Why Aren't Welfare Gains from Trade Increasing Overtime? *

Elizaveta Archanskaia[†]

December 2014

Abstract

This paper shows that the welfare cost of autarky increased by just 1 to 3 percentage points of real income over 1963-2009. This result is attributable to a 72% increase in the trade elasticity that has offset the impact of increased reliance on foreign supply. The estimation is guided by a theoretical framework that obtains a trade elasticity jointly determined by the degree of product substitutability and of producer heterogeneity in the economy. Structural supply and demand parameters are identified in cross-section with parsimonious data requirements. In magnitude, the trade elasticity increased from 2.65 to 4.56 over 1963-2009.

Keywords: Trade elasticity, Heterogeneity, Armington, Welfare gains
JEL codes: F11, F14, F15

*This paper has greatly benefited from helpful discussions with Thomas Chaney, Lionel Fontagné, Guillaume Daudin, Jacques Le Cacheux, Thierry Mayer, André Romahn, Johannes van Biesebroeck, and Frank Verboven. Financial support of the Cournot Foundation is gratefully acknowledged.

†KU Leuven; Naamsestraat 69, 3000 Leuven, Belgium. liza.archanskaia@kuleuven.be

1 Introduction

If a trade economist were lucky to meet a kind-hearted fairy who would agree to reveal the magnitude of one structural parameter, the trade economist would likely enquire about the magnitude of the trade elasticity.

It has long been known from CGE modelling that the single most important parameter for policy analysis measures the sensitivity of trade volumes to cost shocks.[1] The interest intensified once Arkolakis et al. (2012) had shown that this parameter, dubbed the 'trade elasticity', determined almost single-handedly the welfare cost of autarky. This happens because the trade elasticity summarizes the degree of structural heterogeneity in the economy thereby capturing the strength of the incentive to trade. More heterogeneity means stronger complementarity among trade partners in some model-specific dimension. This entails greater losses in real income from shutting down trade.

Microfoundations come into play because the strength of the incentive to trade is likely to be model-specific. Melitz and Redding (2013) show that reliance on foreign supply is model-specific in equilibrium if the trade elasticity is constrained to be equal across microfoundations. Taking the distribution of retail prices and market shares as given, Simonovska and Waugh (2014a) show that it maps into different magnitudes of the trade elasticity in models that single out different dimensions of heterogeneity. Moreover, the incentive to trade is likely increasing in the number of heterogeneity dimensions incorporated in the model (Costinot and Rodriguez-Clare (2014), Levchenko and Zhang (2014)).

The starting point of this paper is the lack of empirical evidence on the evolution of the trade elasticity (Head and Mayer (2013)).[2] Moreover, there remains substantial uncertainty about the relative contribution of heterogeneity in sup-

[1] The focus shifted from substitutability in final goods (Armington (1969), Reinert and Roland-Horst (1992), Imbs and Méjean (2010)) to substitutability in inputs (Johnson and Noguera (2012), Bems et al. (2010, 2011), Bems (2014)).

[2] Eaton and Kortum (2002), Caliendo and Parro (2014), Simonovska and Waugh (2014a,b) provide estimates for a single year.

ply and in demand to determining this parameter. In particular, the finding that demand parameters have no incidence on the magnitude of the trade elasticity is specific to models with producer heterogeneity in which product differentiation occurs at one tier of the CES utility function. Demand comes back into the picture if substitutability of goods of different origin and of goods produced in different sectors do not coincide (Costinot and Rodriguez-Clare (2014), Feenstra et al. (2014), Imbs and Méjean (2014)).

To learn from the data which dimension of heterogeneity matters, the model combines cost heterogeneity in the spirit of Eaton and Kortum (2002) with two-tier CES preferences. At the lower tier, sectoral goods are combined within a country-specific composite good. Composite goods of different national origin are combined at the upper tier. The model delivers the gravity formulation in Anderson and van Wincoop (2003) if all sectoral goods are traded. It delivers the gravity formulation in Eaton and Kortum (2002) if only a subset of goods is traded, and upper- and lower-tier substitutability coincide. In the most general case, the three dimensions jointly determine the trade elasticity.[3]

This simple generalization of the Armington model makes it feasible to identify the lower-tier elasticity, the upper-tier (Armington) elasticity, and the degree of dispersion in sectoral technology draws in cross-section with parsimonious data requirements.[4] The key intuition is that incorporating more dimensions of heterogeneity in the model helps to decompose price and expenditure variation along the three axis uniquely attributable to each structural parameter. While Feenstra et al. (2014) work with a three-tier CES structure and use the time dimension of the data to identify sector-specific demand elasticities, I use the cross-sectional dimension of the data to identify structural parameters relevant

[3] Feenstra et al. (2014) and Costinot and Rodriguez-Clare (2014) obtain a qualitatively similar result but the exact expression of the trade elasticity is model-specific.

[4] The parameters are identified in the absence of data on bilateral frictions used in Caliendo and Parro (2014) and in the absence of data on retail prices used in Eaton and Kortum (2002) and Simonovska and Waugh (2014a).

for the economy as a whole.[5]

The main result is that the magnitude of the trade elasticity has increased from 2.65 in 1963 to 4.56 in 2009. This 72% increase is driven by a 30% reduction in product substitutability within the bundle and a 16% increase in bundle substitutability of different national origin. The evolution of structural demand parameters indicates a shift from ‘location’- to ‘brand’-based product differentiation that reduced the magnitude of variety gains from trade. The incentive to trade has shifted towards the supply channel of cost reductions achieved through access to more efficient producers and resulted in an increased sensitivity of trade to trade costs.

The corollary is that the magnitude of welfare gains from trade is only weakly increasing overtime. The non-increasing welfare cost of autarky is due to the evolution of the trade elasticity that has offset the effect of increased reliance on foreign supply. For the interquartile range of the country sample, the percentage loss in real income that would be incurred by reverting to autarky is situated between 11 and 22% in 2009.[6]

The paper is structured in three parts. Sec.2 presents the theoretical framework. Sec.3 outlines the estimation strategy and reports annual estimates of lower- and upper-tier elasticities. Sec.4 presents empirical evidence on producer cost dispersion, computes the annual magnitude of the trade elasticity, and reports the welfare cost of autarky between 1963 and 2009.

[5] In the spirit of Feenstra (1994) the Armington parameter is identified by implementing the between estimator. In cross-section, I exploit variation in *average* expenditure on the exporter-specific bundle across the set of active destination markets and variation in the price of the bundle as predicted by fundamental exporter ability.

[6] Similar numbers are reported in Costinot and Rodriguez-Clare (2014) for 2008 in the multiple-sector model.

2 The model

In the Armington set-up countries provide the world market with a composite good produced by homogeneous firms (Anderson and van Wincoop (2003), Armington (1969)). This paper proposes a simple generalization of the Armington model that relaxes the implausible assumption of producer homogeneneity and generates a distribution of prices for country-specific sectoral output (sec.2.1). The trade elasticity is directly determined by the Armington elasticity if all sectoral goods are traded (sec.2.1). If only a subset of goods is traded, the magnitude of the trade elasticity is codetermined by the Armington parameter and the gap between producer cost heterogeneity and lower-tier product substitutability (sec.2.2). Hence, truncation does not invalidate the key Armington intuition that the elasticity of aggregate trade to trade costs reflects perceived substitutability of composite goods that countries deliver to the world market.

2.1 A simple generalization of the Armington model

The world contains N countries with labor endowment L_i in each. Output is produced using labor which is perfectly mobile across sectors and immobile across countries. Production technology is non-proprietory within the country and non transferable across countries.

Production technology is linear in labor, with unit labor cost denoted c_i. Output can be produced using one of the production techniques for sector k available in country i. Production techniques vary in efficiency z. Techniques are drawn independently in each sector from a common distribution. For consistency with the assumption of non-proprietory technology, varieties of good k produced within the same country are taken to be perfect substitutes. Constant returns to scale and within-sectoral product homogeneity entail that the best available technique is used in production of each sector within the country. Nonetheless, techniques may differ across sectors within the country and across countries for any given sector.

Technology improvement follows the Poisson process described in Eaton and Kortum (2010) whereby at each point in time the number of techniques available for producing output in sector k with efficiency $Z > z$ follows a Poisson distribution with parameter $\lambda_i(t) = T_i(t)z^{-\theta}$. This parameter is increasing in $T_i(t)$ which denotes the stock of technology accumulated in country i by time t and in $1/\theta$ which denotes the extent of dispersion in technology draws.[7] This parameter maps fundamental exporter ability into the number of goods that can be produced with efficiency higher than any given threshold z.

Given a Poisson process for the arrival of ideas and a stock of technology T_i, the probability of no technique with efficiency $Z > z$ arriving in a unit interval in sector k is given by the Poisson density for $X = 0$, where X is the number of draws with efficiency higher than z:

$$\Pr[Z \leq z] \quad = \quad \Pr[X = 0] = \frac{(\lambda_i)^0 \exp\{-\lambda_i\}}{0!} = \exp\{-\lambda_i\} \tag{1}$$

The probability that a technique of higher efficiency occurs is given by:

$$\Pr[Z > z] \quad = \quad 1 - \Pr[X = 0] = 1 - \exp\left\{-T_i z^{-\theta}\right\} \tag{2}$$

As the process of technology upgrading takes place independently within each sector in the unit continuum, this probability distribution also characterizes the cross-sectoral distribution of best-of ideas in each country. The structure of production thus replicates Eaton and Kortum (2002) wherein techniques effectively used in production are distributed Fréchet. But the structure of preferences, to which we now turn, replicates the Armington hypothesis of product

[7] θ is the shape parameter of the Pareto distribution from which efficiency is drawn. A lower θ corresponds to a distribution with a fatter tail, e.g. a higher probability of getting a high draw (Eaton and Kortum (2010)).

bundles differentiated by place of origin.[8]

Consumer preferences are assumed well represented by a two-tier CES utility function. At the lower-tier country-specific sectoral goods are combined into a composite product bundle. At the upper-tier composite goods of different national origin are combined into an aggregate consumption good.

This set-up is chosen for two reasons. First, we seek to give substance to the Anderson and van Wincoop (2003) concept of country-specific 'composite goods' exchanged on the world market while accomodating price heterogeneity of sectoral output in the data. Second, this set-up makes identification of structural parameters feasible notwithstanding the lack of firm-level information in the widely available trade data.[9]

Overall utility is:

$$U = \sum_{i=1}^{N} \left\{ Q_i{}^{(\sigma-1)/\sigma} \right\}^{\sigma/(\sigma-1)} \tag{3}$$

where the country-specific composite good Q_i is:

$$Q_i = \left[\int_0^1 Q_i(k)^{\frac{\sigma'-1}{\sigma'}} \, \mathrm{d}k \right]^{\frac{\sigma'}{\sigma'-1}} \tag{4}$$

Parameter restrictions $1 < \sigma \leq \sigma'$ ensure that finite positive utility is attained in autarky. The welfare cost of autarky corresponds to the reduction in real income brought about by restricting consumption to the domestic composite good (Arkolakis et al. (2012)).

Define expenditure on the country-specific sectoral good $X_i(k) = P_i(k)Q_i(k)$

[8] Recall that in Eaton and Kortum (2002) countries supply homogeneous sectoral goods, and the consumer only cares about the combination of least-cost goods in the unit continuum.

[9] If firm-level information were available the set-up would be modified to contain within-sectoral combination of varieties into composite sectoral goods of different national origin at the lower tier. Armington elasticities and productivity dispersion would be identified for each sector using the methodology presented in sec.3.

and expenditure on the corresponding composite good $X_i = P_i Q_i$. The share of expenditure on each sector is:

$$\frac{X_i(k)}{X_i} = \left[\frac{P_i(k)}{P_i}\right]^{1-\sigma'} \tag{5}$$

where the price index across the unit continuum of sectors is:

$$P_i = \left[\int_0^1 P_i(k)^{1-\sigma'}\,\mathrm{d}k\right]^{\frac{1}{1-\sigma'}} \tag{6}$$

Alternatively, denoting F_i the price distribution in each source i, we obtain the lower-tier price index by aggregating across the distribution of realized prices:

$$P_i(p) = \left\{\int_0^\infty p^{1-\sigma'}\,\mathrm{d}F_i(p)\right\}^{\frac{1}{1-\sigma'}} \tag{7}$$

Efficiency is the realization of the random variable Z with independent draws for each sector from the Fréchet distribution with parameter λ. The unit cost of producing k in i is then the realization of the random variable $W = c_i/Z$. Consequently, the number of techniques which allow production of output with cost lower than some threshold w is distributed Poisson with parameter $\lambda_i = T_i(c_i/w)^{-\theta}$ (using $z = c_i/w$) where the time subscript is suppressed given our focus on expenditure allocation in cross-section. Applying (1) the probability of no technique allowing production with cost less than w arriving in a unit interval is given by $\exp\{-\lambda_i\}$. Applying (2) the probability of a lower cost draw arriving is given by $1-\exp\{-\lambda_i\}$. The distribution of lowest costs is Weibull with parameter λ_i (Eaton and Kortum (2002)):

$$F(w) = \Pr[W \le w] = 1 - \exp\left\{-T_i c_i^{-\theta} w^{\theta}\right\} \tag{8}$$

and the corresponding pdf is:

$$f(w) = T_i c_i^{-\theta} \theta w^{\theta-1} \exp\left\{-T_i c_i^{-\theta} w^{\theta}\right\} \tag{9}$$

The assumption of perfect competition within each sector entails that the distribution of realized prices is directly given by the distribution of least costs. The structure of preferences entails that all domestic goods survive to compose the country-specific composite good:

$$P_i(p)^{1-\sigma'} = P_i(w)^{1-\sigma'} = \int_0^{\infty} w^{1-\sigma'} f(w)\,\mathrm{d}w \tag{10}$$

Hence I can use Lemma 2 in Eaton and Kortum (2010) together with parameter restrictions $1 < \sigma \leq \sigma' < \theta + 1$ to compute the price of the country-specific composite good:

$$P_i = \left\{T_i c_i^{-\theta}\right\}^{-1/\theta} \{\Gamma(\gamma)\}^{1/1-\sigma'} \tag{11}$$

where $\gamma = (\theta + 1 - \sigma')/\theta$ is the parameter of the Gamma function.[10]

At the upper-tier we get the set-up in Anderson and van Wincoop (2003) whereby each country produces a single country-specific composite good Q_i, and supply of this good is perfectly inelastic. Under the assumption of iceberg trade costs t_{ij}, the scaled price of the composite good delivered from i to j is:

$$\kappa^{-1} P_{ij} = T_i^{-1/\theta} c_i \tau_{ij} \tag{12}$$

where $\tau_{ij} = 1 + t_{ij}$ and $\kappa = \{\Gamma(\gamma)\}^{1/(1-\sigma')}$ is a source-invariant scalar.

