.057)
$\ln \pi_{ddt}$	.857	.858	.862	.869	-.455	-.454	-.441	-.417
	(.004)***	(.004)***	(.007)***	(.010)***	(.013)***	(.013)***	(.024)***	(.032)***
Year Effects		✓		✓		✓		✓
Country Effects			✓	✓			✓	✓
Obs.	444	444	444	444	444	444	444	444
R^2	.991	.991	.998	.998	.749	.752	.949	.952

Notes: Revealed-Comparative-Advantage (RCA)-weighted correlation index: $\rho_{dt}^{RCA} = \sum_s \rho_s \frac{RCA_{sdt}}{\sum_{s'} RCA_{s'dt}}$. Standard errors in parenthesis with levels of significance denoted by ** p < 0.01, and * p < 0.05.

We first consider a series of trade liberalization counterfactuals where the United States unilaterally reduces trade costs. In each counterfactual, the cost of importing into the United States from a single origin country falls by ten percent. Using the procedure outlined in Section 4, we compute the counterfactual change in real wages for both the case of the cross-nested CES model where we account for correlation in technology and the case of the CES model without correlation.

The top panel of Table 3 presents the results. The presence of correlation in technology changes country rankings, which we can see by comparing the first and second columns of Table 3. For example, while the CES model implies that real wages in the United States are most sensitive to changes in Canadian trade costs, the cross-nested CES model predicts that China has the largest impact. A notable example of the importance of correlation is Mexico—who moves from being the third-best partner from the perspective of the CES model to being the fifth-best partner after accounting for correlation, jumping below Japan (third best) and Germany (fourth best). Similarly, India moves down the ranking from eight to ten, passing over Korea (eight) and Italy (nine).

Why does correlation matter for these rankings? The gains from trade come from two potentially offsetting effects—a price effect and a wage effect. The price effect is direct. Cutting trade costs on Chinese goods reduces prices for U.S. consumers. The size of this effect is just the elasticity of the price in the United States to the price of imports from

China, which equals the expenditure share,

$$\frac{\partial \ln P_{USA}}{\partial \ln \tau_{CHN,USA}} = \frac{\partial \ln G^{USA}(P_{1,USA}^{-\theta}, \ldots, P_{N,USA}^{-\theta})^{-\frac{1}{\theta}}}{\partial \ln P_{CHN,USA}} = \frac{P_{CHN,USA}^{-\theta} G_{CHN,USA}}{P_{USA}^{-\theta}} = \pi_{CHN,USA}.$$

In contrast, the wage effect is indirect and operates through general equilibrium effects across labor markets. The market clearing condition for the United States is

$$W_{USA} L_{USA} = \sum_{d=1}^{N} \frac{P_{USA,d}^{-\theta} G_{USA}^{d}(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})}{G^{d}(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})} X_d.$$

If the change in trade costs with China induces U.S. consumers to substitute expenditure towards Chinese goods and away from U.S. goods, the right-hand side of this condition—capturing labor demand—will fall and U.S. wages will fall. How rapidly expenditure shifts away from U.S. goods and towards Chinese goods depends on correlation in technology. If the United States and China were specialized in similar sectors, and those sectors had a high degree of correlation, then U.S. and Chinese goods would be substitutable and labor demand in the United States would fall when trade costs for imports from China fall. In this case, the wage effect would be large and would offset the gains from liberalization through falling prices in the United States. In contrast, if China and the United States were specialized in distinct sectors, or in sectors with a low degree of correlation, then the fall in labor demand would be small, leading to a small wage effect and large gains from trade.

As we see, the later is the case for China and the former is the case for Canada. Accounting for correlation in technology leads to an increase in the gains from liberalizing with China (from 0.059 to 0.065 percent) while it leads to a decrease from liberalizing with Canada (from 0.070 to 0.063 percent). For small changes in trade costs, this difference between countries comes from the wage effect because the direct price effect is measured by the expenditure share which doesn't change between the CES and cross-nested CES models. Incorporating correlation introduces additional heterogeneity in the gains from trade, and changes the ranking of trading partners for the United States.

These unilateral counterfactuals provide intuition for why correlation matters for the gains from trade liberalization. However, do these ranking reversals remain in the more realistic scenario of bilateral trade agreements?

To answer this question, we consider how real wages in the United States would be af-

Table 3: Gains from Trade Liberalization for the United States.

	Ranking		Gains (in %)	
	correlation	no correlation	correlation	no correlation
		10% unilateral liberalization		
China	1	2	0.065	0.059
Canada	2	1	0.063	0.070
Japan	3	4	0.036	0.030
Germany	4	5	0.034	0.029
Mexico	5	3	0.033	0.040
Great Britain	6	6	0.022	0.020
France	7	7	0.015	0.013
Korea	8	9	0.014	0.011
Italy	9	10	0.014	0.011
India	10	8	0.013	0.012
		10% bilateral liberalization		
China	1	2	0.127	0.077
Canada	2	1	0.123	0.119
Japan	3	4	0.076	0.047
Mexico	4	3	0.075	0.069
Germany	5	5	0.068	0.044
Great Britain	6	6	0.052	0.035
Korea	7	8	0.038	0.021
France	8	7	0.033	0.022
India	9	9	0.025	0.017
Taiwan	10	10	0.024	0.016

Notes: The upper panel shows results from lowering, by ten percent, the trade cost for imports into the United States from a trading-partner country, one partner at a time. The lower panel shows results from lowering, simultaneously and by ten percent, the trade cost for exports from the United States into a trading-partner country and imports to the United States from the same trading-partner country, one partner at a time. Ranks refer to welfare gains for the United States for each trade-liberalization exercise.

fected from entering trade agreements with individual countries. That is, we consider bilateral trade liberalizations where both the United States and a trading partner agree to reduce trade costs by ten percent. As before, we compute the gains from trade using the procedure in Section 4.

The bottom panel of Table 3 presents the results. Correlation continues to generate ranking reversals. According to the cross-nested CES model where we account for within-sector correlation in technology, China continues to have the most impact on real wages in the United States, and, if correlation were neglected, we would instead conclude that Canada has the most influence. Accounting for sectoral correlation leads to additional changes in the rankings of countries relative to the case of unilateral liberalization. For example, Taiwan is now among the ten countries with the most impact on U.S. real wages, while, for unilateral liberalization, it was not. Here, ranking reversals reflect the same mechanism as for unilateral liberalizations, except that now they involve the combined effect of trade cost reductions by both countries.

These counterfactual exercises show that the effects of trade liberalization can change substantially once we account for correlation in technology. Incorporating correlation does not only simply increase gains on average, but also introduces heterogeneity in gains that leads to changes in the ranking of trading partners. Counterfactual analysis based on quantitative trade models can change significantly once we incorporate Ricardo's second insight that differences in technological similarity matter for the gains from trade.

7 Conclusions

This paper is motivated by the old Ricardian idea that a country gains the most from trading with those countries that are technologically less similar. We argue that Ricardo's insight is absent from the canonical Ricardian model of trade and develop a theory of trade that allows for arbitrary patterns of correlation in technology between countries. We start from technology primitives that generate a multivariate Fréchet Ricardian model with a general correlation structure and yet preserve all the tractability of EK-type tools. Importantly, our representation of productivity is equivalent to the entire class of GEV import demand systems and, as such, approximates *any* Ricardian model—not only the ones with Fréchet-distributed productivity.

We show that the gains from trade coming from a GEV import demand system can be written as a simple correction to self-trade shares—i.e. the CES case. Moreover, the theory, by relating macro substitutability patterns to underlying micro structures, provides guidance on incorporating standard micro estimates into macro counterfactual exercises.

Our empirical application to a multi-sector model of trade reveals that the adjustment implied by our correlation structure matters: Gains are much more heterogeneous across countries, and within countries across time, than otherwise. Moreover, the correlation adjustment has implications for the gains from unilateral and bilateral trade liberalizations. These results suggest that our framework—which fully captures Ricardo's insights—has the potential to change quantitative conclusions in any literature applying Fréchet tools.

References

Adao, R., A. Costinot, and D. Donaldson (2017). Nonparametric counterfactual predictions in neoclassical models of international trade. *The American Economic Review 107*(3), 633–689.

Ahlfeldt, G., S. Redding, D. Sturm, and N. Wolf (2015). The economics of density: Evidence from the berlin wall. *Econometrica 83*(6), 2127–2189.

Alvarez, F. and R. E. Lucas (2007). General equilibrium analysis of the eaton-kortum model of international trade. *Journal of Monetary Economics 54*(6), 1726–1768.

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *The American Economic Review 69*(1), 106–116.

Antràs, P. and A. de Gortari (2017). On the geography of global value chains. *Mimeo, Harvard University*.

Arkolakis, C., A. Costinot, D. Donaldson, and A. Rodriguez-Clare (2017). The elusive pro-competitive effects of trade. *Review of Economic Studies, forthcoming*.

Arkolakis, C., A. Costinot, and A. Rodríguez-clare (2012). New trade models, same old gains? *The American Economic Review*, 94–130.

Bas, M., T. Mayer, and M. Thoenig (2017). From micro to macro: Demand, supply, and heterogeneity in the trade elasticity. *Journal of International Economics 108*, 1–21.

Berry, S., A. Gandhi, and P. Haile (2013). Connected substitutes and invertibility of demand. *Econometrica 81*(5), 2087–2111.

Bertoletti, P., F. Etro, and I. Simonovska (2017). International trade with indirect additivity. *American Economic Journal: Microeconomics, forthcoming*.

Brooks, W. and P. Pujolas (2017). Non linear gravity.

Buera, F. J. and E. Oberfield (2016). The global diffusion of ideas. Technical report, National Bureau of Economic Research.

Caliendo, L., F. P. annd Esteban Rossi-Hansberg, and P.-D. Sarte (2017). The impact of regional and sectoral productivity changes on the u.s. economy. *Review of Economic Studies, forthcoming*.

Caliendo, L. and F. Parro (2015). Estimates of the trade and welfare effects of nafta. *The Review of Economic Studies 82*(1), 1–44.

Caron, J., T. Fally, and J. Markusen (2014). International trade puzzles: a solution linking production factors and demand. *The Quarterly Journal of Economics 129*(3), 1501–1552.

Chaney, T. (2008). Distorted gravity: the intensive and extensive margins of international trade. *The American Economic Review 98*(4), 1707–1721.

Costinot, A., D. Donaldson, and I. Komunjer (2012). What goods do countries trade? a quantitative exploration of ricardo's ideas. *Review of Economic Studies 79*, 581–608.

Costinot, A., D. Donaldson, J. Vogel, and I. Werning (2015). Comparative advantage and optimal trade policy. *The Quarterly Journal of Economics 130*(2), 659–702.

Costinot, A. and A. Rodrìguez-Clare (2014). Trade theory with numbers: Quantifying the consequences of globalization. Technical Report 4.

Dagsvik, J. K. (1994). Discrete and continuous choice, max-stable processes, and independence from irrelevant attributes. *Econometrica: Journal of the Econometric Society*, 1179–1205.

Dagsvik, J. K. (1995). How large is the class of generalized extreme value random utility models? *Journal of Mathematical Psychology 39*(1), 90–98.

De Haan, L. (1984). A spectral representation for max-stable processes. *The Annals of Probability 12*(4), 1194–1204.

DeLoecker, J., P. Goldberg, A. Khandelwal, and N. Pavcnik (2016). Prices, markups and trade reform. *Econometrica 84*(2), 445Ð510.

Eaton, J. and S. Kortum (1999). International technology diffusion: Theory and measurement. *International Economic Review 40*(3), 537–570.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica*, 1741–1779.

Eaton, J. and S. Kortum (2010). Technology in the global economy: A framework for quantitative analysis. *Unpublished manuscript*.

Eaton, J. and S. Kortum (2012). Putting ricardo to work. *Journal of Economic Perspectives 26*(2), 65–90.

Fajgelbaum, P. and S. Redding (2014). External integration, structural transformation and economic development: Evidence from argentina 1870-1914. *NBER Working Paper 20217*.

Feenstra, R. C., P. Luck, M. Obstfeld, and K. N. Russ (2017). In search of the armington elasticity. *The Review of Economics and Statistics 0*(0), null.

Feenstra, R. C. and D. E. Weinstein (2017). Globalization, markups, and the U.S. price level. *Journal of Political Economy 125*(4), 1040–1074.

Fosgerau, M., D. McFadden, and M. Bierlaire (2013). Choice probability generating functions. *Journal of Choice Modelling 8*, 1–18.

Hanoch, G. (1975). Production and demand models with direct or indirect implicit additivity. *Econometrica: Journal of the Econometric Society*, 395–419.

Hanson, G. H., N. Lind, and M.-A. Muendler (2015). The dynamics of comparative advantage. Technical report, National bureau of economic research.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2013). The allocation of talent and u.s. economic growth. Technical report, National Bureau of Economic Research.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *The American Economic Review 70*(5), 950–959.

Lagakos, D. and M. E. Waugh (2013). Selection, agriculture, and cross-country productivity differences. *The American Economic Review 103*(2), 948–980.

Lashkari, D. and M. Mestieri (2016). Gains from trade with heterogeneous income and price elasticities. *Mimeo, Harvard University*.

Lashkaripour, A. and V. Lugovskyy (2017). Industry-level scale economies: from micro-estimation to macro-implications. *Mimeo, Indiana University*.

Levchenko, A. and J. Zhang (2016). The evolution of comparative advantage: Measurement and welfare implications. *Journal of Monetary Economics 78*, 96–111.

Mas-Collell, A., M. Whinston, and J. R. Green (1995). *Microeconomic Theory*. Oxford University Press.

McFadden, D. (1978). Modeling the choice of residential location. *Transportation Research Record* (673).

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71*(6), 1695–1725.

Monte, F., S. Redding, and E. Rossi-Hansberg (2015). Commuting, migration and local employment elasticities. *NBER Working Paper 21706*.

Ossa, R. (2015). Why trade matters after all. *Journal of International Economics 97*(2), 266–277.

Penrose, M. D. (1992). Semi-min-stable processes. *The Annals of Probability 20*(3), 1450–1463.

Ramondo, N. and A. Rodríguez-Clare (2013). Trade, multinational production, and the gains from openness. *Journal of Political Economy 121*(2), 273–322.

Ramondo, N., A. Rodríguez-Clare, and M. Saborío-Rodríguez (2016). Trade, domestic frictions, and scale effects. *American Economic Review 106*(10), 3159–84.

Redding, S. (2016). Goods trade, factor mobility and welfare. *Journal of International Economics 101*, 148–167.

Redding, S. and E. Rossi-Hansberg (2017). Quantitative spatial economics. *Annual Review of Economics 9*, 21–58.

Redding, S. J. and D. E. Weinstein (2017). Aggregating from micro to macro patterns of trade. *Mimeo, Princeton University*.

Rudin, W. et al. (1964). *Principles of mathematical analysis*, Volume 3. McGraw-hill New York.

Sato, R. (1977). Homothetic and non-homothetic CES production functions. *The American Economic Review 67*(4), 559–569.

Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes 5*(1), 33–44.

Smith, T. E. (1984). A choice probability characterization of generalized extreme value models. *Applied Mathematics and Computation 14*(1), 35–62.

Tintelnot, F. (2017). Global production with export platforms. *The Quarterly Journal of Economics 132*(1), 157–209.

A Properties of Fréchet Random Variables

Lemma A.1. *Let X be distributed as an Fréchet random variable with location $T > 0$ and shape $\alpha > 0$ so that it has cumulative distribution function of $\mathbb{P}[X \leq x] = e^{-Tx^{-\alpha}}$. Then if $\alpha > 1$, it has mean $\mathbb{E}[X] = \Gamma(1 - 1/\alpha)T^{1/\alpha}$. Also, for any $S > 0$ and $\beta > 0$, $(S^{1/\alpha}X)^{\beta}$ is Fréchet with location ST and shape α/β.*

Proof.

$$\begin{aligned}\mathbb{E}[X] &= \int_0^\infty z \frac{\partial}{\partial z}\mathbb{P}\left[X \leq z\right] \mathrm{d}z = \int_0^\infty z\frac{\partial}{\partial z}e^{-Tz^{-\alpha}}\mathrm{d}z \\ &= \int_0^\infty ze^{-Tz^{-\alpha}}\alpha T z^{-\alpha-1}\mathrm{d}z = \int_0^\infty t^{-1/\alpha}e^{-t}\mathrm{d}tT^{1/\alpha} = \Gamma(1-1/\alpha)T^{1/\alpha}.\end{aligned}$$

and

$$\mathbb{P}[(S^{1/\alpha}X)^{\beta} \leq z] = \mathbb{P}[X \leq S^{-1/\alpha}z^{1/\beta}] = e^{-T(S^{-1/\alpha}z^{1/\beta})^{-\alpha}} = e^{-STz^{-\alpha/\beta}}.$$

□

Lemma A.2. *Let $\{X_i\}_{i=1,\dots,N}$ be α-Fréchet with scale parameters $\{T_i\}_{i=1}^N$ and correlation function $G : \mathbb{R}_+^N \to \mathbb{R}_+$. Then, for any $S_i \geq 0$ $i = 1, \dots, N$ and $\beta > 0$, the random variable $\{(S_i^{1/\alpha}X_i)^{\beta}\}_{i=1}^N$ is α/β-Fréchet with location parameters of $\{S_iT_i\}_{i=1}^N$ and correlation function G.*

Proof.

$$\begin{aligned}\mathbb{P}[(S_i^{1/\alpha}X_i)^{\beta} \leq y_i, i = 1, \dots, N] &= \mathbb{P}[X_i \leq S_i^{\alpha}y_i^{1/\beta}, i = 1, \dots, N] \\ &= \exp\left[-G(T_1S_1y_1^{-\alpha/\beta}, \dots, T_NS_Nx_N^{-\alpha/\beta})\right].\end{aligned}$$

□

Lemma A.3. *Let $\{X_i\}_{i=1,\dots,N}$ be θ-Fréchet with scale parameters $\{T_i\}_{i=1}^N$ and correlation function $G : \mathbb{R}_+^N \to \mathbb{R}_+$. Then, the random variable $\max_{i=1,\dots,N} X_i$ is θ-Fréchet with location $G(T_1, \dots, T_N)$. Moreover, let $\{\mathcal{I}_j\}_{j=1}^M$ be any partition of $\{1, \dots, N\}$ and define the random variable $\{Y_1, \dots, Y_M\}$ as*

$$Y_j = \max_{i \in \mathcal{I}_j} X_i.$$

Let $j : \{1, \dots, N\} \to \{1, \dots, M\}$ be the unique mapping such that $j = j(i)$ if and only if $i \in \mathcal{I}_j$. Define $\tilde{T}_j = G(T_1\mathbf{1}\{1 \in \mathcal{I}_j\}, \dots, T_N\mathbf{1}\{N \in \mathcal{I}_j\})$ and $\omega_i = \frac{T_i}{\tilde{T}_j}\mathbf{1}\{i \in \mathcal{I}_j\}$. Then:

1. *$\{Y_1, \ldots, Y_M\}$ is θ-Fréchet with correlation function $H : \mathbb{R}_+^M \to \mathbb{R}_+$ satisfying*

$$H(z_1, \ldots, z_M) = G(\omega_1 z_{j(1)}, \ldots, \omega_N z_{j(N)}).$$

2. $$\mathbb{P}\left[Y_j = \max_i X_i\right] = \frac{\sum_{i \in \mathcal{I}_j} T_i G_i(T_1, \ldots, T_N)}{G(T_1, \ldots, T_N)}.$$

   where $G_i(x_1, \ldots, x_N) \equiv \partial G(x_1, \ldots, x_N)/\partial x_i$.