[10] The procedure in Eaton and Kortum (2010) is: plug (9) into (10); use the definition of λ to write $d\lambda = T_i c_i^{-\theta} \theta w^{\theta-1} dw$ and $(\lambda / T_i c_i^{-\theta})^{(1-\sigma')/\theta} = w^{1-\sigma'}$; change the variable of integration and rearrange (10) to get $P_i^{1-\sigma'} = \left\{T_i c_i^{-\theta}\right\}^{-(1-\sigma')/\theta} \int_0^{\infty} \lambda^{(1-\sigma')/\theta} \exp\{-\lambda\}\,\mathrm{d}\lambda$. The latter integral is equal to $\Gamma[1 + (1-\sigma')/\theta]$.

Denoting total expenditure $Y_j = \sum_{i \in N} P_{ij} Q_{ij}$, the share spent on goods from i is:

$$\frac{X_{ij}}{Y_j} = \frac{\left(P_{ij}\right)^{1-\sigma}}{\sum_{n=1}^{N} \left(P_{nj}\right)^{1-\sigma}} \tag{13}$$

where the value of bilateral trade is obtained by maximizing (3) subject to the constraint that expenditure not exceed total income. Total income is given by the landed value of exports from j to all partners $\sum_{n \in N} P_{jn} Q_{jn}$ and is equal to total expenditure in equilibrium.

The gravity structure of aggregate bilateral trade replicates Anderson and van Wincoop (2003) whereby the magnitude of the trade elasticity is determined by the Armington elasticity σ which captures perceived substitutability of country-specific product bundles:

$$X_{ij} = \frac{Y_i Y_j}{Y_w} \left(\frac{\tau_{ij}}{\Pi_i \Phi_j}\right)^{1-\sigma} \tag{14}$$

where Y_w is world expenditure, $\Phi_j = \left[\sum_{n=1}^{N} \left(P_{nj}\right)^{1-\sigma}\right]^{1/(1-\sigma)}$ is the overall price index of the importer, $\Pi_i = \left[\sum_j s_j (\tau_{ij}/\Phi_j)^{1-\sigma}\right]^{1/(1-\sigma)}$ is the multilateral trade resistance term of the exporter, and $s_j = Y_j/Y_w$ is the expenditure share of each country.[11]

2.2 The incidence of the truncated product set

In theory, only the intensive margin is operational whereby higher production or trade costs leave the set of traded goods unaffected. In practice, product coverage of the world market is highly fragmented (App.A). This section

[11] Anderson and van Wincoop (2003): use (13), sum over i's partners to get income: $Y_i = \sum_j X_{ij} = \sum_j (T_i^{-1/\theta} c_i \tau_{ij})^{1-\sigma} \Phi_j^{\sigma-1} Y_j$. Solve for $\left(T_i^{-1/\theta} c_i\right)^{1-\sigma} = Y_i \left[\sum_j (\tau_{ij}/\Phi_j)^{1-\sigma} Y_j\right]^{-1}$, plug this back into (13) to get $X_{ij} = Y_i Y_j \left(\frac{\tau_{ij}}{\Phi_j}\right)^{1-\sigma} \left[\sum_j (\tau_{ij}/\Phi_j)^{1-\sigma} Y_j\right]^{-1}$. Multiply and divide the RHS by Y_w and replace Π_i by its value.

shows that zeros can be accomodated as a statistical feature of the data instead of modifying the production side of the economy to generate structural zeros as in Helpman et al. (2008) or Eaton et al. (2012).

2.2.1 The incidence of statistical zeros on the price of the bundle

Assume there exists a statistical threshold $\bar{X}$ common to all countries such that the nominal value of sectoral bilateral trade is registered iff it is at least equal to this threshold. Sectors in which the least cost draw is sufficiently high carry marginal weight in expenditure on the exporter-specific composite good ($\sigma' > 1$). Define $\bar{w}$ the maximal production cost associated with the smallest observed nominal value and apply (5) to the price of each sectoral good: $X_i(k) \geq \bar{X}$ implies $P_i(k) \leq \bar{w}$. The fraction of high cost draws determines observed bundle variety on the world market. Destination-specific characteristics and bilateral trade frictions determine bilateral variation in bundle variety.

The cost threshold $\bar{w}$ is incorporated in the lower-tier price index to obtain the landed price of the truncated product bundle $\bar{P}_{ij}$:

$$\bar{P}_{ij}(p) \quad = \quad \bar{P}_{ij}(w) = \left\{ \int_0^{\bar{w}} w^{1-\sigma'} f(w) \, \mathrm{d}w \right\}^{1/(1-\sigma')} \tag{15}$$

To derive the lower-tier price index for the truncated product set, I follow Eaton and Kortum (2010) and rewrite (15) as the product of two terms: the expected number of bilateral draws below the threshold and the expected cost of such draws.

The number of techniques that allow production with cost less than $\bar{w}$ is given by $\lambda_i(\bar{w})$ (sec.2.1). Augmenting unit labor cost c_i with bilateral trade frictions τ_{ij} redefines this statistic at the bilateral level: $\lambda_{ij}(\bar{w}) = T_i(c_i\tau_{ij}/\bar{w})^{-\theta}$. To simplify notation, denote $\tilde{z}_i$ the scale parameter of the Fréchet distribution and use the definition of the mean $T_i^{1/\theta}\Gamma(1-1/\theta)$ to define $\tilde{z}_i = T_i^{1/\theta}$. The scale parameter of the Fréchet is a sufficient statistic of fundamental exporter productivity (Costinot et al. (2012)). The expected bilateral number of draws is:

$$\lambda_{ij}(\bar{w}) \quad = \quad \left(c_i \tau_{ij}/\tilde{z}_i\right)^{-\theta} \bar{w}^{\theta} \tag{16}$$

The expected cost of such draws is obtained by integrating the conditional density function over effectively observed cost draws. Recall that in Eaton and Kortum (2010) techniques are drawn from a Pareto distribution with parameter θ. Hence, the distribution of costs conditional on the cost threshold $\bar{w}$ is $F_c(w) = \Pr(W \leq w | W \leq \bar{w}) = (w/\bar{w})^{\theta}$. The corresponding conditional density function is $f_c(w) = \theta w^{\theta-1} \bar{w}^{-\theta}$. The lower-tier price index is given by:

$$\bar{P}_{ij}(w) = \left\{ \underbrace{\lambda_{ij}(\bar{w})}_{\text{nbr draws}} \underbrace{\int_0^{\bar{w}} w^{1-\sigma'} f_c(w) \, \mathrm{d}w}_{\text{expected cost}} \right\}^{1/(1-\sigma')} \tag{17}$$

Replacing $\lambda_{ij}(\bar{w})$ and $f_c(w)$ and solving for the integral defines the price of the truncated product bundle as a function of exporter characteristics, trade frictions, and the cost threshold:

$$\bar{P}_{ij} \quad = \quad \left\{ \frac{\theta}{\theta - \sigma' + 1} \left(\frac{c_i \tau_{ij}}{\tilde{z}_i} \right)^{-\theta} \bar{w}^{\theta - \sigma' + 1} \right\}^{1/(1-\sigma')} \tag{18}$$

The next step is to derive the cost threshold. Using the two-tier structure of expenditure allocation, the landed value of sectoral trade that is effectively observed at the bilateral level is:

$$X_{ij}(k)\big|_{X_{ij}(k) \geq \bar{X}} \quad = \quad \left[\frac{P_{ij}(k)}{P_{ij}} \right]^{1-\sigma'} \left[\frac{P_{ij}}{\Phi_j} \right]^{1-\sigma} Y_j \tag{19}$$

The expression of the cost threshold is obtained by solving for the upper bound of the observed landed sectoral price in (19):

$$P_{ij}(k) \;\leq\; \left[\frac{Y_j}{\bar{X}}\,\Phi_j^{\sigma-1}\right]^{1/(\sigma'-1)} P_{ij}^{\;\frac{\sigma'-\sigma}{\sigma'-1}} \;=\; \bar{w} \tag{20}$$

Plug (12) into (20) to visualize the four components of the cost threshold:

$$\bar{w} \;=\; \left[\kappa^{(\sigma-\sigma')}\bar{X}\right]^{\frac{1}{(1-\sigma')}} \left[Y_j\Phi_j^{\sigma-1}\right]^{\frac{1}{(\sigma'-1)}} \left[\frac{c_i}{\tilde{z}_i}\right]^{\frac{\sigma'-\sigma}{\sigma'-1}} \tau_{ij}^{\frac{\sigma'-\sigma}{\sigma'-1}} \tag{21}$$

Define $\upsilon = \frac{\theta}{\theta-\sigma'+1}\left[\kappa^{(\sigma-\sigma')}\bar{X}\right]^{(\theta-\sigma'+1)/(1-\sigma')}$ and $\alpha = (\sigma'-\sigma)(\theta-\sigma'+1)/(\sigma'-1)$, with $0 \leq \alpha < \theta$.[12] Plugging (21) into (18) gives the landed price of the truncated bundle:

$$\bar{P}_{ij} \;=\; \left\{\upsilon\,\left[Y_j\Phi_j^{\sigma-1}\right]^{\frac{\theta-\sigma'+1}{\sigma'-1}} \left(\frac{c_i\tau_{ij}}{\tilde{z}_i}\right)^{-(\theta-\alpha)}\right\}^{1/(1-\sigma')} \tag{22}$$

Direct comparison of the exponent in (22) and (12) gives $(\theta-\alpha)/(\sigma'-1) > 1$. Hence, truncation unambiguously enhances the sensitivity of bundle prices to bilateral trade frictions.

2.2.2 Substitutability of truncated product bundles

With truncation, the object of interest becomes the degree of substitutability of effectively traded product bundles ($\bar{\sigma}$). To characterize the wedge that truncation introduces between structural and measured bundle substitutability, it is helpful to work out effective expenditure allocation among truncated product bundles.

Truncated expenditure allocation at the upper-tier is obtained by conditioning utility $\bar{U}_j$ to be derived from registered quantities $\bar{Q}_{ij}$ according to the truncated analog of (3). Total expenditure $\bar{Y}_j$ is set equal the sum of registered

[12] Parameter restrictions $\theta+1 > \sigma' \geq \sigma > 1$ entail $\theta > \alpha \geq 0$.

bilateral imports: $\bar{Y}_j = \sum_i \bar{X}_{ij}$ where $\bar{X}_{ij} = \sum_k X_{ij}(k)\left\{X_{ij}(k) : X_{ij}(k) \geq \bar{X}\right\}$.[13] A vector of trade deficits D_j equalizes truncated expenditure to truncated income: $\bar{Y}_j = \sum_n \bar{P}_{jn}\bar{Q}_{jn} + D_j$. This gives (13) in terms of observed expenditure:

$$\frac{\bar{X}_{ij}}{\bar{Y}_j} = \frac{\left(\bar{P}_{ij}\right)^{1-\bar{\sigma}}}{\sum_{n=1}^{N}\left(\bar{P}_{nj}\right)^{1-\bar{\sigma}}} \tag{23}$$

To show that $\bar{\sigma} < \sigma'$, consider relative truncated expenditure on the world market for some pair $\{i, i'\}$ such that $\tilde{z}_i/c_i > \tilde{z}_{i'}/c_{i'}$:

$$\frac{\bar{X}_i}{\bar{X}_{i'}} = \frac{X_i P_i^{\sigma'-1} \int_0^{\bar{w}_i} p^{1-\sigma'} f_i(p)\mathrm{dp}}{X_{i'} P_{i'}^{\sigma'-1} \int_0^{\bar{w}_{i'}} p^{1-\sigma'} f_{i'}(p)\mathrm{dp}} \tag{24}$$

Consider the numerator on the right hand side (RHS) of (24). The last component is equal to $\bar{P}_i^{1-\sigma'}$ and is a monotonic transformation of the truncated price index. The second component is $P_i^{\sigma'-1} = P_i^{\sigma-1}P_i^{\sigma'-\sigma}$. Since $[X_i/X_{i'}]/(P_i/P_{i'})^{1-\sigma} = 1$, the expression simplifies to:

$$\frac{\bar{X}_i}{\bar{X}_{i'}} = \left[\frac{\bar{P}_i}{\bar{P}_{i'}}\right]^{1-\sigma'}\left[\frac{P_i}{P_{i'}}\right]^{\sigma'-\sigma} < \left[\frac{\bar{P}_i}{\bar{P}_{i'}}\right]^{1-\sigma'} \tag{25}$$

where the inequality is established by $[P_i/P_{i'}] < 1$ given $\tilde{z}_i/c_i > \tilde{z}_{i'}/c_{i'}$. Hence, truncated bundles are perceived to be less substitutable than sectoral goods that compose the bundle.