3. *For any $j = 1, \ldots, M$, the distribution of Y_j conditional on the event $Y_j = \max_{i=1,\ldots,N} X_i$ is identical to the distribution of $\max_{i=1,\ldots,N} X_i$:*

$$\mathbb{P}\left[Y_j \leq y \mid Y_j = \max_i X_i\right] = e^{-G(T_1,\ldots,T_N)y^{-\theta}} = \mathbb{P}\left[\max_{i=1,\ldots,N} X_i \leq y\right].$$

Proof. We first prove part (1). Let $\{\mathcal{I}_j\}_{j=1}^M$ be a partition of $\{1, \ldots, N\}$ and define $Y_j = \max_{i \in \mathcal{I}_j} X_i$. Let the function $j : \{1, \ldots, N\} \to \{1, \ldots, M\}$ satisfy $i \in \mathcal{I}_{j(i)}$ for all $i = 1, \ldots, N$. Note that there is a unique function satisfying this condition since $\{\mathcal{I}_j\}_{j=1}^M$ is a partition of $\{1, \ldots, N\}$. Then,

$$\begin{aligned}\mathbb{P}\left[Y_j \leq y_j, \forall j = 1, \ldots, M\right] &= \mathbb{P}\left[X_i \leq y_j, \forall i \in \mathcal{I}_j, \forall j = 1, \ldots, M\right] \\ &= e^{-G(T_1 y_{j(1)}^{-\theta}, \ldots, T_N y_{j(N)}^{-\theta})}.\end{aligned}$$

Therefore $\{Y_1, \ldots, Y_M\}$ is θ-Fréchet. It's scale parameters are

$$\lim_{y_k \to \infty, k \neq j} G(T_1 y_{j(1)}^{-\theta}, \ldots, T_N y_{j(N)}^{-\theta}) = G(T_1 \mathbf{1}\{1 \in \mathcal{I}_j\}, \ldots, T_N \mathbf{1}\{N \in \mathcal{I}_j\}) = \tilde{T}_j$$

and its correlation function must then be

$$G(T_1/\tilde{T}_{j(1)} z_{j(1)}, \ldots, T_N/\tilde{T}_{j(N)} z_{j(N)}) = G(\omega_1 z_{j(1)}, \ldots, \omega_N z_{j(N)}) = H(z_1, \ldots, z_M).$$

Note that if we take $M = 1$ so that $\mathcal{I}_1 = \{1, \ldots, N\}$ we get

$$\begin{aligned}\mathbb{P}\left[\max_{i=1,\ldots,N} X_i \leq y\right] &= \mathbb{P}\left[Y_1 \leq y\right] = \mathbb{P}\left[Y_j \leq y, \forall j = 1, \ldots, M\right] \\ &= e^{-G(T_1 y^{-\theta}, \ldots, T_N y^{-\theta})} = e^{-G(T_1, \ldots, T_N) y^{-\theta}}.\end{aligned}$$

That is, $\max_{i=1,\ldots,N} X_i$ is a Fréchet random variable with location $G(T_1, \ldots, T_N)$ and shape θ.

Next we prove part (2). We have

$$
\begin{aligned}
&\mathbb{P}\left[\max_i X_i \leq y \text{ and } Y_j = \max_i X_i\right] = \mathbb{P}\left[Y_j \leq y \text{ and } Y_j = \max_i X_i\right] \\
&= \mathbb{P}\left[Y_j \leq y \text{ and } X_i \leq Y_j, \forall i = 1, \dots, N\right] = \mathbb{P}\left[Y_j \leq y \text{ and } X_i \leq Y_j, \forall i \notin \mathcal{I}_j\right] \\
&= \int_0^y \mathbb{P}\left[X_i \leq t, \forall i \notin \mathcal{I}_j \mid Y_j = t\right] \frac{\partial}{\partial t}\mathbb{P}[Y_j \leq t]\mathrm{d}t \\
&= \int_0^y \frac{\partial}{\partial t}\, \mathbb{P}\left[X_i \leq z, \forall i \notin \mathcal{I}_j, \text{ and } X_i \leq t, \forall i \in \mathcal{I}_j\right]\Big|_{z=t} \mathrm{d}t \\
&= \int_0^y \sum_{i \in \mathcal{I}_j} \frac{\partial}{\partial y_i}\, e^{-G(T_1 y_1^{-\theta}, \dots, T_N y_N^{-\theta})}\Big|_{y_i = t, \forall i = 1, \dots, N} \mathrm{d}t \\
&= \int_0^y \sum_{i \in \mathcal{I}_j} e^{-G(T_1 y_1^{-\theta}, \dots, T_N y_N^{-\theta})} G_i(T_1 y_1^{-\theta}, \dots, T_N y_N^{-\theta}) T_i \theta y_i^{-\theta-1}\Big|_{y_i = t, \forall i = 1, \dots, N} \mathrm{d}t \\
&= \int_0^y e^{-G(T_1, \dots, T_N) t^{-\theta}} \sum_{i \in \mathcal{I}_j} T_i G_i(T_1, \dots, T_N) \theta t^{-\theta-1} \mathrm{d}t \\
&= \frac{\sum_{i \in \mathcal{I}_j} T_i G_i(T_1, \dots, T_N)}{G(T_1, \dots, T_N)} \int_0^y e^{-G(T_1, \dots, T_N) t^{-\theta}} G(T_1, \dots, T_N) \theta t^{-\theta-1} \mathrm{d}t \\
&= \frac{\sum_{i \in \mathcal{I}_j} T_i G_i(T_1, \dots, T_N)}{G(T_1, \dots, T_N)} e^{-G(T_1, \dots, T_N) y^{-\theta}},
\end{aligned}
$$

where $G_i(x_1, \dots, x_N) = \partial G(x_1, \dots, x_N)/\partial x_i$. Let $y \to \infty$ to get

$$
\mathbb{P}\left[Y_j = \max_i X_i\right] = \frac{\sum_{i \in \mathcal{I}_j} T_i G_i(T_1, \dots, T_N)}{G(T_1, \dots, T_N)}.
$$

Finally, we can prove part (3) using the previous results:

$$
\begin{aligned}
\mathbb{P}\left[\max_i X_i \leq y \mid Y_j = \max_i X_i\right] &= \frac{\mathbb{P}\left[\max_i X_i \leq y \text{ and } Y_j = \max_i X_i\right]}{\mathbb{P}\left[Y_j = \max_i X_i\right]} \\
&= \frac{\frac{\sum_{i \in \mathcal{I}_j} T_i G_i(T_1, \dots, T_N)}{G(T_1, \dots, T_N)} e^{-G(T_1, \dots, T_N) z^{-\theta}}}{\frac{\sum_{i \in \mathcal{I}_j} T_i G_i(T_1, \dots, T_N)}{G(T_1, \dots, T_N)}} \\
&= e^{-G(T_1, \dots, T_N) z^{-\theta}} \\
&= \mathbb{P}\left[\max_i X_i \leq y\right].
\end{aligned}
$$

$\square$

Corollary A.1 (Multivariate Max-Stability). *Let the random vector $(A_1, \dots, A_K)$ be multivariate θ-Fréchet with scale parameters T_k for $k = 1, \dots, K$ and correlation function F. Define an-*

other random vector $(B_1, \ldots, B_J)$ formed by max-linear combinations: $B_j \equiv \max_{k=1,\ldots,K} \alpha_{jk} A_k$ for for weights $\alpha_{jk} \geq 0$ for each $j = 1, \ldots, J$ and $k = 1, \ldots, K$. Then $(B_1, \ldots, B_J)$ is multivariate θ-Fréchet with scale parameters

$$S_j \equiv F(\alpha_{j1}^{1/\theta} T_1, \ldots, \alpha_{jK}^{1/\theta} T_K) \quad \text{, for} \quad j = 1, \ldots, J,$$

and correlation function

$$G(x_1, \ldots, x_J) \equiv F(\mathbf{\Omega}_1 \cdot \mathbf{X}, \ldots, \mathbf{\Omega}_K \cdot \mathbf{X}) \quad \text{with} \quad \mathbf{X} \equiv (x_1, \ldots, x_J),$$

for weight vectors $\mathbf{\Omega}_k = (\alpha_{1k}^{1/\theta} T_k / S_1, \ldots, \alpha_{Jk}^{1/\theta} T_k / S_J)$ for $k = 1, \ldots, K$ and $\mathbf{\Omega} \cdot \mathbf{X}$ denoting the inner product of vectors.

B Proof of Lemma 1

Proof. First, we show that if productivity is θ-Fréchet, then there must exist a correlation function $G^d : \mathbb{R}_+^N \to \mathbb{R}_+$ such that (3) is the joint distribution of productivity across origins. Consider any $(x_1, \ldots, x_N) \in \mathbb{R}_+^N$. Then $x_o^{1/\theta} \geq 0$ for each o. From the definition of a multivariate θ-Fréchet random variable, the random variable $\max_{o=1,\ldots,N} x_o^{1/\theta} A_{od}(v)$ must be distributed as a θ-Fréchet random variable. That is, there exists some $T > 0$ such that

$$\mathbb{P}\left[\max_{o=1,\ldots,N} x_o^{1/\theta} A_{od}(v) \leq a\right] = e^{-Ta^{-\theta}}.$$

Let $T^d : \mathbb{R}_+^N \to \mathbb{R}_+$ be the map $(x_1, \ldots, x_N) \mapsto T$. We then have that for any $(x_1, \ldots, x_N) \in \mathbb{R}_+^N$

$$\mathbb{P}\left[\max_{o=1,\ldots,N} x_o^{1/\theta} A_{od}(v) \leq a\right] = \exp\left[-T^d(x_1, \ldots, x_N) a^{-\theta}\right].$$

Note that the joint distribution of productivity can be written as

$$\begin{aligned}\mathbb{P}[A_{1d}(v) \leq a_1, \ldots, A_{Nd}(v) \leq a_N] &= \mathbb{P}[A_{1d}(v)/a_1 \leq 1, \ldots, A_{Nd}(v)/a_N \leq 1] \\ &= \mathbb{P}\left[\max_{o=1,\ldots,N} A_{od}(v)/a_o \leq 1\right].\end{aligned}$$

Choosing $x_o = a_o^{-\theta}$ and $a = 1$ we can use the properties of our function T^d and get

$$\mathbb{P}\left[\max_{o=1,\ldots,N} A_{od}(v)/a_o \leq 1\right] = \exp\left[-T^d(a_1^{-\theta}, \ldots, a_N^{-\theta})\right].$$

therefore the joint distribution of productivity satisfies

$$\mathbb{P}[A_{1d}(v) \leq a_1, \ldots, A_{Nd}(v) \leq a_N] = e^{-G^d(T_{1d}a_1^{-\theta}, \ldots, T_{Nd}a_N^{-\theta})},$$

for the function $G^d : \mathbb{R}_+^N \to \mathbb{R}_+$ defined by $(x_1, \ldots, x_N) \mapsto T^d(x_1/T_{1d}, \ldots, x_N/T_{Nd})$.