To show that $\bar{\sigma} > \sigma$, use $\bar{P}_i^{1-\sigma'} = \bar{P}_i^{1-\sigma}\bar{P}_i^{\sigma-\sigma'}$ to write:

$$\frac{\bar{X}_i}{\bar{X}_{i'}} = \left[\frac{\bar{P}_i}{\bar{P}_{i'}}\right]^{1-\sigma}\left[\frac{\bar{P}_i}{\bar{P}_{i'}}\right]^{\sigma-\sigma'}\left[\frac{P_i}{P_{i'}}\right]^{\sigma'-\sigma} = \left[\frac{\bar{P}_i}{\bar{P}_{i'}}\right]^{1-\sigma}\left[\frac{\bar{P}_{i'}/P_{i'}}{\bar{P}_i/P_i}\right]^{\sigma'-\sigma} \tag{26}$$

[13] This simply says that the solution to the non-truncated problem directly gives expenditure allocation in the truncated problem by conditioning on some threshold $\bar{X}$.

Focus on the last term on the RHS of (26). The ratio of truncated to non-truncated prices is always greater than one.[14] Moreover, this ratio is monotonically decreasing in adjusted exporter ability $\tilde{z}_i/c_i$.[15] As $\tilde{z}_i/c_i > \tilde{z}_{i'}/c_{i'}$, it must be that $[\bar{P}_{i'}/P_{i'}]\,/\,[\bar{P}_i/P_i] > 1$ whereby $\bar{\sigma} > \sigma$. Hence, truncated bundles are perceived to be more substitutable than non-truncated product bundles:

$$\frac{\bar{X}_i}{\bar{X}_{i'}} = \left[\frac{\bar{P}_i}{\bar{P}_{i'}}\right]^{1-\sigma} \left[\frac{\bar{P}_{i'}/P_{i'}}{\bar{P}_i/P_i}\right]^{\sigma'-\sigma} > \left[\frac{\bar{P}_i}{\bar{P}_{i'}}\right]^{1-\sigma} \tag{27}$$

The main implication of truncation is that the price sensitivity of demand becomes specific to the exporter pair and is increasing in the pair-specific ability gap.[16] The corollary is that truncation has to be sufficiently severe to entail a sensible magnification of the aggregate price. In particular, the relative magnification factor $[\bar{P}_{i'}/P_{i'}]\,/\,[\bar{P}_i/P_i]$ approaches 1 for all but very small exporters whenever σ' is sufficiently high.[17] This finding has an immediate implication for the choice of the estimator used to identify the magnitude of upper-tier substitutability in sec.3. The upward bias in the estimated parameter relatively to the structural parameter is expected to be reduced if the estimator places relatively little weight on small trade volumes. This motivates the choice of the Poisson Pseudo Maximum Likelihood (PPML) estimator (Head and Mayer (2013), Santos Silva and Tenreyro (2006)).[18]

By determining expenditure allocation across and within product bundles, the interplay of upper- and lower-tier substitutability may generate statistical zeros in aggregate bilateral trade. The intuition is the following. If σ is high,

[14] This is established by taking the partial derivative of the truncated price index: $\partial \bar{P}_i/\partial \bar{w} < 0$ and observing that $\bar{P}_i \to P$ when $\bar{w} \to \infty$.

[15] Focus on the bundle exported to the world market and use (12) and (22) to establish $\bar{P}_i/P_i = \varpi(\tilde{z}_i/c_i)^{-\rho}$ with $\rho = (\theta - \sigma' + 1)(\sigma - 1)/(1-\sigma')^2$ and ϖ constant across exporters.

[16] Waugh (2010) finds that asymmetric trade frictions are needed to rationalize relative expenditure on exports from developed and developing countries. This paper suggests a complementary mechanism through the ability gap that determines relative truncation.

[17] Simulation is conducted for 100 countries with 10000 products each, with the max/min ability ratio set at 100, and the parameter range defined as $2 \leq \sigma < 4$ and $4 < \sigma' \leq 8$.

[18] Sec.3 finds relatively high lower-tier substitutability in the data ($\sigma' \geq 8$). Consequently, the incidence of truncation on relative bundle prices is expected to be reduced whereby $\bar{\sigma} \approx \sigma$.

a relatively small share of expenditure is allocated to low-ability exporters. If σ' is low, the best draws of low ability exporters get a relatively low share of total expenditure on the bundle. As a consequence, the trade flow for the least cost good of the low-ability exporter may be below the registration threshold. This is all the more likely if the receiving country is itself a low-ability exporter. It follows that low-ability exporters are more likely to export positive amounts to high-ability importers while high-ability exporters are more likely to export positive amounts to low-ability importers. Aggregate zeros are all the more likely if world ability is low and ability dispersion across exporters is high.

2.2.3 The incidence of truncation on the trade elasticity

Truncation modifies the elasticity of trade flows to variable trade costs. Plugging (22) in the numerator and denominator of truncated upper-tier demand (23) leaves fundamental exporter characteristics and bilateral trade frictions raised to the power $\varepsilon = -(\theta-\alpha)\bar{\gamma}$ where $0 < \bar{\gamma} = (\bar{\sigma}-1)/(\sigma'-1) < 1$:

$$\frac{\bar{X}_{ij}}{\bar{Y}_j} = \frac{\left(c_i\tau_{ij}/\tilde{z}_i\right)^{-(\theta-\alpha)\bar{\gamma}}}{\sum_{n=1}^{N}\left(c_n\tau_{nj}/\tilde{z}_n\right)^{-(\theta-\alpha)\bar{\gamma}}} \tag{28}$$

The expression of the trade elasticity is simplified by plugging (12) and (22) into (25):

$$\frac{\bar{X}_{ij}}{\bar{X}_{i'j}} = \left[\frac{c_i\tau_{ij}/\tilde{z}_i}{c_{i'}\tau_{i'j}/\tilde{z}_{i'}}\right]^{-(\theta-\alpha)}\left[\frac{c_i\tau_{ij}/\tilde{z}_i}{c_{i'}\tau_{i'j}/\tilde{z}_{i'}}\right]^{\sigma'-\sigma} \tag{29}$$

Defining $\gamma = (\sigma-1)/(\sigma'-1) < \bar{\gamma}$ and rearranging to simplify the exponent gives:

$$\frac{\bar{X}_{ij}}{\bar{X}_{i'j}} = \left[\frac{c_i\tau_{ij}/\tilde{z}_i}{c_{i'}\tau_{i'j}/\tilde{z}_{i'}}\right]^{-\theta\gamma} \tag{30}$$

The magnitude of the trade elasticity $|\varepsilon| = \theta\gamma$ is magnified relatively to the

Armington model without truncation ($\theta > \sigma' - 1$) but dampened relatively to the Ricardian model with coinciding upper- and lower-tier substitutability ($\gamma < 1$). Its bounds are determined by the bounds of the Armington elasticity: $|\varepsilon|$ tends to $(\sigma - 1)$ when $\sigma \to 1$ and to θ when $\sigma \to \sigma'$.

The impact of an increasing Armington elasticity on the sensitivity of trade to variable trade costs is magnified relatively to the benchmark Armington model. The magnification factor is increasing in the extent to which producer heterogeneity exceeds within-bundle substitutability. This happens because θ regulates variation in the number of observed draws as a consequence of a change in trade costs while σ' regulates the incidence of these marginal draws on the price of the truncated composite good (Chaney (2008)). Whenever σ' is relatively high ($\theta/(\sigma' - 1) \to 1$), marginal draws have little incidence on the price of the truncated bundle. This dampens the incidence of the extensive margin on the trade elasticity.[19]

The gap between measured and structural substitutability can be circumscribed using the two expressions of the trade elasticity. Rearranging and simplifying $\theta\gamma = (\theta - \alpha)\bar{\gamma}$ gives:

$$\frac{\bar{\sigma}}{\sigma} = \frac{\theta(\sigma' - 1) - (\sigma' - \sigma)(\theta - \sigma' + 1)/\sigma}{\theta(\sigma' - 1) - (\sigma' - \sigma)(\theta - \sigma' + 1)} \tag{31}$$

The ratio tends to 1 whenever $(\sigma' - \sigma)(\theta - \sigma' + 1) \to 0$. For the magnification factor to be significantly different from 1, it must be that $\sigma \ll \sigma' \ll \theta$. But in this case $\bar{\sigma}/\sigma \ll \sigma'/\sigma$. Hence, $\bar{\sigma}$ is always a better approximation of σ than of σ'.[20]

[19] The incidence of parameter changes on the magnitude of $|\varepsilon|$ is qualitatively similar to Feenstra et al. (2014) and Costinot and Rodriguez-Clare (2014) although the expression of the trade elasticity is model-specific.

[20] Empirical evidence on the magnitude of lower-tier substitutability ($\sigma' > 8$) indicates that $\bar{\sigma} \approx \sigma$ in our data. This is because the incidence of truncation on relative price distortion is reduced whenever σ' is relatively high.

3 Estimation of demand parameters in cross-section

3.1 Lower-tier substitutability

3.1.1 Methodology

The lower-tier elasticity σ' is needed to compute the price of the truncated bundle (App.C). According to the model, variation in sectoral expenditure within the bundle is determined by the variability of productivity draws: $X_{ij}(k)/X_{ij}(k') = [z_i(k)/z_i(k')]^{\sigma'-1}$. I use the structure of the model to estimate sectoral productivity $z_i(k)$, and then use the fact that these draws are inversely proportional to sectoral prices to obtain annual estimates of σ'.

Exporter-sector productivity draws $z_i(k)$ are identified by focusing on variation in sectoral expenditure within the bundle across the full set of destination markets, i.e. by estimating the set of exporter-sector fixed effects $f_i(k)$.[21] Following Costinot et al. (2012) I use a flexible specification in which pair fixed effects η_{ij} pick up bilateral trade costs while destination-sector fixed effects $\eta_j(k)$ capture systematic variation in sectoral trade costs:

$$\ln\left[X_{ij}(k)\right] \quad = \quad \eta_0 + f_i(k) + \eta_j(k) + \eta_{ij} + \eta_{ij}(k) \tag{32}$$

Exporter-sector fixed effects $f_i(k) = (\sigma'-1)\ln(z_i(k))$ identify sectoral productivity up to the scalar $(\sigma'-1)$ relatively to a benchmark country and sector.

Sectoral bilateral prices $P_{ij}(k)$ are regressed on estimated exporter-sector fixed effects $\hat{f}_i(k)$ while controlling for pair-specific determinants of trade with pair fixed effects β_{ij}:

$$\ln\left[P_{ij}(k)\right] \quad = \quad \beta_0 - \zeta \hat{f}_i(k) + \beta_{ij} + \beta_{ij}(k) \tag{33}$$

[21] The approach is similar in spirit to Hummels and Schaur (2013) who use exporter-specific sales to the world as a predictor of latent product profitability on the US market.

Within-bundle variation in sectoral prices is determined by variation in sectoral productivity whereby $-\zeta \hat{f}_i(k) = -\ln(\hat{z}_i(k))$. The magnitude of lower-tier substitutability is computed as $\mathrm{E}(\sigma'_f) = 1 + 1/\mathrm{E}(\zeta)$ where the subindex f indicates that the estimate is obtained with the first approach.[22] This estimate is an approximation since by Jensen's inequality $E(\sigma'_f - 1) \geq 1/E(\zeta)$. The quality of the approximation is checked by taking the Taylor expansion about the expectation and evaluating the magnitude of higher order terms.[23]

The second approach delivers an estimate of lower-tier substitutability by regressing bilateral sectoral expenditure on the price component predicted by sectoral productivity draws $\ln\left[\hat{P}_i(k)\right] = -\zeta \hat{f}_i(k)$ while controlling for pair fixed effects $\tilde{\eta}_{ij}$:

$$\ln\left[X_{ij}(k)\right] \quad = \quad \tilde{\eta}_0 - (\sigma'_s - 1)\ln \hat{P}_i(k) + \tilde{\eta}_{ij} + \tilde{\eta}_{ij}(k) \tag{34}$$

where the subindex s indicates that this estimate is obtained with the second approach. The implicit assumption in the latter estimation is that $\zeta \hat{f}_i(k) = \ln(\hat{z}_i(k))$ whereby $\sigma'_f = \sigma'_s$.

Estimation of sectoral productivity draws $\hat{f}_i(k)$ in (32) may encounter feasibility constraints because of the sheer number of fixed effects. The data can be demeaned to reduce dimensionality. I opt for an alternative strategy whereby the relationship in (32) is estimated separately for each exporter while normalizing sectoral productivity by the best exporter-specific draw.[24] In the model, better draws have higher probability of being exported to any market. The best draw

[22] The reciprocal transformation applied to ζ entails that $\mathrm{E}(\sigma'_f)$ is well-defined iff ζ exhibits negligible probability in the finite neighbourhood of 0 (Johnson et al. (1994)). If this is the case, the standard error of the transformed parameter can be computed using the delta method whereby $\mathrm{Var}(\sigma'_f) = \mathrm{Var}(1/\zeta) = \mathrm{Var}(\zeta)/(\mathrm{E}(\zeta))^4$.

[23] Define $\zeta = \hat{\zeta} + \eta$ with $\mathrm{E}(\eta) = 0$ and show that the Taylor expansion is a convergent sequence whenever $|\eta| < \hat{\zeta}$ (this follows from footnote 22) whereby $E(1/\zeta) = 1/E(\zeta) + \mathscr{O}\left[\mathrm{Var}(\zeta)/(\mathrm{E}(\zeta))^3\right]$ when $\eta/\mathrm{E}(\zeta) \to 0$.

[24] This normalization is consistent with the characterization of exporter-specific productivity dispersion in sec.4.

is therefore identified with the most frequently exported product.[25]

Conducting the estimation separately for each exporter may reduce precision of retrieved exporter-sector dummies. Additional data cleaning is implemented as follows. Exporters with $< 50\%$ of significant dummies are dropped. The productivity distribution is truncated for the remaining exporters by keeping only negative and significant coefficients. This places the focus on the segment of the estimated productivity distribution that verifies the assumption of normalization by the best draw. The set of dummies is adjusted for precision in the estimation, and exporters who exhibit a correlation coefficient below .3 between raw and standardized dummies are dropped. The set of standardized dummies for the remaining exporters is used in the estimation of σ'.

Whenever the relationship in (32) is estimated by exporter, (33) is adjusted to include a set of controls for bilateral trade frictions (T_{ij}) provided in Mayer and Zignago (2011) together with exporter and destination fixed effects (resp. β_i and β_j) instead of pair fixed effects:

$$\ln\left[P_{ij}(k)\right] = \beta_0 - \zeta \hat{f}_i(k) + \beta_i + \beta_j + {T_{ij}}'\beta + \beta_{ij}(k) \tag{35}$$

Analogously, (34) becomes:

$$\ln\left[X_{ij}(k)\right] = \tilde{\eta}_0 - (\sigma_s' - 1)\ln \hat{P}_i(k) + \tilde{\eta}_i + \tilde{\eta}_j + {T_{ij}}'\tilde{\eta} + \tilde{\eta}_{ij}(k) \tag{36}$$

3.1.2 BACI: lower-tier substitutability in 1995-2009

The estimation is first implemented in 1995-2009 on the BACI dataset. As explained by Gaulier and Zignago (2010), BACI has the advantage of offering an extensive (212 countries) and detailed (HS 6-digit) coverage of bilateral trade while providing more complete and accurate information on unit values than the raw data supplied in UN COMTRADE.

[25] Eaton et al. (2011) show that more productive firms enter more markets. Firm productivity maps into sectoral productivity here because each good is produced by the least cost firm.

The estimation is conducted in the balanced and square samples. The balanced panel covers $> 96\%$ of total trade in 1995-2009 and contains 209 destination markets. The square panel covers 70-80% of total trade and contains 50 destination markets (App.B). The restriction of the sample to the set of stable relationships is motivated by the main identification assumption whereby expenditure variation within the bundle maps into (unobserved) sectoral productivity. This assumption is more likely to hold within the set of stable trade relationships. The drawback is that sample truncation reduces the number of markets included in the identification of sectoral productivity draws through expenditure variation within the bundle.

The second trade-off is linked to choosing the extent of data disaggregation. Each sectoral good has to be observed sufficiently frequently to credibly estimate the extent of variability in sectoral expenditure within the bundle. However, aggregation may blur the difference between lower and upper-tier substitutability. The incidence of aggregation is evaluated by conducting the estimation at the 4-digit (1222 goods) and 2-digit (93 goods) levels.[26]

Fig.1 reports annual magnitudes of lower-tier substitutability σ' obtained in the balanced sample for the 1222-good bundle. Fig.2 reports these magnitudes for the 93-good bundle. The left pane of each figure reports in black the magnitude of σ'_f for the approximation $E(\sigma'_f) = 1/E(\zeta) + 1$ and in red the central value adjusted for the maximum approximation error. The right pane reports lower-tier substitutability obtained with the second approach (σ'_s).

In all specifications the elasticity is stable in 1995-2009.[27] Two features suggest that the parameter is identified. Its magnitude is increasing in the extent of data disaggregation: it doubles from about 5 for a 93-good bundle to about 10 for a 1222-good bundle. Moreover, at a given level of data disaggregation,

[26] Exporters observed on < 10 markets or with $< 40(10)$ goods at 4(2)-digit are dropped. Price aggregation from the product to the sectoral level uses bundle-specific expenditure weights.

[27] Fixed effects are estimated separately for each exporter. About 25% of estimated fixed effects are dropped due to lack of precision.

Figure 1: BACI balanced panel 4-digit

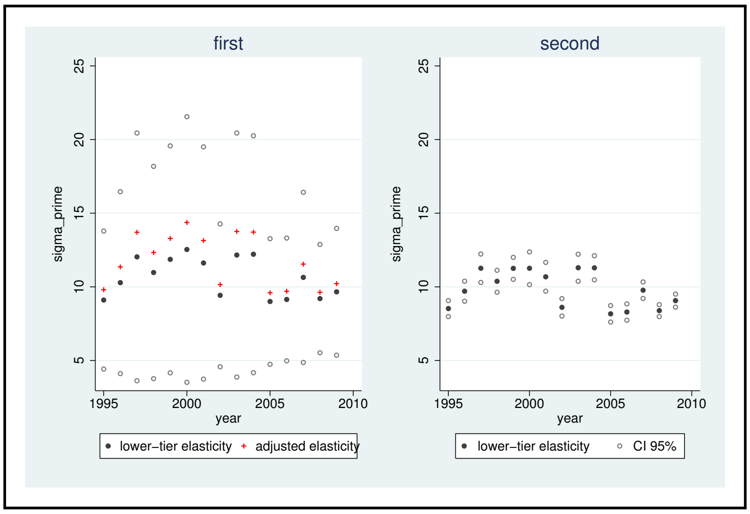
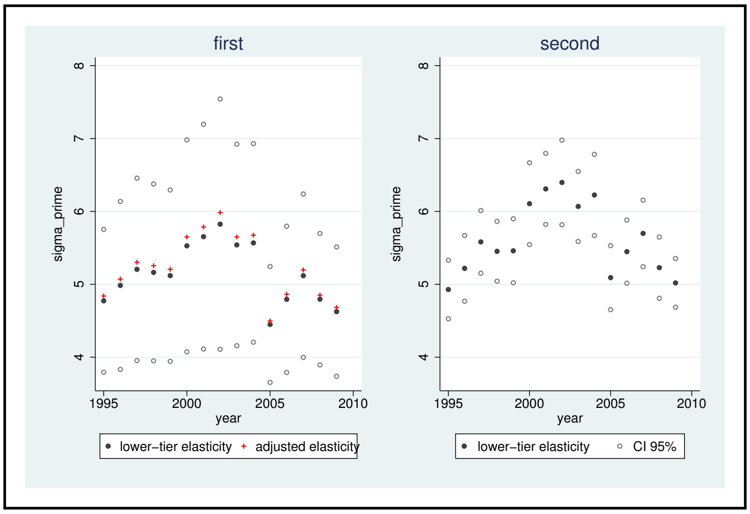


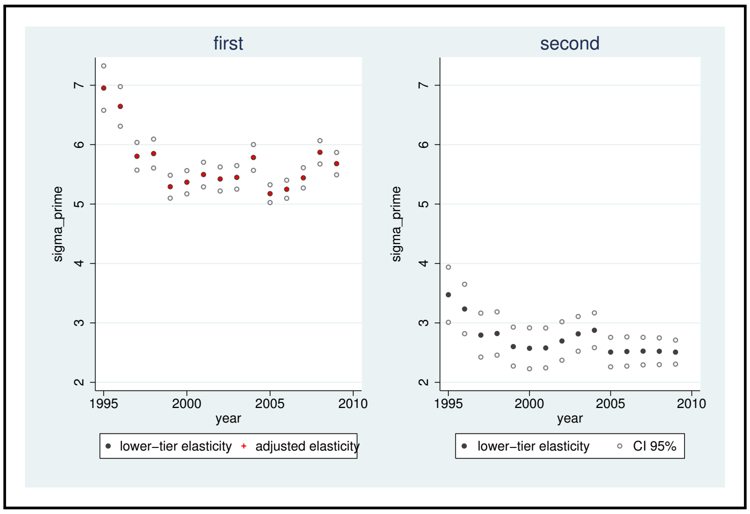
Figure 2: BACI balanced panel 2-digit



the central value of the estimate is of similar magnitude in the first and in the second approach whereby the assumption $\sigma'_f = \sigma'_s$ is verified.

Fig.3 reports estimates of lower-tier substitutability for the 93-good bundle in the square sample. This sample contains the set of 50 countries that trade positive amounts with every other country in the set in each year (App.B).[28] The error of the approximation is nil in the first approach, and the central value is estimated at $\sigma'_f \in \{5,7\}$. The same range was obtained in the balanced sample at this level of disaggregation. However, the magnitude obtained with the second approach is now significantly lower $\sigma'_s \in \{2,4\}$. This finding suggests that the estimate retrieved with the first approach is more robust to sample truncation.

Figure 3: BACI square panel 2-digit



The complementary finding is that estimation performance does not hinge on the extent of data disaggregation as much as on the number of destination markets included in the regression. In particular, sectoral dummies are weakly correlated with sectoral prices within the 1222-good bundle when the estimation is conducted separately for each exporter in the square sample (50 markets). This correlation becomes very strong if the estimation is conducted in the balanced sample (200 markets). It follows that price and expenditure information provided in BACI suffices to identify exporter-specific sectoral productivity at the 4-digit level of the HS classification. This allows obtaining annual estimates of aggregate productivity (see App.D) and of productivity dispersion (see

[28] Fixed effects are estimated simultaneously for 49 exporters relatively to the USA.

sec.4.1) for exporters of the BACI balanced sample in 1995-2009.

3.1.3 UN COMTRADE: lower-tier substitutability in 1963-2009

To move back in time I work with trade reported at the SITC 4-digit level in UN COMTRADE (UNC). Unit values are arguably a worse proxy of underlying prices in this dataset. Identification of sectoral productivity from expenditure variation within the bundle is also trickier. Estimated productivity is only weakly correlated with unit values at any level of disaggregation if the estimation is conducted separately for each exporter. Hence, the sample is cleaned to eliminate exporters with intermittent coverage of the world market, and the estimation is carried out on pooled data.[29]

In the stable sample, defined as the set of 132 exporters who are active in 10 or more markets in each year, estimation is conducted at the 2-digit level (55 goods).[30] To reduce measurement error and make identification feasible at a higher level of data disaggregation, I further restrict the sample to the square that covers 40-60% of total trade and contains 24 exporters who trade positive amounts with every other country in the set (App.B). Sectoral productivity is estimated at the 3-digit level (175 goods) in the square sample.

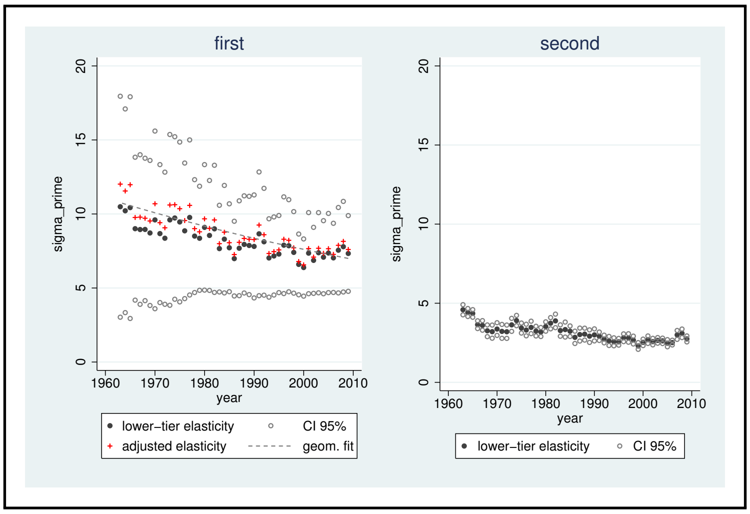
Fig.4 reports estimates of lower-tier substitutability for the 175-good bundle. The central value obtained with the first approach is $\sigma'_f \in \{7, 11\}$. The estimate obtained with the second approach is significantly lower $\sigma'_s \in \{3, 5\}$. The degree of data disaggregation is situated in-between the 2- and 4-digit levels of the HS classification. The magnitude of the estimated parameter is also intermediate to the range obtained on 2- and 4-digit data in BACI. This finding conforms to our prior that substitutability is increasing in the degree of data disaggregation.

Lower-tier substitutability has decreased by 30 (36)% if the parameter is estimated with the first (second) approach. This result is consistent with BACI

[29] Sectoral productivity is only identified relatively to a benchmark country in this dataset. Hence, measures of productivity dispersion are obtained in relative terms (sec.4.1).

[30] This sample contains all destination markets and covers 93-98% of total trade (App.B).