We now show that this G^d is a correlation function. First we show that is must be homogenous. Fix $(x_1, \ldots, x_N) \in \mathbb{R}_+^N$ and let $\lambda > 0$. We have

$$\begin{aligned}
\exp\left[-G^d(\lambda x_1, \ldots, \lambda x_N)\right] &= \mathbb{P}[T_{1d}A_{1d}(v)^{-\theta} \geq \lambda x_1, \ldots, T_{Nd}A_{Nd}(v)^{-\theta} \geq \lambda x_N] \\
&= \mathbb{P}[(x_1/T_{1d})^{1/\theta}A_{1d}(v) \leq \lambda^{-1/\theta}, \ldots, (x_N/T_{Nd})^{-1/\theta}A_{Nd}(v) \leq \lambda^{-1/\theta}] \\
&= \mathbb{P}[\max_{o=1,\ldots,N}(x_o/T_{od})^{-1/\theta}A_{od}(v) \leq \lambda^{-1/\theta}] \\
&= \exp\left[-T^d(x_1/T_{1d}, \ldots, x_N/T_{Nd})\lambda\right] \\
&= \exp\left[-\lambda G^d(x_1, \ldots, x_N)\right],
\end{aligned}$$

so that $G^d(\lambda x_1, \ldots, \lambda x_N) = \lambda G^d(x_1, \ldots, x_N)$ as desired.

Now consider the normalization restriction. Fix o. The distribution of $A_{od}(v)$ is

$$\exp\left(-T_{od}a^{-\theta}\right) = \mathbb{P}[A_{od}(v) \leq a] = \mathbb{P}\left[\max_{n=1,\ldots,N} x_n^{1/\theta}A_{nd}(v) \leq a\right],$$

for the choice of $x_n = 0$ for $n \neq o$ and $x_o = 1$. But then,

$$\begin{aligned}
\exp\left(-T_{od}a^{-\theta}\right) &= \exp\left[-T^d(x_1, \ldots, x_N)a^{-\theta}\right] \\
&= \exp\left[-T^d(0, \ldots, 0, 1, 0, \ldots, 0)a^{-\theta}\right] \\
&= \exp\left[-G^d(0, \ldots, 0, T_{od}, 0, \ldots, 0)a^{-\theta}\right] \\
&= \exp\left[-G^d(0, \ldots, 0, 1, 0, \ldots, 0)T_{od}a^{-\theta}\right],
\end{aligned}$$

where the last line comes from homogeneity. We therefore must have $G^d(0, \ldots, 0, 1, 0, \ldots, 0) = 1$ as desired.

The unboundedness restriction follows from the limiting properties of a joint distribution. Fix o.

$$\begin{aligned}
\lim_{x_o \to \infty} e^{-G^d(x_1, \ldots, x_N)} &= \lim_{x_o \to \infty} \mathbb{P}[T_{1d}A_{1d}(v)^{-\theta} \geq x_1, \ldots, A_{Nd}(v) \geq x_N] \\
&= \lim_{x_o \to \infty} \mathbb{P}[T_{1d}^{-1/\theta}A_{1d}(v) \leq x_1, \ldots, T_{Nd}^{-1/\theta}A_{Nd}(v) \leq x_N] = 0.
\end{aligned}$$

Therefore $\lim_{x_o \to \infty} G^d(x_1, \ldots, x_N) = \infty$ as desired.

Finally, the differentiability restrictions are necessary because the productivity distribution is continuous and therefore has a joint density function. Smith (1984) shows that the differentiability condition is necessary for this joint density to exist.

Therefore, the function G^d must be a correlation function, and we have proven that if productivity is θ-Fréchet then there exists of a correlation function $G^d : \mathbb{R}^N_+ \to \mathbb{R}_+$ such that (3) holds.

We now prove the converse. Let $T_{od} > 0$ for each $o = 1, \ldots, N$, and let $G^d : \mathbb{R}^N_+ \to \mathbb{R}_+$ be a correlation function. Suppose that $\{A_{od}(v)\}_{o=1}^N$ satisfies

$$\mathbb{P}[A_{1d}(v) \leq a_1, \ldots, A_{Nd}(v) \leq a_N] = \exp\left[-G^d(T_{od}a_1^{-\theta}, \ldots, T_{Nd}a_N^{-\theta})\right].$$

We want to show that $\{A_{od}(v)\}_{o=1}^N$ is θ-Fréchet. Let $(x_1, \ldots, x_N) \in \mathbb{R}^N_+$ and consider the distribution of $\max_{o=1,\ldots,N} x_o A_{od}(v)$. It is

$$\begin{aligned}
\mathbb{P}[\max_{o=1,\ldots,N} x_o A_{od}(v) \leq a] &= \mathbb{P}[x_1 A_{1d}(v) \leq a, \ldots, x_N A_{Nd}(v) \leq a] \\
&= \mathbb{P}[A_{1d}(v) \leq a/x_1, \ldots, A_{Nd}(v) \leq a/x_N] \\
&= \exp\left[-G^d(T_{od}x_1^{\theta}a^{-\theta}, \ldots, T_{Nd}x_N^{\theta}a^{-\theta})\right] \\
&= \exp\left[-G^d(T_{od}x_1^{\theta}, \ldots, T_{Nd}x_N^{\theta})a^{-\theta}\right],
\end{aligned}$$

where the last line uses homogeneity of G^d. Therefore $\max_{o=1,\ldots,N} x_o A_{od}(v)$ is a θ-Fréchet random variable with location parameter $G^d(T_{od}x_1^{\theta}, \ldots, T_{Nd}x_N^{\theta})$. As a result, we can conclude that $\{A_{od}(v)\}_{o=1}^N$ is θ-Fréchet. □

C Proof of Theorem 1

Proof. Normalize by $\mathbb{E}A_{iod}(v)^{\theta}$ and transform by a power of θ to define

$$\widetilde{A}_{od}(v) \equiv \frac{A_{od}(v)^{\theta}}{\mathbb{E}A_{iod}(v)^{\theta}} = \max_{i=1,2,\ldots} Z_i(v)^{\theta} \frac{A_{iod}(v)^{\theta}}{\mathbb{E}A_{iod}(v)^{\theta}} = \max_{i=1,2,\ldots} Z_i(v)^{\theta} \widetilde{A}_{iod}(v),$$

for $\widetilde{A}_{iod}(v) \equiv \frac{A_{iod}(v)^{\theta}}{\mathbb{E}A_{iod}(v)^{\theta}}$. By theorem 2 of De Haan (1984), theorem 3 of Penrose (1992), or, most directly, theorem 2 of Schlather (2002), we get that $\widetilde{A}_{od}(v)$ is unit-Fréchet with standard unit-Fréchet marginals if and only if we can take $\{Z_i(v)^{\theta}\}_{i=1,2,\ldots}$ as a Poisson point process on $[0, \infty)$ with intensity measure $g^{-2}\mathrm{d}g$ and $\mathbb{E}\widetilde{A}_{iod}(v) = 1$. □

D Proof of Proposition 1

Proof. Perfect competition implies that potential import prices are

$$P_{od}(v) = \frac{W_o}{A_{od}(v)}.$$

Then

$$\begin{aligned}
\mathbb{P}[P_{1d}(v) \geq p_1, \ldots, P_{Nd}(v) \geq p_N] &= \mathbb{P}[P_{1d}(v)/W_1 \geq p_1/W_1, \ldots, P_{Nd}(v)/W_N \geq p_N/W_N] \\
&= \mathbb{P}[1/A_{1d}(v) \geq p_1/W_1, \ldots, 1/A_{Nd}(v) \geq p_N/W_N] \\
&= \mathbb{P}[A_{1d}(v) \leq W_1/p_1, \ldots, A_{Nd}(v) \leq W_N/p_N].
\end{aligned}$$

By Theorem 1,

$$\begin{aligned}
\mathbb{P}[A_{1d}(v) \leq W_1/p_1, \ldots, A_{Nd}(v) \leq W_N/p_N] &= \exp\left[-\mathbb{E} \max_{o=1,\ldots,N} \left(\frac{A_{iod}(v)}{W_o/p_o}\right)^\theta\right] \\
&= \exp\left[-\mathbb{E} \max_{o=1,\ldots,N} \frac{A_{iod}(v)^\theta}{\mathbb{E}[A_{iod}(v)^\theta]} \mathbb{E}[A_{iod}(v)^\theta] \left(\frac{p_o}{W_o}\right)^\theta\right] \\
&= \exp\left[-\mathbb{E} \max_{o=1,\ldots,N} \frac{A_{iod}(v)^\theta}{T_{od}} \left(\frac{p_o}{\tau_{od} W_o/A_o}\right)^\theta\right] \\
&= \exp\left[-\mathbb{E} \max_{o=1,\ldots,N} \frac{A_{iod}(v)^\theta}{T_{od}} P_{od}^{-\theta} p_o^\theta\right].
\end{aligned}$$

Since

$$G^d(x_1, \ldots, X_n) = \mathbb{E} \max_{o=1,\ldots,N} \frac{A_{iod}(v)^\theta}{T_{od}} x_o,$$

we have

$$\begin{aligned}
\mathbb{P}[P_{1d}(v) \geq p_1, \ldots, P_{Nd}(v) \geq p_N] &= \mathbb{P}[A_{1d}(v) \leq W_1/p_1, \ldots, A_{Nd}(v) \leq W_N/p_N] \\
&= \exp\left[-G^d(P_{1d}^{-\theta} p_1^\theta, \ldots, P_{Nd}^{-\theta} p_n^\theta)\right].
\end{aligned}$$