Figure 4: UN COMTRADE square panel 3-digit



estimates because the reduction occurs between 1963 and 1995.[31]

Figure 5: UN COMTRADE stable sample 2-digit

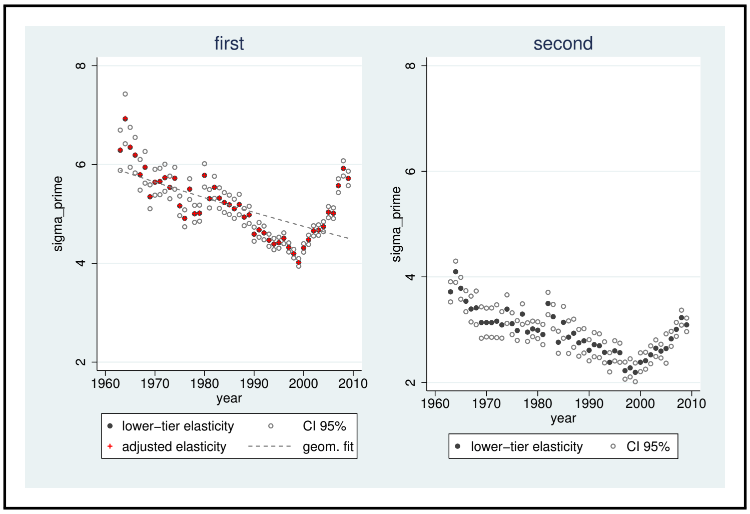


Fig.5 reports estimates for the 55-good bundle. The magnitude of $\sigma_f' \in \{4,7\}$ is lower than for the 175-good bundle. Contrary to previous results, the

[31] The annualized growth rate in 1963-1995 is $-.93\%$ per year with the first approach (-26% total change) and -1.2% per year with the second approach (-32% total change).

evolution of the parameter is U-shaped. This discrepancy may be due to the difficulty of disentangling the evolution of lower- and upper-tier elasticities at this level of data aggregation.

To sum up, lower-tier substitutability has decreased by about 30% over 1963-1995. The parameter is best described as stable in 1995-2009.

3.2 Upper-tier substitutability

3.2.1 The price of the truncated composite good

The price of the truncated product bundle for each bilateral relationship is computed using estimates of lower-tier substitutability obtained with the first approach in the BACI balanced sample at the 4-digit level and in the UNC square sample at the 3-digit level. I work with lower-tier substitutability estimated at the highest level of disaggregation to ensure separate identification of upper-tier substitutability. The choice of σ'_f in UNC is motivated by the fact that the first approach is more robust to sample truncation. The choice of σ'_f in BACI has no incidence on aggregate prices since $\sigma'_f \approx \sigma'_s$. Price aggregation is restricted to the set of bilateral relationships included in the estimation of σ'_f.[32] By direct implementation of the CES formula in (15), the price of the bundle $\bar{P}_{ij}$ is obtained by raising each observed sectoral price to the power $(1-\sigma'_f)$ and raising the sum of these components to the power $1/(1-\sigma'_f)$.

3.2.2 The instruments

A non-instrumented estimation of the price elasticity in the truncated demand equation (23) may run into the classical endogeneity concern whereby unobserved quality pushes up observed prices and observed expenditure and introduces a downward bias in the estimate of the demand elasticity (Feenstra and Romalis (2014); Crozet and Erkel-Rousse (2004)). Unobserved quality corre-

[32] This eliminates bundles with < 40 goods and exporters who cover < 10 destination markets.

sponds to unobserved ability in the model of this paper. The downward bias occurs if observed prices exceed the true underlying prices that determine expenditure allocation at the upper tier by some component of fundamental exporter ability that is unobserved.

The only bilateral component of truncated prices $\bar{P}_{ij}$ that contains information on cost-driven price variation corresponds to bilateral trade frictions (22). The premise of this paper is that information on trade frictions is unavailable.[33] The alternative is to find a variable that picks up fundamental exporter ability $\tilde{z}_i$ or cost-adjusted exporter ability $\tilde{z}_i/c_i$ and can be used to instrument aggregate prices. The three variables used in this paper are: physical capital stocks, aggregate TFP, and bundle variety.

The model renders explicit the mapping between bundle variety and cost-adjusted exporter ability.[34] The parameter of the Poisson distribution $\lambda_{ij} = \left(c_i\tau_{ij}/\tilde{z}_i\right)^{-\theta}\bar{w}^{\theta}$ gives the number of techniques available for production with cost lower than some threshold $\bar{w}$. Using (21) to solve for the cost threshold and rearranging gives:

$$\lambda_{ij} \quad = \quad \left[\bar{X}\kappa^{\sigma-\sigma'}\right]^{-\frac{\theta}{\sigma'-1}} \left[Y_j\Phi_j^{\sigma-1}\right]^{\frac{\theta}{\sigma'-1}} \tau_{ij}^{-\theta\gamma}\left[\tilde{z}_i/c_i\right]^{\theta\gamma} \tag{37}$$

Conditional on destination-specific characteristics and bilateral trade frictions, the number of goods delivered to the world market is increasing in cost-adjusted exporter ability. I assume that bilateral bundle variety is distributed Poisson and follow Gourieroux et al. (1984) in fitting a linear exponential model in each year:

$$\lambda_{ij} = \exp\left\{\chi_0 + f_i + T_{ij}{}'\chi + \chi_j\right\}\chi_{ij} \tag{38}$$

where χ_0 is a constant, T_{ij} is a vector of bilateral trade cost controls (distance,

[33] If it were, ε would be estimated directly using Caliendo and Parro (2014) methodology.
[34] Bundle variety is defined at the HS 6-digit in BACI and SITC 4-digit in UNC.

common language...), and χ_j are destination fixed effects. Cost-adjusted ability relatively to the benchmark country (USA) is captured by exporter fixed effects $\hat{f}_i = \theta\gamma \ln(\tilde{z}_i/c_i)$.

I find that exporter ability captured through bundle variety picks up the same type of variation as information on stocks of physical capital provided in the Penn World Tables (Feenstra et al. (2013)).[35] Moreover, capital stocks 'trump' bundle variety in that the latter has no additional power in predicting prices of exporter-specific bundles when both variables are used in the estimation. Consequently, these variables are used separately to isolate the price component that covaries with exporter ability. The third specification combines estimates of aggregate TFP with information on stocks of physical capital to instrument bundle prices (App.D).

Denote the instrument $\delta_i = \left\{\hat{f}_i, \ln(K_i)\right\}$ where $\hat{f}_i$ is the standardized coefficient of the exporter fixed effect estimated in (38), and K_i is the stock of physical capital. The ability component of bilateral prices is identified by estimating (39) in each year, with standard errors clustered by exporter to take into account the use of a repeated regressor:

$$\ln\left(\bar{P}_{ij}\right) = \mu_0 - \mu_1\delta_i + T_{ij}{}'\mu + \mu_j + \mu_{ij} \tag{39}$$

where μ_0 is a constant, and μ_j is the destination fixed effect.

Fig.6 reports annual estimates of the coefficient $-\mu_1$ obtained for each instrument in BACI. Fig.7 reports these results for UNC. The coefficient is always significant and negatively signed. This conforms to the prediction of the model that cost-adjusted ability reduces the price of the product bundle (see (22)).

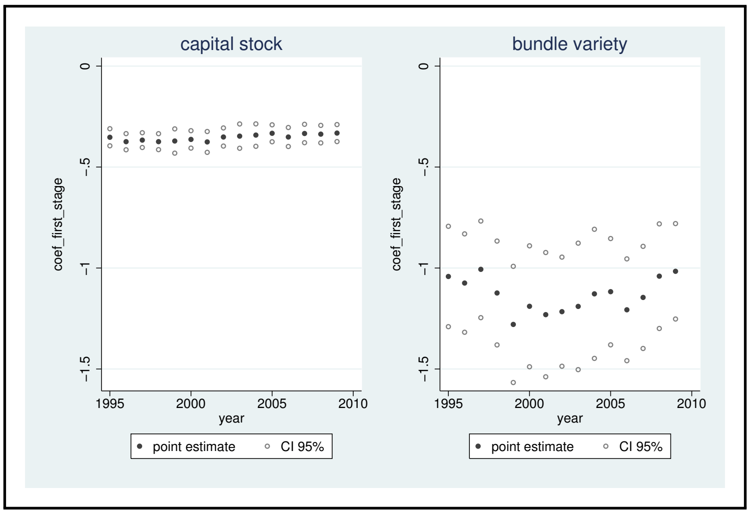
The right pane of each figure documents the relationship between aggregate prices and the variety of the product mix delivered to the world market. The magnitude of the coefficient is stable in BACI while it doubles in UNC. In

[35] The correlation coefficient exceeds .7. PWT 8.0 is available at http:\www.ggdc.net/pwt.

App.A it is shown that dispersion in bundle variety is largely maintained in the BACI balanced sample while it shrinks to nil in the UNC square sample by the end of the 1980s. As the instrument may be picking up spurious price variation, it is not used to estimate upper-tier substitutability in UNC.

The left pane of each figure documents the relationship between aggregate prices and physical capital stocks. Results are quantitatively similar in BACI and in UNC although the precision of the estimation in UNC is gradually reduced. This is due to reduced variation in capital stocks among the countries included in the square sample. The sensitivity of results to this shortcoming is checked in App.D by using estimates of aggregate TFP together with information on physical capital stocks to instrument bundle prices.

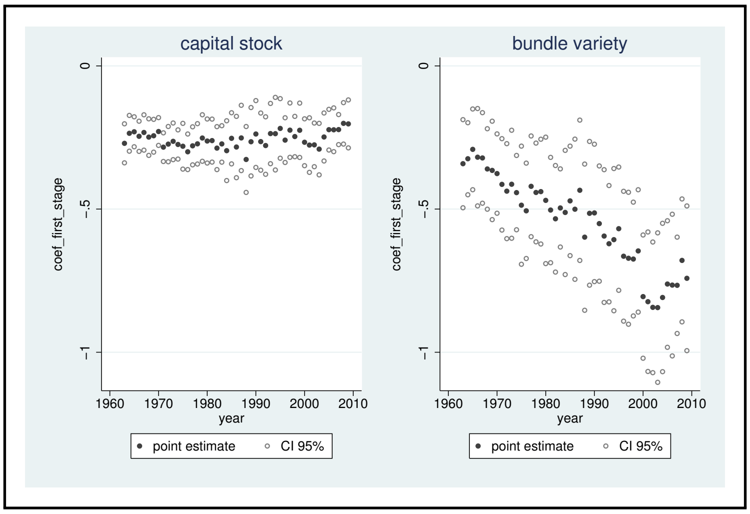
Figure 6: BACI balanced 4-digit: bundle price and underlying ability



3.2.3 Results on upper-tier substitutability

Upper-tier substitutability is identified by regressing bilateral expenditure on instrumented prices of truncated product bundles while controlling for destination fixed effects $\bar{\eta}_j$ and the vector of bilateral trade costs T_{ij}. To obtain a consistent point estimate of the parameter, the estimation is conducted in multi-

Figure 7: UNC square 3-digit: bundle price and underlying ability



plicative form by implementing the PPML estimator advocated by Santos Silva and Tenreyro (2006), and with standard errors clustered by exporter:

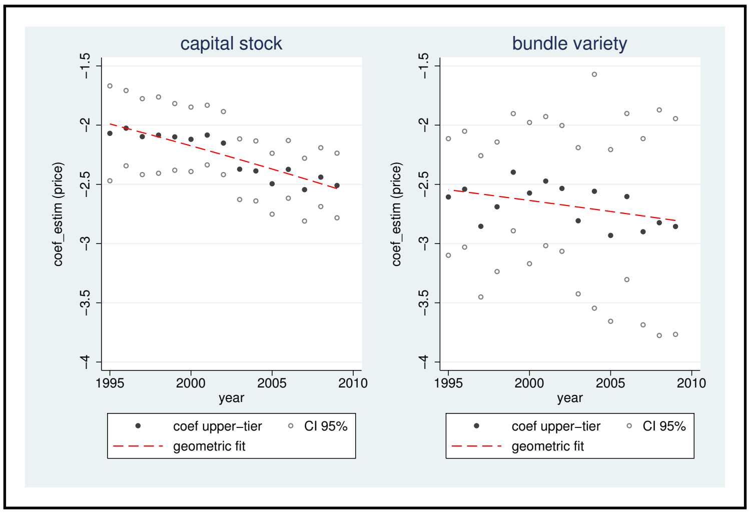
$$\bar{X}_{ij} = \exp\left\{\bar{\eta}_0 - (\bar{\sigma}-1)\ln\hat{P}_i + \bar{\eta}_j + T'_{ij}\bar{\eta}\right\}\bar{\eta}_{ij} \tag{40}$$

Fig.8 reports annual estimates of the coefficient $-(\bar{\sigma}-1)$ in the BACI balanced sample. Upper-tier substitutability $\bar{\sigma}$ is situated in the 3-4 range. Similar magnitudes are obtained in Feenstra et al. (2014) and Broda and Weinstein (2006).[36] Upper-tier substitutability $\bar{\sigma}$ increases from 3.54 in 1995 to 3.81 in 2009 (+7%) when prices are instrumented with bundle variety (right pane). This increase is magnified to 18% (from 2.99 in 1995 to 3.54 in 2009) when prices are instrumented with capital stocks (left pane). The parameter increases by 12% if prices are instrumented with estimated TFP together with capital stocks (App.D).

Fig.9 reports annual estimates of the coefficient $-(\bar{\sigma}-1)$ in the UNC square sample. Upper-tier substitutability increases from 3.87 in 1963 to 4.49 in 2009

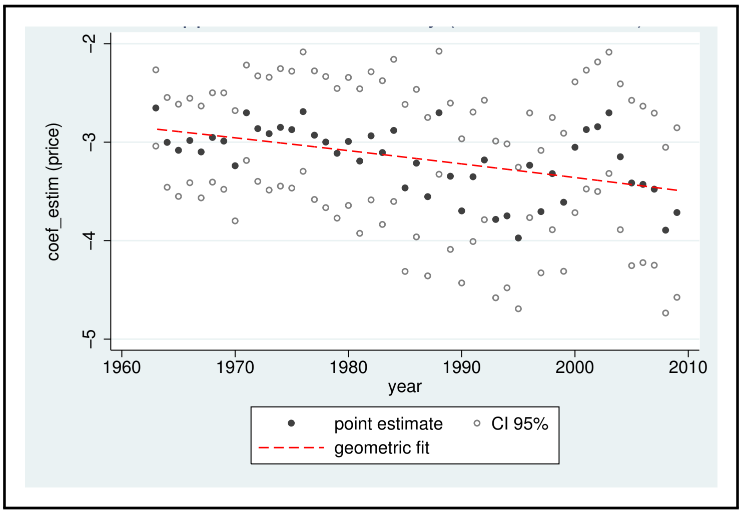
[36] Feenstra et al. (2014) obtain a median micro-elasticity for the U.S. of 3.24 in TSLS (4.12 in two-step GMM) while Broda and Weinstein (2006) obtain a median elasticity of 3.1.

Figure 8: BACI balanced 4-digit: upper-tier substitutability



(+16%) when prices are instrumented with capital stocks.[37] The parameter is stable in 1963-1995 and increases by 16.5% in 1995-2009 if prices are instrumented with aggregate TFP together with capital stocks (App.D).

Figure 9: UNC square 3-digit: upper-tier substitutability



[37] Archanskaia and Daudin (2012) report a 13% increase in 1963-2009 using a different estimation approach.

To sum up, the parameter $\bar{\sigma}$ that measures perceived bundle substitutability is comprised between 3 and 4.5. Under the assumption that product substitutability evolves similarly for truncated and non-truncated product bundles,[38] the Armington elasticity σ has increased by 16-18% between 1963 and 2009.

The combined evolution of upper- and lower-tier substitutability corresponds to an annualized growth rate of 1.16% in the magnitude of the trade elasticity $|\varepsilon|$ and amounts to a 71% increase between 1963 and 2009.[39] This partial effect measures the impact of changes in demand parameters on the sensitivity of trade to variable trade costs. It provides a lower bound on the total change in $|\varepsilon|$ if, as argued in Levchenko and Zhang (2011), dispersion in sectoral productivity draws has been reduced.

4 The trade elasticity and the cost of autarky

4.1 Empirical evidence on productivity dispersion (θ)

To the best of our knowledge, Levchenko and Zhang (2011) is the only paper that characterizes the evolution of dispersion in sectoral productivity in 1960-2010. The authors obtain estimates of sectoral technology stocks cleaned from the effect of country differences in factor abundance and sectoral differences in factor intensity for 19 industrial sectors at the ISIC 2-digit level. Levchenko and Zhang (2011) find evidence of swifter technological upgrading in the worst-performing sectors and of reduced dispersion in sectoral productivity.

I check whether this pattern is present in the data at a more disaggregate level by computing annual estimates of exporter-specific productivity dispersion in BACI at the 4-digit level (1222 goods) and in UNC at the 3-digit level (175

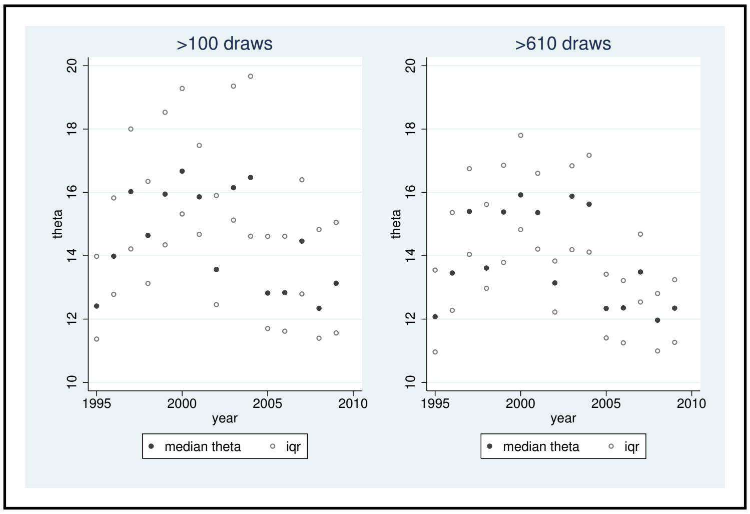
[38] This assumption is expected to hold because bundle variety is increasing in our sample, and $\bar{\sigma}$ is expected to converge to σ from above as the extent of truncation is reduced.

[39] Take the most conservative estimate: σ' decreases from 9.95 to 7.38, whereby $(\sigma' - 1)$ is reduced by 29% (-.735% per year); σ increases from 3.87 to 4.49 whereby $(\sigma - 1)$ increases by 21.6% (.425% per year).

goods). Sample-specific estimates of lower-tier substitutability σ'_f are used to extract productivity draws $z_i(k)$ from estimated exporter-sector dummies $\hat{f}_i(k) = (\sigma' - 1)\ln(z_i(k))$. Exporter-specific productivity dispersion $1/\theta$ corresponds to the standard deviation of observed sectoral draws $\tilde{SD}_{it}(z)$. [40]

Measured dispersion is mechanically decreasing in the extent of truncation (Costinot et al. (2012)). To reduce the incidence of small bundles on measured dispersion, exporters that cover less than 8% of world variety are dropped.[41] To reduce the incidence of changes in bundle variety on measured dispersion, annual samples of productivity draws are truncated so that exporter-specific bundle variety is the same in each year. According to the model, the best productivity draws are observed first. Hence, sectoral draws are ordered by decreasing magnitude and draws below exporter-specific minimum bundle variety are dropped.

Figure 10: Measured heterogeneity in sectoral productivity (BACI)



Results for the BACI balanced sample are presented first. Recall that sectoral productivity estimates are obtained for each exporter separately, with sectoral

[40] Sectoral productivity is identified within a simplified set-up in which, as in Costinot et al. (2012), capital and labor are combined in all sectors in the same proportion. Hence, productivity measurement is not directly comparable to Levchenko and Zhang (2011).

[41] This restricts the sample to exporters who cover $> 50\%$ of world variety in UNC.

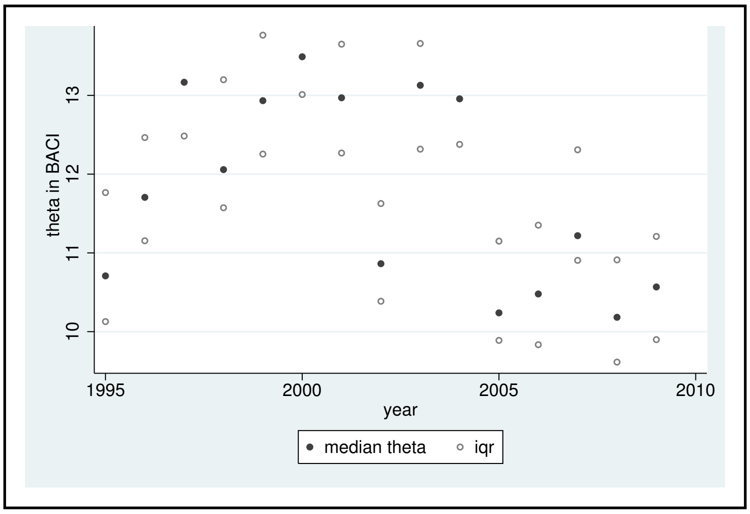
draws normalized by the best exporter-specific draw (sec.3.1). This normalization is inherited by the standard deviation:

$$\tilde{SD}_{it}(z) = SD_{it}(z)/\max(z) \tag{41}$$

As long as the best draw is increasing in average productivity, this normalization is comparable to the standard approach of normalization by the mean (Levchenko and Zhang (2011)).

Fig.10 reports the median and interquartile range of the annual θ-distribution in BACI. The left pane comprises 74 exporters whose bundle variety exceeds 100 products in each year. The median θ fluctuates between 12 and 16, with no clear evolution overtime. The right pane restricts the sample to 43 exporters whose bundle variety exceeds 610 products in each year. This narrows the interquartile range but leaves the median θ unchanged.

Figure 11: Structural heterogeneity in sectoral productivity (BACI)



These results indicate that the sample median is anchored at the extent of heterogeneity observed for exporters with the least amount of truncation while dispersion in θ is attributable to bundles with reduced variety. Since the gap between measured and true dispersion is smallest when truncation is close to nil (Costinot et al. (2012)), the sample median is expected to provide the best approximation to the level of the underlying structural parameter.

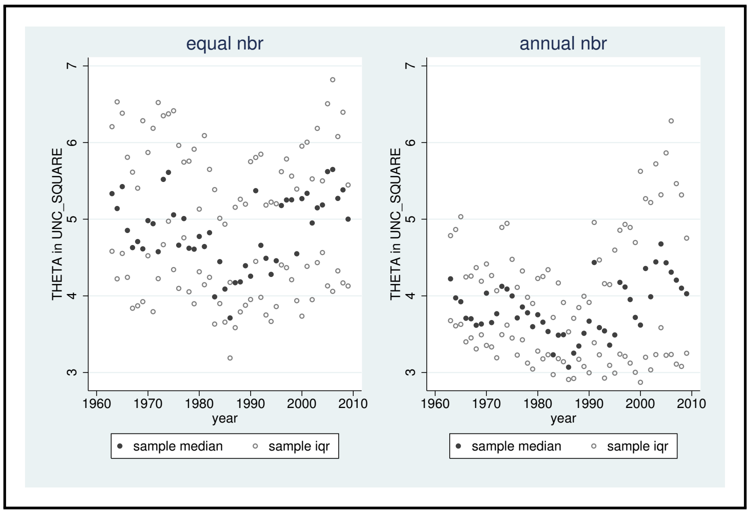
The distribution of θ pictured in fig.10 is obtained when the number of draws in each year is constrained to equal mimimum bundle variety observed for the exporter in 1995-2009. To check whether this leads to an underestimation of productivity dispersion, I recompute the θ-distribution while keeping the full set of annual draws for the 43 exporters who have the best coverage of world variety. Fig.11 shows that this leads to a downward shift of the median θ and establishes the level of structural heterogeneity in the 10-13 range.[42]

Obtaining information on structural heterogeneity becomes trickier if we go further back in time. This is because the normalization implemented to obtain estimates of sectoral productivity in UNC at the SITC 3-digit level is more intricate. Each draw is normalized relatively to the sectoral draw in the benchmark country and further normalized by the productivity draw in the benchmark sector for the exporter relatively to the benchmark country. Consequently, measured dispersion corresponds to the normalized dispersion for the exporter relatively to the normalized dispersion in the benchmark country (USA).

Fig.12 presents results on relative productivity dispersion in UNC for the 23 countries that cover more than 50% of world variety. When the number of exporter-specific draws is constrained to be the same in each year (left pane), the parameter fluctuates in the 4-6 range. When bundle variety is not constrained (right pane), the level of the parameter is reduced to 3-5, and the dispersion in the parameter strongly increases overtime. Focusing on the median exporter, productivity dispersion relatively to the USA has remained broadly unchanged in 1963-2009. Hence, either dispersion has remained broadly unchanged for

[42] $\theta \in \{7,9\}$ at the 2-digit level. The magnitude of θ and σ' are both specific to the extent of data disaggregation. Hence, σ is the fundamental structural parameter of the model: the magnitude of the trade elasticity $|\varepsilon|$ hinges on the ratio $\theta/(\sigma'-1)$ while bundle substitutability is invariant to the degree of detail at the product level.

Figure 12: Relative productivity dispersion (UN COMTRADE)



exporters with the best coverage of world variety, or dispersion has evolved similarly for such exporters and for the USA.

4.2 The magnitude of the trade elasticity

The magnitude of the trade elasticity $|\varepsilon|$ in 1995-2009 is computed using point estimates of σ', $\bar{\sigma}$, and θ obtained in the BACI balanced sample together with the assumption $\bar{\sigma} \approx \sigma$. The level of structural heterogeneity is set at $\theta \in [10.1, 13.5]$. This range corresponds to the set of annual median values obtained in the sample of 43 exporters with the best world variety coverage at the 4-digit level (fig.11). Upper and lower bounds of $|\varepsilon|$ are computed using the magnitude of θ at the upper and lower bounds of the interquartile range. The smoothed estimate of the trade elasticity is obtained by fixing the ratio $\theta/(\sigma' - 1)$ to its median value in 1995-2009 and by using the annualized growth rate to compute the fitted value of $\bar{\sigma}$ in each year.

Results are reported in fig.13. When $\bar{\sigma}$ is identified by instrumenting bundle prices with stocks of physical capital, the magnitude of the trade elasticity increases by 27.4% between 1995 and 2009 (left pane). When $\bar{\sigma}$ is identified by instrumenting bundle prices with stocks of physical capital together with estimates of aggregate TFP (see app.D), the magnitude of the trade elasticity increases by 18.3% (right pane).The level of the parameter increases from 2.42 (2.44) in 1995 to 3.08 (2.89) in 2009 in the former (latter) case.