□

E Proof of Proposition 2

Proof. The proof of Proposition 2 follows directly from the properties of Fréchet random variables. The probability that variety v is imported by destination d from origin o is

$$\pi_{od} \equiv \mathbb{P}[P_{od}(v) \geq P_{o'd}(v) \quad \forall o' \neq o] = \frac{P_{od}^{-\theta} G_o^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})}{G^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})},$$

using Proposition 1 and Lemma A.3. The distribution of prices among those goods imported by destination d from country o satisfies

$$\mathbb{P}\left[P_{od}(v) \geq p \mid P_{od}(v) = \min_{o'=1,\ldots,N} P_{o'd}(v)\right] = \mathbb{P}\left[\min_{o'=1,\ldots,N} P_{o'd}(v) \geq p\right] = e^{-G(P_{1d}^{-\theta},\ldots,P_{Nd}^{-\theta})p^{\theta}},$$

by Proposition 1 and Lemma A.3. The price index in destination d is then

$$P_d = \left[\int_0^1 \min_{o=1,\ldots,N} P_{od}(v)^{-\sigma} \mathrm{d}v\right]^{-\frac{1}{\sigma}} = \left[\mathbb{E}\left(\min_{o=1,\ldots,N} P_{od}(v)^{-\sigma}\right)\right]^{-\frac{1}{\sigma}} = \gamma G^d\left(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta}\right)^{-\frac{1}{\theta}},$$

where $\gamma = \Gamma\left(\frac{\theta-\sigma}{\theta}\right)^{-\frac{1}{\sigma}}$, $\Gamma(\cdot)$ is the gamma function, and the last equality follows from the fact that $\min_{o=1,\ldots,N} P_{od}(v)^{-\sigma} = (\max_{o=1,\ldots,N} 1/P_{od}(v))^{\sigma}$ is a Fréchet random variable with location $G^d\left(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta}\right)$ and shape $\theta/\sigma > 1$ due to the assumption that $\theta > \sigma$ and due to Lemma A.1. □

F Proof of Proposition 3

Proof. First, the set of varieties from o imported to d is $\{v \in [0,1] \mid P_{od}/U_{od}(v) = \min_{o'} P_{o'd}/U_{o'd}(v)\}$ and for any variety in this set, expenditure is

$$X_d(v) = \left(\frac{W_o/A_{od}(v)}{P_d}\right)^{-\sigma} X_d = \left(\frac{P_{od}/U_{od}(v)}{P_d}\right)^{-\sigma} X_d.$$

Any v not in this set must get imported from a different origin.

Note that the price index is

$$P_d = \left[\int_0^1 \left(\min_{o'} W_{o'}/A_{o'd}(v)\right)^{-\sigma} \mathrm{d}v\right]^{-\frac{1}{\sigma}} = \left[\int_0^1 \left(\min_{o'} P_{o'd}/U_{o'd}(v)\right)^{-\sigma} \mathrm{d}v\right]^{-\frac{1}{\sigma}}$$

so we can write the expenditure share as

$$
\begin{aligned}
\pi_{od}(\mathbf{P}_d, X_d) &\equiv \int_0^1 \frac{X_d(v)}{X_d} \mathbf{1}\left\{P_{od}/U_{od}(v) = \min_{o'} P_{o'd}/U_{o'd}(v)\right\} \mathrm{d}v \\
&= \frac{\int_0^1 (P_{od}/U_{od}(v))^{-\sigma} \mathbf{1}\left\{P_{od}/U_{od}(v) = \min_{o'} P_{o'd}/U_{o'd}(v)\right\} \mathrm{d}v}{\int_0^1 \left(\min_{o'} P_{o'd}/U_{o'd}(v)\right)^{-\sigma} \mathrm{d}v} \\
&= \frac{\mathbb{E}\left[(P_{od}/U_{od}(v))^{-\sigma} \mathbf{1}\left\{P_{od}/U_{od}(v) = \min_{o'} P_{o'd}/U_{o'd}(v)\right\}\right]}{\mathbb{E}\left[\left(\min_{o'} P_{o'd}/U_{o'd}(v)\right)^{-\sigma}\right]}.
\end{aligned}
$$

It remains to show that there exists a correlation function which approximates this demand system. The proof is similar to the proof of Theorem 1 in Dagsvik (1995), differing in the functional form of the demand system to be approximated. We will construct an approximating GEV demand system using the following correlation function for some $\theta > \sigma$:

$$
G^d(x_1, \ldots, x_N) = \left[\mathbb{E}\left(\sum_o (U_{od}(v)^\theta x_o)\right)^{\frac{\sigma}{\theta}}\right]^{\frac{\theta}{\sigma}}.
$$

This choice will give the result because it implies a price level which approximates the true price level. Recall that the price level associated with a correlation function is $P_d = \Gamma\left(\frac{\theta-\sigma}{\theta}\right) G^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})^{-\frac{1}{\theta}}$. Then

$$
\begin{aligned}
P_d &= \Gamma\left(\frac{\theta-\sigma}{\theta}\right)\left[\mathbb{E}\left(\sum_o (U_{od}(v)/P_{od})^\theta\right)^{\frac{\sigma}{\theta}}\right]^{-\frac{1}{\sigma}} \\
&\overset{\theta\to\infty}{\to} \left[\mathbb{E}\left(\max_o U_{od}(v)/P_{od}\right)^\sigma\right]^{-\frac{1}{\sigma}} = \left[\mathbb{E}\left(\min_o P_{od}/U_{od}(v)\right)^{-\sigma}\right]^{-\frac{1}{\sigma}}.
\end{aligned}
$$

That is, the price level implied by this correlation function approximates the true price level associated with true productivity.

The implied GEV import demand system is

$$\begin{aligned}
\pi_{od}^{GEV}(\mathbf{P}_d, X_d; \theta) &= \frac{P_{od}^{-\theta} G_o^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})}{G^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})} \\
&= \frac{P_{od}^{-\theta}\left[\mathbb{E}\left(\sum_{o'}(U_{o'd}(v)^\theta P_{o'd}^{-\theta})\right)^{\frac{\sigma}{\theta}}\right]^{\frac{\theta}{\sigma}-1} \mathbb{E}\left[\left(\sum_{o'}(U_{o'd}(v)^\theta P_{o'd}^{-\theta})\right)^{\frac{\sigma}{\theta}-1} U_{od}(v)^\theta\right]}{G^d(P_{1d}^{-\theta}, \ldots, P_{Nd}^{-\theta})} \\
&= \frac{\mathbb{E}\left[\left(\sum_{o'}(U_{o'd}(v)^\theta P_{o'd}^{-\theta})\right)^{\frac{\sigma}{\theta}-1} U_{od}(v)^\theta P_{od}^{-\theta}\right]}{\mathbb{E}\left(\sum_{o'}(U_{o'd}(v)^\theta P_{o'd}^{-\theta})\right)^{\frac{\sigma}{\theta}}} \\
&\overset{\theta\to\infty}{\to} \frac{\mathbb{E}\left[(P_{od}/U_{od}(v))^{-\sigma}\mathbf{1}\left\{P_{od}/U_{od}(v) = \min_{o'} P_{o'd}/U_{o'd}(v)\right\}\right]}{\mathbb{E}\left[(\min_{o'} P_{o'd}/U_{o'd}(v))^{-\sigma}\right]} = \pi_{od}(\mathbf{P}_d, X_d).
\end{aligned}$$

That is, the implied GEV import demand system converges point wise to the true demand system. To establish uniform convergence across $\mathbf{P}_d \in K$ for K compact, note that if the sequence $\{\pi_{od}^{GEV}(\mathbf{P}_d, X_d; \theta_j)\}_{j=1}^\infty$ is convergent, there exists a positive sequence $\{\theta_k\}_{k=1}^\infty$ that diverges such that $\{\pi_{od}^{GEV}(\mathbf{P}_d, X_d; \theta_k)\}_{k=1}^\infty$ is monotone and converges. Then since $\pi_{od}(\mathbf{P}_d, X_d)$ is continuous we can apply Theorem 7.13 in Rudin et al. (1964) to establish uniform convergence. □

G Proof of Proposition 5

Proof. Define the excess demand function $E : \mathbb{R}_+^N \to \mathbb{R}^N$ as satisfying

$$E_o(\mathbf{W}) = -W_o L_o + \sum_{d=1}^N \pi_{od}(W_d L_d - TB_d) \quad \text{for each} \quad o = 1, \ldots, N.$$

Since productivity has a global innovation representation, by Theorem 1, it is Fréchet. Thus, by Lemma 1, G^d is a correlation function. As a result, we must have $\pi_{od} > 0$ for any finite competitiveness indices.