Figure 13: The Magnitude of the Trade Elasticity in 1995-2009

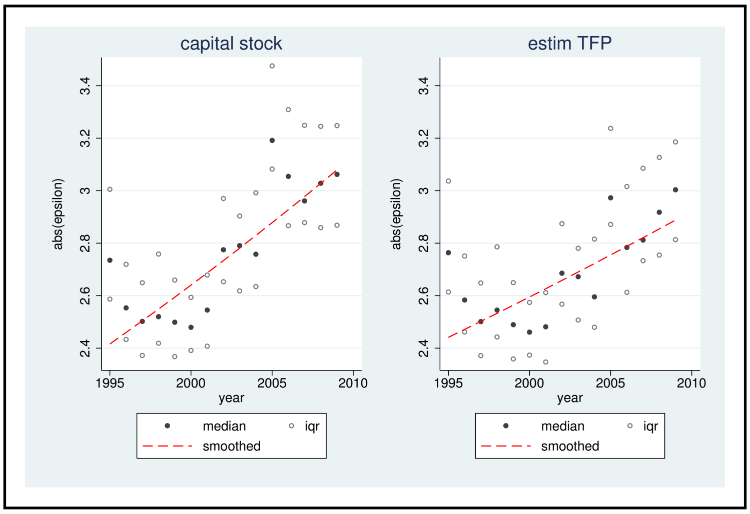
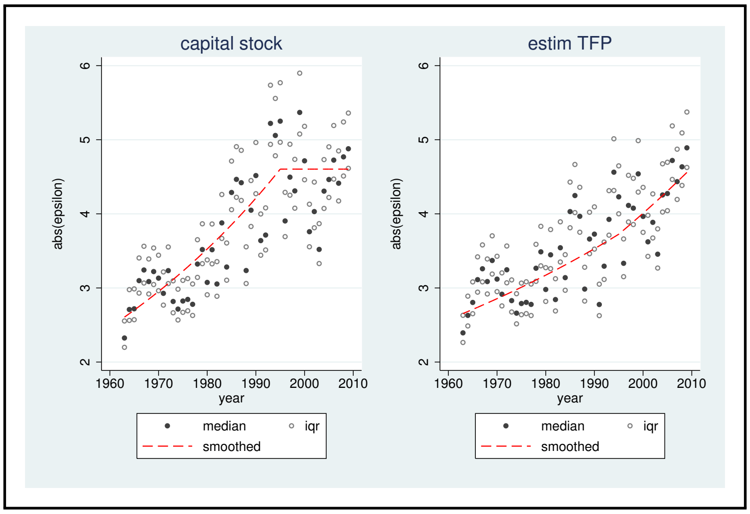


Figure 14: The Magnitude of the Trade Elasticity in 1963-2009



These magnitudes are somewhat lower than the 3.7-5.5 range reported in Simonovska and Waugh (2014b) for 2004 for the benchmark Armington and Ricardian models. This discrepancy may be due to the fact that our model takes into account the contribution of multiple dimensions of heterogeneity in generating the observed distribution of prices and expenditure. The focus in Simonovska and Waugh (2014b) is different. They point out that different magnitudes of the trade elasticity are needed to generate the observed distribution of prices and expenditure in models that focus on a single model-specific dimension of heterogeneity.[43]

The magnitude of the trade elasticity $|\varepsilon|$ in 1963-2009 is computed using point estimates of σ' and $\bar{\sigma}$ obtained in UNC at the 3-digit level. The level of θ is anchored in 1995 by setting the ratio $\theta/(\sigma'-1)$ in UNC equal to the corresponding ratio in BACI in 1995. Upper and lower bounds for this ratio in 1995 are established by using the magnitude of θ at the 25-th and 75-th percentiles in BACI. This establishes the median value of structural heterogeneity at 8.3 in UNC. The corresponding interquartile range is $\theta \in [7.9, 9.1]$).

The estimation of σ' in BACI and in UNC indicates that the lower-tier elasticity is best described as stable in 1995-2009. Results on the θ-distribution in BACI indicate that the ratio $\theta/(\sigma'-1)$ is also best described as stable. This information is taken into account in computing the trade elasticity in UNC by fixing the ratio $\theta/(\sigma'-1)$ in 1995-2009 to its level in 1995. As the magnitude of θ prior to 1995 is unknown, the magnitude of the trade elasticity is computed in 1963-1994 under the assumption that productivity dispersion for the median exporter remains unchanged at the level observed in 1995.[44]

Results are reported in fig.14. When $\bar{\sigma}$ is identified by instrumenting bundle prices with stocks of physical capital (left pane), the magnitude of the trade

[43] Results are consistent with Simonovska and Waugh (2014b) to the extent that these authors show that more flexible models map into lower magnitudes of the trade elasticity.

[44] This assumption is conservative in light of Levchenko and Zhang (2011) results on sectoral technology convergence. Results can also be viewed as reporting a partial effect.

elasticity increases from 2.61 in 1963 to 4.60 in 1995 (+76.6%) and remains unchanged thereafter. This is due to the fact that $\bar{\sigma}$ increases in magnitude prior to 1995 and remains unchanged thereafter.

In the right pane, the upper-tier elasticity is identified by instrumenting bundle prices with stocks of physical capital and estimates of aggregate TFP. Results obtained in this specification for 1995-2009 are consistent with the evolution of $\bar{\sigma}$ in BACI. The magnitude of the trade elasticity increases from 2.65 in 1963 to 3.73 in 1995 (+40.3%) as a consequence of reduced within-bundle substitutability. It further increases to 4.56 by 2009 (+22.4%) as a consequence of increasing substitutability of country-specific product bundles. The total increase of the parameter amounts to 71.7% between 1963 and 2009.

4.3 The magnitude of welfare gains from trade

The model of this paper belongs to the class of models discussed in Arkolakis et al. (2012) in which welfare gains from trade can be computed with help of two sufficient statistics: the trade elasticity ε and the share of expenditure on domestic supply Λ_{ii}. The magnitude of welfare gains from trade is identified with the 'welfare cost of autarky' and defined as the percentage loss in real income incurred by moving to autarky from an equilibrium in which the country relies on foreign supply: $\Delta = 1 - \Lambda_{ii}^{1/|\varepsilon|}$. I characterize the extent of reliance on foreign supply and use this information to quantify the welfare cost of autarky in 1963-2009.

4.3.1 Reliance on foreign supply has increased

Domestic expenditure is obtained by combining production data reported in INDSTAT2 at the ISIC 2-digit level with data on total exports and imports reported in UN COMTRADE at the SITC 3-digit level and converted to the ISIC

2-digit classification.[45] Domestic absorption (Λ_{ii}) is computed as the share of expenditure on domestically supplied goods in the subset of ISIC 2-digit categories for which information on production, exports, and imports is simultaneously available. The indicator measures reliance on domestic supply in manufacturing and is thought of as providing an upper bound on total reliance of the economy on foreign supply.

Figure 15: Reliance on domestic supply (Λ_{ii})

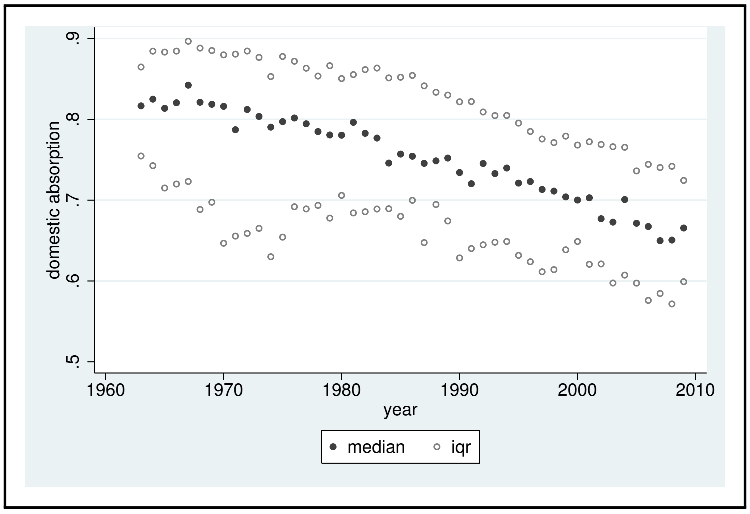


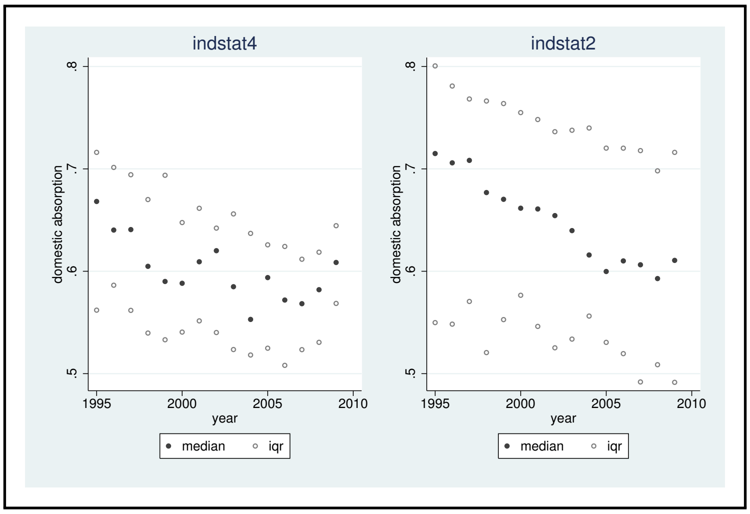
Fig.15 reports the interquartile range of the indicator in the sample of 32 countries observed in ≥ 40 years. Reliance on domestic supply was reduced from 82% in 1963 to 67% in 2009 for the median country. This corresponds to a reduction of 3 percentage points (pp) per decade.[46]

Fig.16 provides a sensitivity check for 1995-2009 by comparing the level of the INDSTAT2 indicator with annual measures of domestic absorption in the UNIDO Supply and Demand Database (SD). The latter dataset provides information on production, exports, and imports at the ISIC 4-digit level. Reliance on domestic supply decreases by .4 pp per year in the sample of 24 countries

[45] SITC is mapped into ISIC using the SITC-HS88 and the HS88-ISIC correspondences.

[46] Λ_{ii} is reduced by .17 pp per year in the full sample in which the number of countries increases from 37 to 64. The reduction in Λ_{ii} at the median is unchanged: -15 pp from 76% to 61%.

Figure 16: Sensitivity of Λ_{ii} to data disaggregation



present in the latter database while it decreases by .5 pp per year in the sample of 35 countries present in INDSTAT2. Reliance on domestic supply for the median country is reduced from 72% in INDSTAT2 (67% in UNIDO SD) in 1995 to 61% in 2009 in both datasets.

The indicator is not highly sensitive to the use of gross output data at the 2- or 4-digit level while country coverage is wider in INDSTAT2. Consequently, I use annual indicators of reliance on domestic supply that cover the full sample of countries in INDSTAT2 to compute the welfare cost of autarky.

4.3.2 The welfare cost of autarky is weakly increasing in 1963-2009

The welfare cost of autarky is obtained by combining information on the magnitude of the trade elasticity with information on domestic absorption in the median country of the INDSTAT2 sample.[47] The magnitude of the trade elasticity $|\varepsilon|$ is taken from the most conservative specification in which the upper-tier elasticity is identified by instrumenting bundle prices with aggregate TFP and stocks of physical capital (App.D).

[47] Upper and lower bounds are computed using the interquartile range of Λ_{ii} in INDSTAT2.

Fig.17 reports the welfare cost of autarky on the basis of BACI estimates of the trade elasticity. In the median country, the cost of autarky increased from 14 to 15.8% of real income between 1995 and 2009 (+2 pp).[48] If the magnitude of the trade elasticity is kept fixed at its maximum (minimum), the increase in the cost of autarky for the median country is magnified to 3.8 (4.4) pp (resp. red and blue dash in fig.17). These magnitudes are very much in line with the 15.3% reported by Costinot and Rodriguez-Clare (2014) in 2008 for the multiple-sector model of the world economy as the average welfare cost of autarky.[49]

Figure 17: The welfare cost of autarky 1995-2009

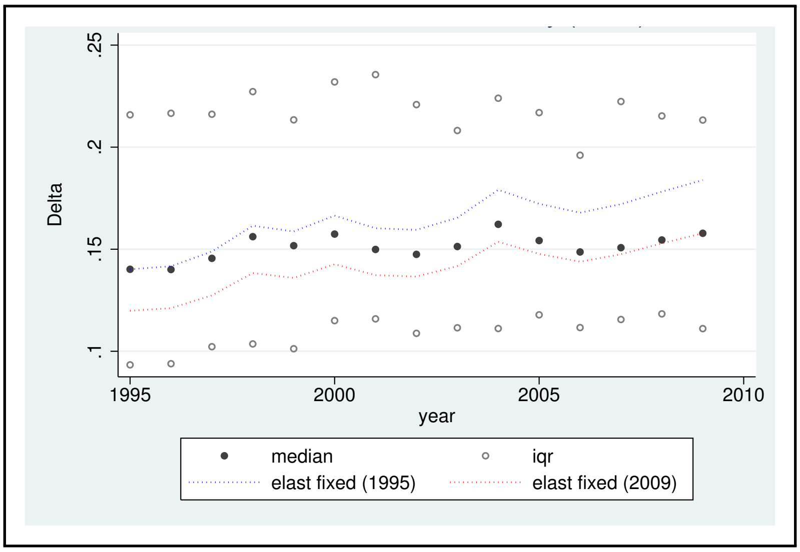


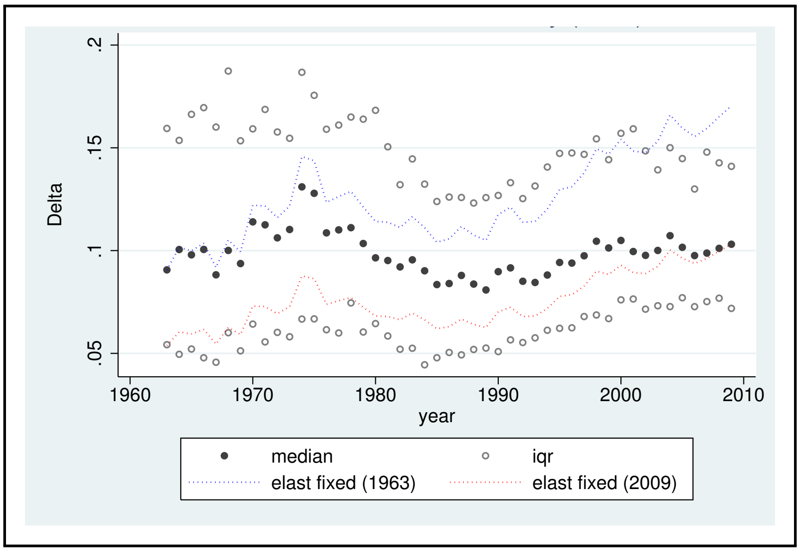
Fig.18 reports the welfare cost of autarky in 1963-2009 on the basis of UNC estimates of the trade elasticity.[50] The welfare cost of autarky is stable at 9.1-9.4% until 1995 and increases to just 10.3% by 2009 at the sample median. The autarky cost increases by roughly 2 pp of real income for the country situated at the 25-th percentile: from 5.4% to 7.2% over 1963-2009. If the trade elasticity

[48] The autarky cost increases from 12.8 to 15.7% for countries present in each year.