The implication is that the excess demand system satisfies strict gross substitutability. For each $o = 1, \ldots, N$ and each $n \neq o$ we have

$$\begin{aligned}\frac{\partial E_o(\mathbf{W})}{\partial W_n} &= \sum_{d=1}^{N} \frac{\partial}{\partial W_n} \frac{P_{od}^{-\theta} G_o^d}{G^d}(W_d L_d - TB_d) \\ &= \sum_{d=1}^{N} \frac{P_{od}^{-\theta}}{G^d} \underbrace{\left(G_{on}^d - \frac{G_o^d G_n^d}{G^d}\right)}_{\leq 0} \underbrace{(W_d L_d - TB_d)}_{=X_d>0} \underbrace{\frac{\partial P_{Nd}^{-\theta}}{\partial W_n}}_{<0} + \frac{P_{on}^{-\theta} G_o^n}{G^d} L_n \\ &\geq \frac{P_{on}^{-\theta} G_o^n}{G^d} L_n = \pi_{on} L_n > 0.\end{aligned}$$

The first inequality in the second line follows from the differentiability restriction on the correlation function. The final strict inequality follows from $\pi_{on} > 0$. Given that the excess demand function is homogenous of degree one and satisfies strict gross substitutability, we can apply Proposition 17.F.3 of Mas-Collell et al. (1995) to establish existence and uniqueness. Moreover, a *tâtonnement* process, as in Alvarez and Lucas (2007), can be used to solve for the equilibrium. □

H Proof of Proposition 6

Proof. Since micro productivity has a global innovation representation, by Theorem 1 extended to account for the micro dimension m, its distribution is multivariate θ-Fréchet with location parameters of $T_{mod} \equiv \mathbb{E} A_{imod}(v)^\theta$ and micro correlation function

$$F^d(\mathbf{X}_1, \ldots, \mathbf{X}_N) \equiv \mathbb{E} \frac{A_{imod}(v)^\theta}{T_{mod}} x_{mo}.$$

Define implied macro applicability of

$$A_{iod}(v) \equiv \max_{m=1,\ldots,M_o} A_{imod}(v).$$

Then, since $A_{mod}(v) \equiv \max_{i=1,2,\ldots} Z_i(v) A_{imod}(v)$ we have macro productivity of

$$A_{od}(v) \equiv \max_{m=1,\ldots,M_o} A_{mod}(v) = \max_{m=1,\ldots,M_o} \max_{i=1,2,\ldots} Z_i(v) A_{imod}(v) = \max_{i=1,2,\ldots} Z_i(v) A_{iod}(v).$$

As a result, $A_{od}(v)$ satisfies Assumption 1. Since micro applicability is $i.i.d.$ over (i, v), so is macro applicability, and Assumption 2 is satisfied. Finally, Assumption 3 holds for $Z_i(v)$ since it is global productivity from the global innovation representation of micro productivity. Therefore, macro productivity has a global innovation representation.

Then, by Theorem 1, macro productivity is multivariate θ-Fréchet. Its location parameters are

$$T_{od} \equiv \mathbb{E}A_{iod}(v)^{\theta} = \mathbb{E} \max_{o'=1,\ldots,N} \max_{m=1,\ldots,M_{o'}} \frac{A_{imo'd}(v)^{\theta}}{T_{mo'd}} \mathbf{1}\{o = o'\} T_{mod},$$

and so

$$T_{od} = F^{d}\left(\mathbf{0}_{1}, \ldots, \mathbf{0}_{o-1}, \mathbf{T}_{od}, \mathbf{0}_{o+1}, \ldots, \mathbf{0}_{N}\right) \equiv F^{od}\left(\mathbf{T}_{od}\right),$$

for $\mathbf{T}_{od} \equiv (T_{1od}, \ldots, T_{M_o od})$. Its correlation function is

$$G^{d}(x_1, \ldots, x_N) \equiv \mathbb{E} \max_{o=1,\ldots,N} \max_{m=1,\ldots,M_{o}} \frac{A_{imod}(v)^{\theta}}{T_{mod}} \omega_{mod} x_o \equiv F^{d}(\mathbf{\Omega}_{1d} x_1, \ldots, \mathbf{\Omega}_{Nd} x_N),$$

for $\omega_{mod} \equiv T_{mod}/T_{od}$ and $\mathbf{\Omega}_{od} \equiv (\omega_{1od}, \ldots, \omega_{M_o od})$.

□

I Extensions

I.1 Armington

This section introduces an extension of the framework that accommodates models of monopolistic competition. It also provides an alternative micro foundation based on random utility at the individual consumer level which aggregates to imply that price-adjusted utility of individual consumers is multivariate θ-Fréchet.

Variety $v \in V$ is produced by a single firm in origin country $o(v)$ and their product has characteristic $m(v) \in \{1, \ldots, M_o\}$. The product characteristic can index any countable set of micro factors relevant for the production of the firm's good—e.g., the classification of a product into a specific sector, the production method used to produce the good, the home country of a multinational operating the firm, a sub-region of the country where the variety is produced, or the sequence of locations (value chain) along which the good is produced.

The set of destinations where v is delivered is $D(v)$. The set of products with characteristic m produced in origin o and available in destination d is $V_{mod} = \{v \in V \mid m(v) = m, o(v) = o, d \in D(v)\}$. We denote the set of varieties available in d and produced in o by $V_{od} \equiv \cup_{m=1}^{M_o} V_{mod}$ and the set of all varieties available in d by $V_d \equiv \cup_{o=1}^{N} V_{od}$.

Consumers in each country have random utility over varieties and are endowed with a unit of time that they supply inelastically in a competitive labor market. The set of consumers in destination d is I_d. Let W_d denote the wage in d and $T(i)$ any lump sump transfers to the household (e.g., from firm profits). Household $i \in I_d$ solves

$$\begin{aligned} \max_{C_d(\cdot,i)\geq 0} \quad & \int_{V_d} U_d(v,i)C_d(v,i)\mathrm{d}v \\ \text{s.t.} \quad & \int_{V_d} P_d(v)C_d(v,i)\mathrm{d}v\mathrm{d}v \leq W_d + T(i), \end{aligned}$$

where $P_d(v)$ is the price that firm $v \in V_d$ charges for their product in destination d. We can interpret the random quantity $U_d(v,i)$ as consumer i's perception of the quality of variety v and allows us to incorporate heterogeneity in preferences for products as a source of comparative advantage—in the spirit of Armington models of trade (such as Anderson, 1979).

Since varieties are perfect substitutes, individual consumers will purchase whichever variety they perceive to have the lowest quality-adjust price. We can break this decision into first choosing a variety within V_{mod}

$$C_d(v,i) > 0 \quad \Longrightarrow \quad \frac{P_d(v)}{U_d(v,i)} = \inf_{v\in V_{mod}} \frac{P_d(v)}{U_d(v,i)},$$

and then an origin and product characteristic,

$$C_d(v,i) > 0 \quad \Longrightarrow \quad \inf_{v\in V_{mod}} \frac{P_d(v)}{U_d(v,i)} = \min_{o=1,\ldots,N} \min_{m=1,\ldots,M_o} \inf_{v\in V_{mod}} \frac{P_d(v)}{U_d(v,i)}.$$

Additive random utility at the individual consumer level ensures that the pattern of demand across origins and product characteristics is determined by cost-minimization.

We assume that $\{U_d(v,i)\}_{v\in V}$ is a θ-Fréchet process with scale parameters $\{Q_d(v)^\theta\}_{v\in V}$ and correlation functional $\tilde{F}^d : \{x : V \to \mathbb{R}_+\} \to \mathbb{R}_+$ of the form

$$\tilde{F}^d(x(\cdot)) = F^d(\mathbf{X}_{1d},\ldots,\mathbf{X}_{Nd}) \text{ where } \mathbf{X}_{od} \equiv (x_{1od},\ldots,x_{M_o od}), x_{mod} = \left(\int_{V_{mod}} x(v)^{\frac{1}{1-\rho_{mod}}}\right)^{1-\rho_{mod}},$$

where F^d is a micro correlation function. This nested-CES structure implies that $\sup_{v\in V_{mod}} \frac{U_d(v,i)}{P_d(v)}$ is multivariate θ-Fréchet with scale parameters of $\left[\int_{V_{mod}} (P_d(v)/Q_d(v))^{-\frac{\theta}{1-\rho_{mod}}}\right]^{1-\rho_{mod}}$ and correlation function F^d.

Define the (expenditure) elasticity of substitution within V_{mod}, $\sigma_{mod} \equiv \theta/(1-\rho_{mod})$, and a price index as

$$P_{mod} \equiv \left[\int_{V_{mod}} \left(P_d(v)/Q_d(v) \right)^{-\sigma_{mod}} \right]^{-\frac{1}{\sigma_{mod}}}.$$

Then the density of consumers in d that purchase variety v is

$$\pi_{mod}(v) = \left(\frac{P_d(v)/Q_d(v)}{P_{mod}} \right)^{-\sigma_{mod}} \pi_{mod}, \tag{I.1}$$

with the share of expenditure of consumers in d on type-m goods from country o equal to

$$\pi_{mod} = \frac{P_{mod}^{-\theta} F_{mo}^{d}(\mathbf{P}_{1d}^{-\theta}, \ldots, \mathbf{P}_{Nd}^{-\theta})}{F^d(\mathbf{P}_{1d}^{-\theta}, \ldots, \mathbf{P}_{Nd}^{-\theta})},$$

where $\mathbf{P}_{od} \equiv (P_{1od}, \ldots, P_{M_o od})$. Total expenditure is $X_d = W_d L_d + T_d$, where L_d is the measure of the set I_d and $T_d = \int_{I_d} T(i) \mathrm{d}i$ total transfers to d from profits. Demand across firms with the same characteristic m and origin o is CES and depends on quality-adjusted prices relative to the price index P_{mod}. The pattern of demand across product characteristics takes a GEV form.

I.2 Monopolistic Competition

The specification for preferences in the previous section enables us to incorporate monopolistic competition. Faced with a CES demand curve, monopolistic firm v sets the price in destination d as a constant markup over marginal cost.[22] Denote the marginal cost of firm $v \in V_{mod}$ when delivering to d by $\widetilde{MC}_d(v)$ and quality-adjusted marginal cost by $MC_d(v) \equiv \widetilde{MC}_d(v)/Q_d(v)$.

The quality-adjusted price of variety $v \in V_{mod}$ in d is then

$$\frac{P_d(v)}{Q_d(v)} = \frac{\sigma_{mod} + 1}{\sigma_{mod}} MC_d(v) \quad \forall v \in V_{mod}; \tag{I.2}$$

revenue shares from variety v are

$$\pi_{mod}(v) = \frac{MC_d(v)^{-\sigma_{mod}}}{\int_{V_{mod}} MC_d(v')^{-\sigma_{mod}} \mathrm{d}v'} \pi_{mod}; \tag{I.3}$$

[22] The CES demand curve at the variety level in (I.1) comes from our assumption on the process for household utility. A different distribution will generate non-CES demand and variable markups. See Footnote 5 for references.

profits are

$$\Pi_{mod}(v) = \frac{1}{\sigma_{mod}+1}\pi_{mod}(v)X_d; \tag{I.4}$$

and the price index is

$$P_{mod} = \frac{\sigma_{mod}+1}{\sigma_{mod}}\left[\int_{V_{mod}} MC_d(v)^{-\sigma_{mod}}\right]^{-\frac{1}{\sigma_{mod}}}.$$

We can now consider how different specifications for production determine marginal costs and therefore import prices.

I.3 Global Value Chains / Shipping

We allow for production to use labor inputs from multiple countries. This setup can accommodate models of global value chains (see Antràs and de Gortari, 2017) —where the index m specifies a specific value chain—as well as models incorporating a production structure for shipping goods—in which case m would specify the route along which goods get shipped.

Let $\mathbf{W} = (W_1, \ldots, W_N)$ denote the vector of wages, and $\mathbf{L}_d(v) = (L_{1d}(v), \ldots, L_{Nd}(v))$ the vector of labor used for producing goods for delivery in destination d. We assume that firms have constant returns to scale production and face competitive labor markets. The marginal cost for firm $v \in V_{mod}$ when delivering to d is

$$\begin{aligned} MC_d(v) = \frac{c_{mod}(\mathbf{W})}{A_d(v)} \equiv \min_{\mathbf{L}_d(v)} \quad & \mathbf{W}'\mathbf{L}_d(v) \\ \text{s.t.} \quad & 1 \leq A_d(v)F_{mod}(\mathbf{L}_d(v)), \end{aligned}$$

where $A_d(v)$ is quality-adjusted productivity of firm v when producing for market d. The production function $F_{mod}(\mathbf{L}_d)$ is common across varieties with the same product characteristics, origin, and destination. It is homogenous of degree one and so $c_{mod}(\mathbf{W})$ is also homogenous of degree one in $\mathbf{W}$. To get a demand system in terms of wages, we now must incorporate these cost functions into the demand system. The price indices are related to the cost functions as

$$P_{mod} = \frac{c_{mod}(\mathbf{W})}{A_{mod}} \quad \text{where} \quad A_{mod} \equiv \frac{\sigma_{mod}}{\sigma_{mod}+1}\left(\int_{V_{mod}} A_d(v)^{\sigma_{mod}}\right)^{\frac{1}{\sigma_{mod}}}.$$

Define a productivity index $A_{mo} \equiv A_{moo}$ and iceberg indices as $\tau_{mod} \equiv A_{moo}/A_{mod}$. Then

$$P_{mod} = \tau_{mod} \frac{c_{mod}(\mathbf{W})}{A_{mo}},$$

Our model of production has pinned down import prices as a function of trade costs, productivities, and normalized marginal costs. Note that if we assume that the production function is common across destination markets—imposing that inputs used in production do not depend on the destination where a good gets sold—then the cost function does not depend on d. In this case, we have

$$P_{mod} = \tau_{mod} \frac{c_{mo}(\mathbf{W})}{A_{mo}},$$

so that the underlying structure of production is fully absorbed into a micro-factor-origin fixed effect. Note that it is this fixed effect that captures comparative advantage. The implication is that, as long as such a fixed effect is included in the estimation, we can use exogenous variation in trade costs to get exogenous variation in import prices.

I.4 Selection

The model accommodates standard models of trade based on monopolistic competition and heterogenous firms, such as Melitz (2003) and Chaney (2008), as we show next.

We take the set of global varieties V as fixed. Each variety v is associated with an origin $o(v)$. They then choose which destinations they will deliver their product to. Let $f_{mod}(\mathbf{W})$ denote the fixed cost of introducing a variety with characteristics m produced in o into destination market d. The threshold level of quality-adjusted productivity at which varieties enter d is A^*_{mod} satisfying

$$f_{mod}(\mathbf{W}) = \frac{1}{\sigma_{mod}} \frac{{A^*_{mod}}^{\sigma_{mod}}}{\int_{V_{mod}} A_d(v)^{\sigma_{mod}} \mathrm{d}v} \pi_{mod} X_d. \tag{I.5}$$

As a result, the set of varieties with characteristics m offered in d from o satisfies $V_{mod} = \{v \in V_{mo} \mid A^*_{mod} < A_d(v)\}$, where A^*_{mod} is given by (I.5).

Assume that $A_d(v)$ is distributed Pareto among varieties in $v \in V_{mo}$ and independent across d. The shape parameter is κ_{mod}, with $\kappa_{mod} > \sigma_{mod}$, and the lower bound of the distribution is $\underline{A}_{mod}$. Define $\eta_{mod} \equiv (\kappa_{mod} - \sigma_{mod})/\kappa_{mod}$, with $\eta_{mod} \in (0, 1)$. The measure of varieties from o in d with characteristics m is

$$M_{mod} = \left(\frac{A^*_{mod}}{\underline{A}_{mod}}\right)^{-\kappa_{mod}} M_{mo}, \tag{I.6}$$

where M_{mo} is the measure of V_{mo}. Assuming that not all $v \in V_{mo}$ enter each destination market, and using (I.5) and (I.6), the quality-adjusted productivity threshold is

$$A^*_{mod} = \left(\frac{\sigma_{mod} f_{mod}(\mathbf{W})}{\eta_{mod} X_{mod}} M_{mo}\right)^{\frac{1}{\kappa_{mod}}} \underline{A}_{mod}, \tag{I.7}$$

The implied productivity index A_{mod} is then

$$A_{mod} = \frac{\sigma_{mod}}{\sigma_{mod}+1}\left[\left(\frac{\eta_{mod}}{\sigma_{mod}}\frac{\pi_{mod} X_d}{f_{mod}(\mathbf{W}) M_{mo}}\right)^{\eta_{mod}} \frac{M_{mo}}{\eta_{mod}}\right]^{\frac{1}{\sigma_{mod}}} \underline{A}_{mod} \equiv s_{mod}(\pi_{mod} X_d, \mathbf{W})\underline{A}_{mod}.$$

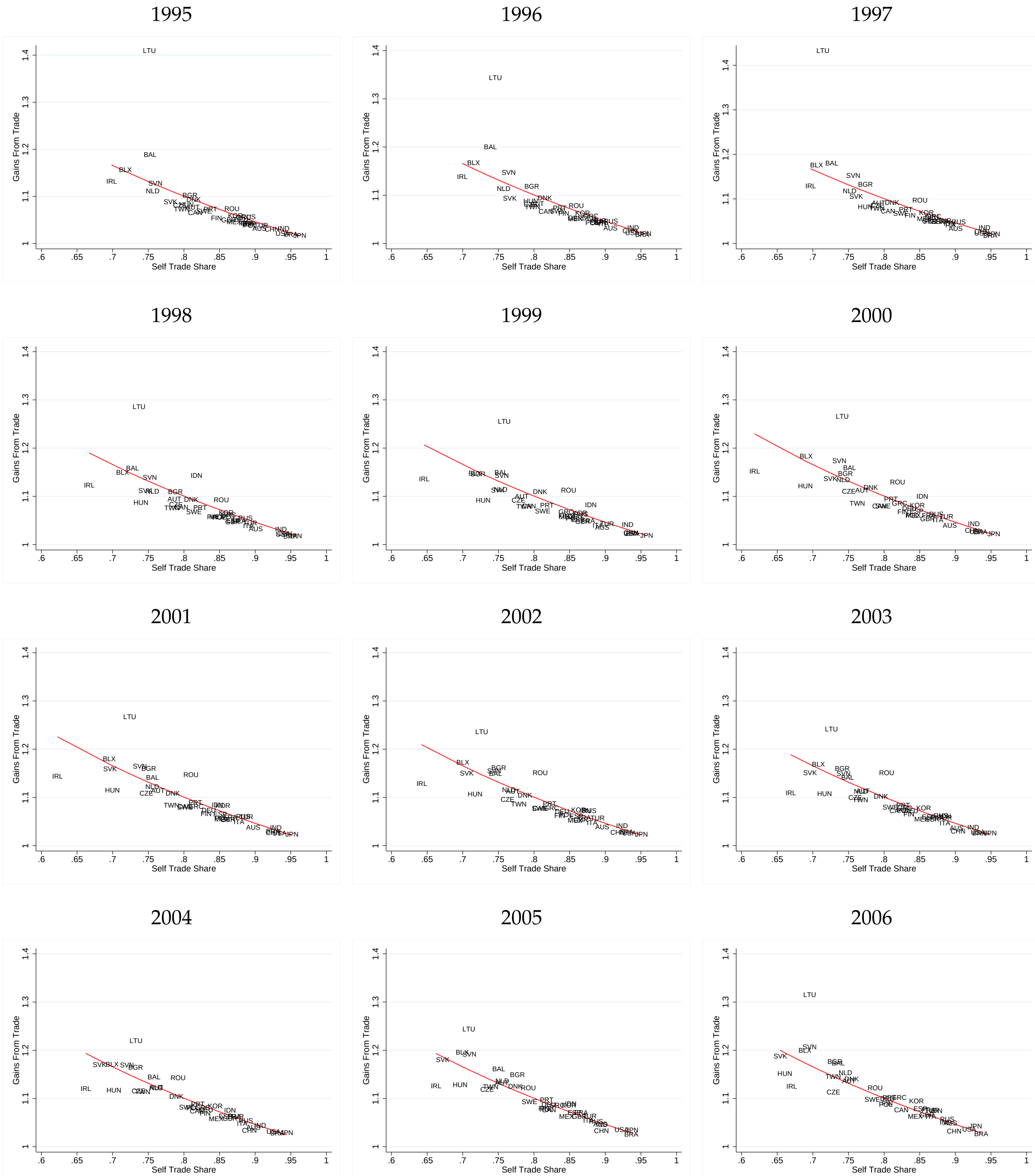
Productivity depends endogenously on expenditure and wages. Selection effects, captured by the homogenous of degree zero function s_{mod}, imply that expenditure influences productivity. The implied import prices are now

$$P_{mod} = \tau_{mod} \frac{c_{mod}(\mathbf{W})/A_{mo}}{s_{mod}(\pi_{mod} X_d, \mathbf{W})}$$

for $A_{mo} \equiv \bar{A}_{moo}$ and $\tau_{mod} \equiv \bar{A}_{moo}/\bar{A}_{mod}$. Combined with our previous import demand system, we now have an implicit relationship between expenditure shares and trade costs (the exogenous component of import prices). Viewed from the perspective of how expenditure depends on import prices, we are still in the same class of import demand systems. Viewed from the perspective of how expenditure depends on trade costs, this implicit demand system may not necessarily remain in the GEV class. The GEV class implies substitutability and the strategic complementarity associated with firms selecting into exporting may induce this implicit relation between trade flows and trade costs to exhibit complementarity.

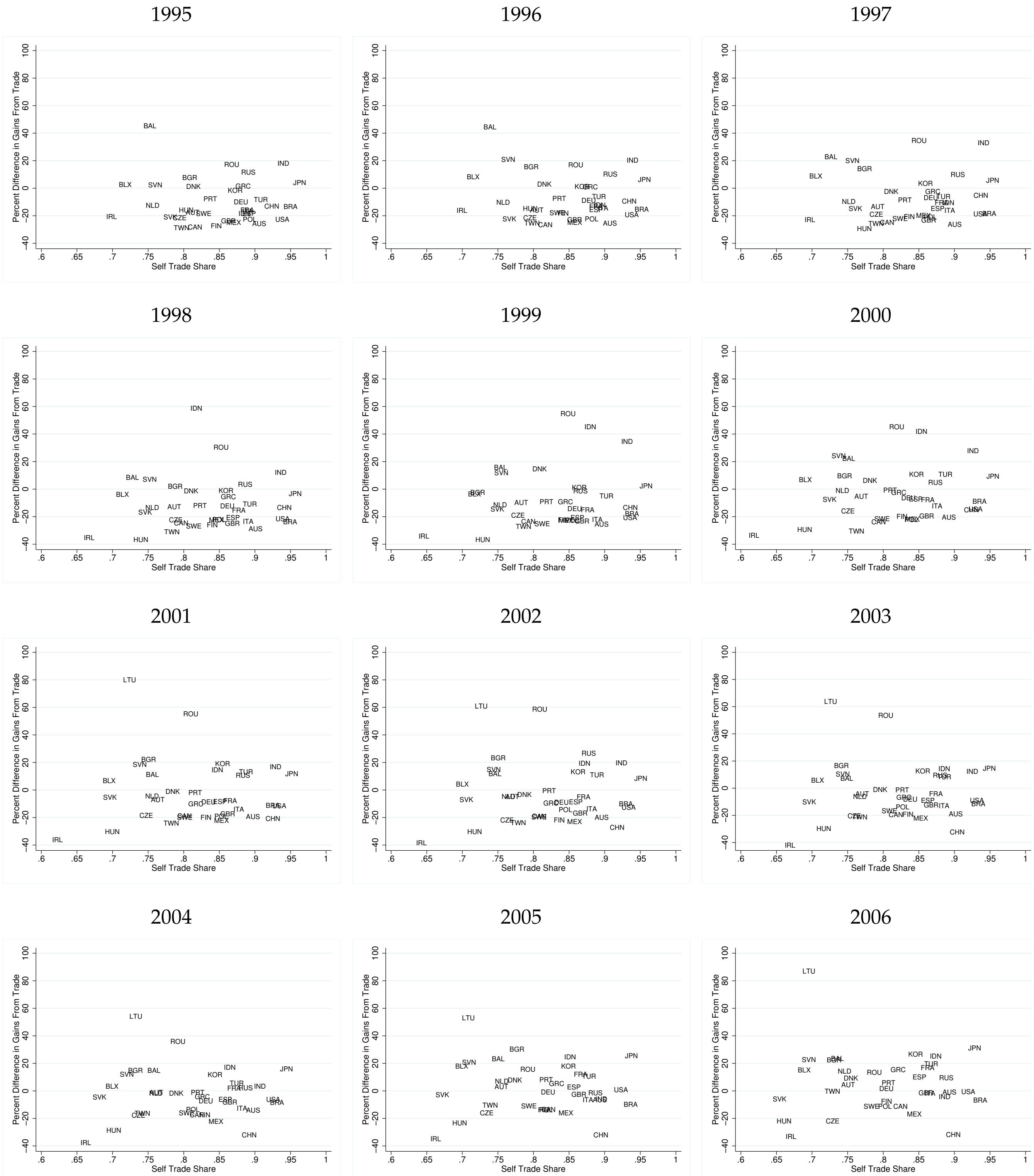
J Additional figures

Figure J.1: Gains from Trade, levels, by year.



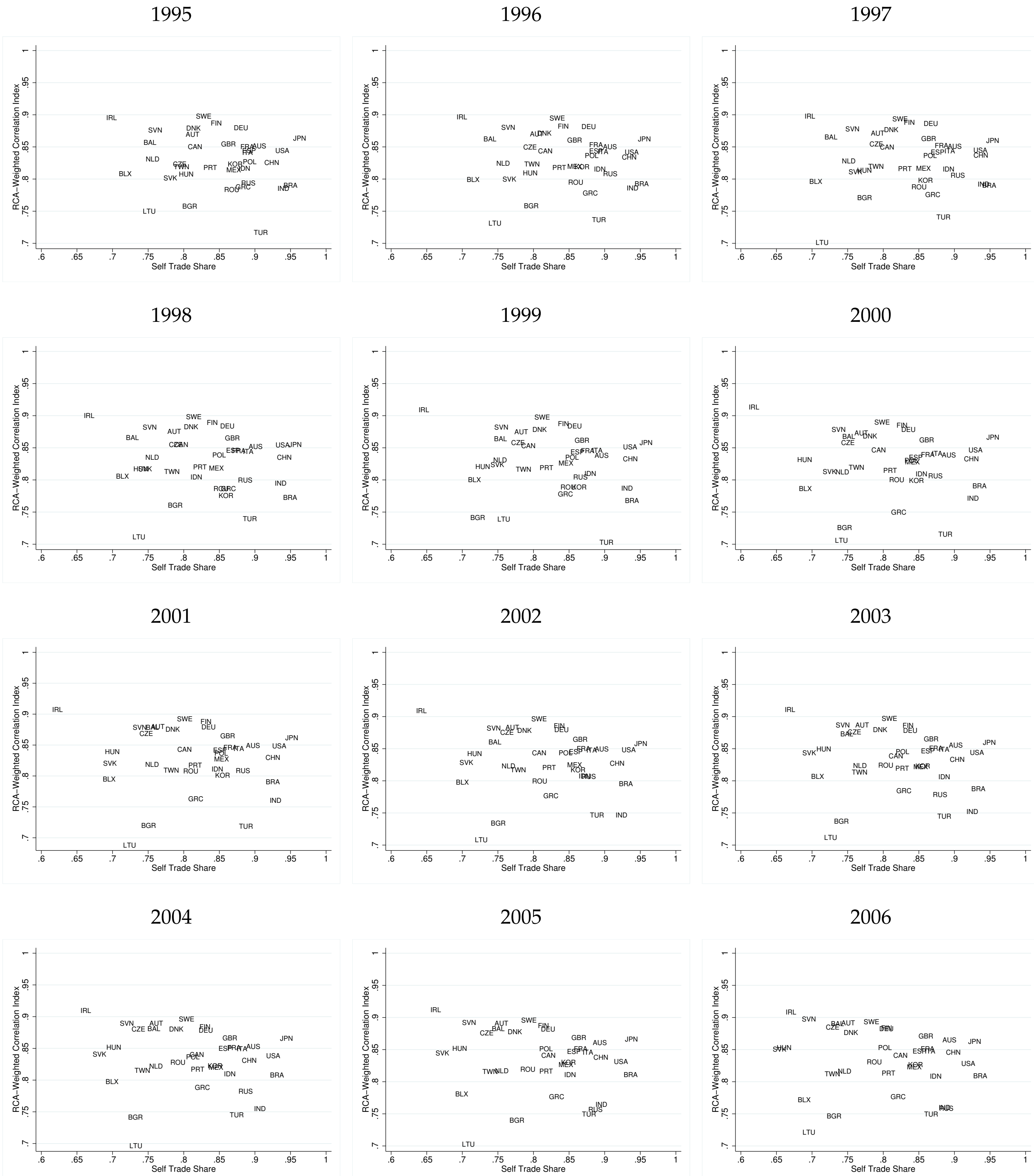
Notes: Black data: $GT_d^{CN} = (\sum_s \pi_{sdd}/G_{dd})^{-1/\theta}$. Red line: $GT_d^{CES} = (\sum_s \pi_{sdd})^{-1/\theta^{ces}}$.

Figure J.2: Gains from Trade, percentage changes, by year.



Notes: Black data: $GT_d^{CN} = (\sum_s \pi_{sdd}/G_{dd})^{-1/\theta}$. Red line: $GT_d^{CES} = (\sum_s \pi_{sdd})^{-1/\theta^{ces}}$. Percent difference calculated as $100 \times \frac{GT_d^{CN} - GT_d^{CES}}{GT_d^{CES} - 1}$.

Figure J.3: RCA-Weighted Correlation Index, by year.



Notes: Revealed-Comparative-Advantage (RCA)-weighted correlation index: $\rho_{dt}^{RCA} = \sum_s \rho_s \frac{RCA_{sdt}}{\sum_{s'} RCA_{s'dt}}$.

Figure J.4: Evolution of RCA-Weighted Correlation Index, by country.

Notes: Revealed-Comparative-Advantage (RCA)-weighted correlation index: $\rho_{dt}^{RCA} = \sum_s \rho_s \frac{RCA_{sdt}}{\sum_{s'} RCA_{s'dt}}$. Baltic Republics are Estonia, Latvia, and Lithuania; Benelux is Belgium, Luxembourg, and Netherlands.

Figure J.5: RCA-Weighted Correlation Index and Self Trade Shares, by country.

Notes: Revealed-Comparative-Advantage (RCA)-weighted correlation index: $\rho_{dt}^{RCA} = \sum_s \rho_s \frac{RCA_{sdt}}{\sum_{s'} RCA_{s'dt}}$. Baltic Republics are Estonia, Latvia, and Lithuania; Benelux is Belgium, Luxembourg, and Netherlands.