[49] The loss of real income at the median is 12.5-13% in Costinot and Rodriguez-Clare (2014).

[50] To ensure the comparability of domestic absorption indicators for 1963-1994 and 1995-2009, the INDSTAT2 sample in 1963-2009 is restricted to the set of 119 countries observed at least once in 1995-2009. This drops 20 countries that are never observed in 1995-2009.

Figure 18: The welfare cost of autarky 1963-2009



were kept fixed at its maximum (minimum), the cost of autarky would have increased from 5.4 to 10.3% (9.1 to 17%) of real income in the median country.

To sum up, the welfare cost of autarky is non-negligible in the 1960s, but its increase over the last 50 years is close to nil. The loss of real income incurred by reverting to autarky has increased by just 1 to 3 percentage points between 1963 and 2009. The mechanical explanation of this remarkable stability is that reliance on foreign supply was low precisely when potential gains from trade were high. Increased reliance on foreign supply has been concomitant to the reduction in the magnitude of potential gains from trade as measured through the increase in the magnitude of the trade elasticity.

5 Conclusion

The main contribution of this paper consists in providing empirical evidence on the evolution of structural parameters that are of paramount importance in workhorse trade models. These are the Armington elasticity, the degree of dispersion in sectoral technology draws, and the trade elasticity. The complementary contribution is methodological. Acknowledging that there remains substantial uncertainty about the role of heterogeneity in supply and heterogeneity in demand in determining the magnitude of the trade elasticity, I propose a simple generalization of the Armington model that makes it feasible to separately identify supply and demand parameters in cross-section with parsimonious data requirements. Hence, I learn from the data which dimensions of heterogeneity determine the evolution of the trade elasticity.

The model spells out how price and expenditure variation reported in the widely available trade data should be decomposed along the three dimensions of heterogeneity present in this economy: within-bundle substitutability, across-bundle substitutability, and producer cost heterogeneity. The structural parameters are identified sequentially. The identification of within-bundle substitutability makes feasible the subsequent identification of producer heterogeneity and of upper-tier substitutability. The magnitude of the trade elasticity is found to have increased from 2.65 in 1963 to 4.56 in 2009 (+72%), with an increase of 18-22% between 1995 and 2009. This increase is mainly attributable to changes in structural demand parameters.

By combining empirical evidence on the magnitude of the trade elasticity with information on effective reliance on foreign supply, I document a non-negligible cost of shutting down trade already in the 1960s. However, welfare gains from trade are only weakly increasing. The loss of real income incurred by reverting to autarky has increased by just 1 to 3 percentage points over 1963-2009, with the median country set to lose 16% of real income in 2009.

This result may appear surprising in the face of overwhelming evidence on the growing interdependence of the economies within the global trade network. The empirical regularity that underpins this result is that increased interdependence has gone hand in hand with the reduction in the extent of structural dissimilarity of these economies. Indeed, if the trade elasticity is held fixed, the percentage point loss in real income incurred by reverting to autarky nearly doubles.

The task of working out whether the intensification of exchange may lead to a reduction in the degree of structural heterogeneity in the economy is beyond the scope of this paper. Nonetheless, the negative relationship between the intensity of exchange and the magnitude of potential gains from trade signals that supply and demand parameters likely result from a dynamic process and likely depend on the past intensity of exchange. The dynamic and possibly cumulative realization of the gains from trade cannot be captured with help of the welfare-cost-of-autarky metric. As a consequence, the numbers reported in this paper are likely understating the magnitude of total gains from trade.

References

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

Archanskaia, E. and G. Daudin (2012). Heterogeneity and the distance puzzle. *FIW Working Paper series* (095).

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012). New trade models, same old gains? *American Economic Review 102*(1), 94–130.

Armington, P. S. (1969). A theory of demand for products distinguished by place of production. *IMF Staff Papers 16*(1), 159–178.

Bems, R. (2014). Intermediate inputs, external rebalancing, and relative price adjustment. *unpublished manuscript*.

Bems, R., R. C. Johnson, and K.-M. Yi (2010). Demand spillovers and the collapse of trade in the global recession. *IMF Economic Review 58*(2), 295–326.

Bems, R., R. C. Johnson, and K.-M. Yi (2011). Vertical linkages and the collapse of world trade. *American Economic Review: Papers and Proceedings 101*(3).

Broda, C. and D. E. Weinstein (2006). Globalization and the gains from variety. *Quarterly Journal of Economics 121*(2), 541–585.

Caliendo, L. and F. Parro (2014). Estimates of the trade and welfare effects of nafta. *Review of Economic Studies (forthcoming)*.

Chaney, T. (2008). Distorted gravity: The intensive and extensive margins of international trade. *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*(2), 581–608.

Costinot, A. and A. Rodriguez-Clare (2014). Trade theory with numbers: Quantifying the consequences of globalization. In G. Gopinath, E. Helpman, and K. Rogoff (Eds.), *Handbook of International Economics*, Volume 4, pp. 197–262. Amsterdam: Elsevier.

Crozet, M. and H. Erkel-Rousse (2004). Trade performances, product quality perceptions, and the estimation of trade price elasticities. *Review of International Economics 12*(1), 108–129.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5).

Eaton, J. and S. Kortum (2010). *Technology in the Global Economy: A Framework for Quantitative Analysis*. Unpublished manuscript.

Eaton, J., S. Kortum, and F. Kramarz (2011). An anatomy of international trade: Evidence from french firms. *Econometrica 79*(5), 1453–1498.

Eaton, J., S. Kortum, and S. Sotelo (2012). International trade: Linking micro and macro. *NBER Working Paper Series* (17864).

Feenstra, R. (1994). New product varieties and the measurement of international prices. *The American Economic Review 84*(1), 157–177.

Feenstra, R., R. Inklaar, and M. Timmer (2013). The Next Generation of the Penn World Table. *NBER Working Paper Series* (19255).

Feenstra, R. C., M. Obstfeld, and K. N. Russ (2014). In Search of the Armington Elasticity. *NBER Working Paper series* (20063).

Feenstra, R. C. and J. Romalis (2014). International prices and endogenous quality. *Quarterly Journal of Economics 129*(2), 477–527.

Gaulier, G. and S. Zignago (2010). BACI: International trade database at the product level. *CEPII working papers* (23).

Gourieroux, C., A. Monfort, and A. Trognon (1984). Pseudo maximum likelihood methods: Theory. *Econometrica 52*(3), 681–700.

Head, K. and T. Mayer (2013). What separates us? Sources of resistance to globalization. *Canadian Journal of Economics 46*(4), 1196–1231.

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

Hummels, D. and G. Schaur (2013). Time as a trade barrier. *American Economic Review 103*, 1–27.

Imbs, J. and I. Méjean (2010). Trade elasticities: A final report for the European Commission. *432*.

Imbs, J. and I. Méjean (2014). Elasticity optimism. *American Economic Journal: Macroeconomics (forthcoming)*.

Johnson, N. L., S. Kotz, and N. Balakrishnan (1994). *Continuous Univariate Distributions*. Wiley.

Johnson, R. C. and G. Noguera (2012). Accounting for intermediates: Production sharing and trade in value added. *Journal of International Economics 86*, 224–236.

Levchenko, A. A. and J. Zhang (2011). The evolution of comparative advantage: Measurement and welfare implications. *NBER Working Paper Series 16806*.

Levchenko, A. A. and J. Zhang (2014). Ricardian productivity differences and the gains from trade. *European Economic Review 65*, 45–65.

Mayer, T. and S. Zignago (2011). Notes on CEPII's distances measures: the GeoDist database. *MPRA Working Paper Series* (36347).

Melitz, M. and S. Redding (2013). New trade models, new welfare implications. *NBER Working Paper Series* (18919).

Reinert, K. and D. Roland-Horst (1992). Armington elasticities for United States manufacturing sectors. *Journal of Policy Modeling 14*(5), 531–539.

Santos Silva, J. M. C. and S. Tenreyro (2006). The log of gravity. *The Review of Economics and Statistics 88*(4), 641–658.

Simonovska, I. and M. Waugh (2014a). The elasticity of trade: Estimates and evidence. *Journal of International Economics 92*(1), 34–50.

Simonovska, I. and M. Waugh (2014b). Trade models, trade elasticities, and the gains from trade. *NBER Working Paper Series* (20495).

Waugh, M. (2010). International trade and income differences. *American Economic Review 100*(5), 2093–2124.

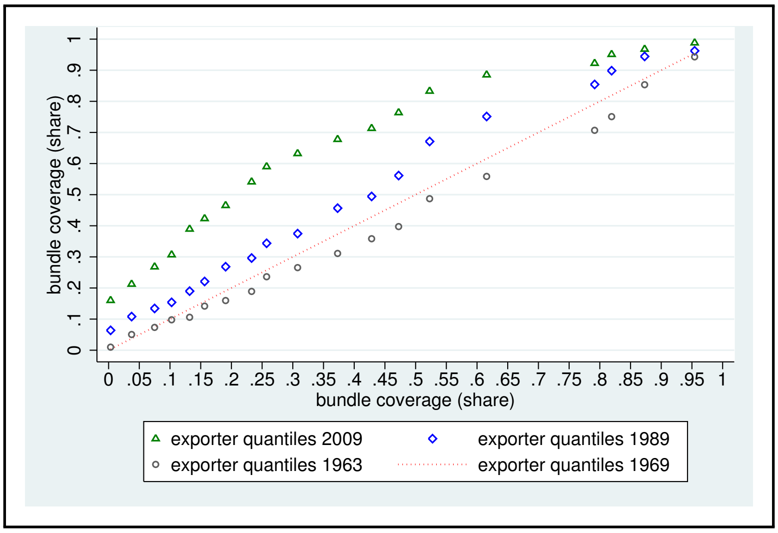
Appendices

A Bundle variety and product overlap

A.1 Product coverage in UN COMTRADE (SITC Rev.1, 1963-2009)

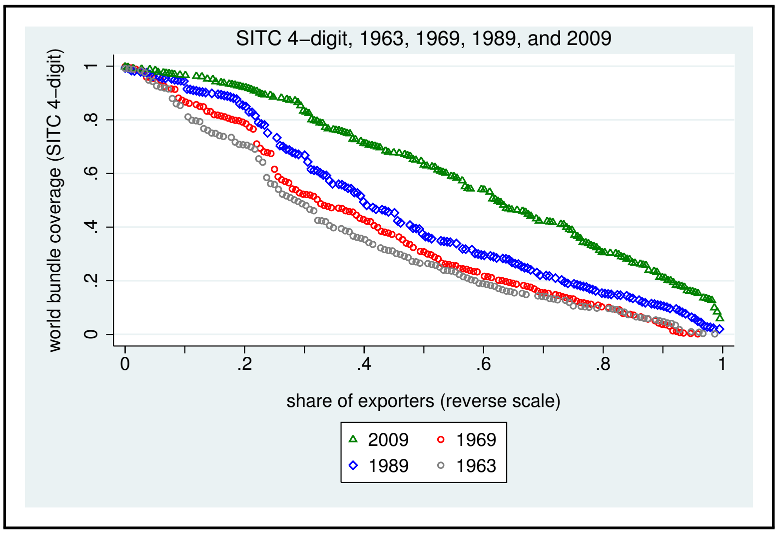
Fragmented product coverage of world markets is documented in 1963-2009 using the SITC 4-digit classification (600-800 products). This classification provides the longest year coverage in UN COMTRADE (UNC). I compute the total number of goods traded on the world market in each year together with the number of goods that each exporter effectively supplies to the world. Sorting exporters in the order of increasing coverage of the world variety and grouping them in 20 bins, I document the distribution of product coverage between 1963 and 2009 with help of a QQ-plot where the 45-degree line corresponds to product coverage in 1969 (fig.19). Moving from left to right, each point on the graph corresponds to bundle coverage reached by successive exporter quantiles. The upward movement of the distribution overtime documents increasing coverage of world variety by exporter-specific product bundles.

Figure 19: QQ-plot: distribution of product coverage in 1963-2009



To further illustrate fragmented product coverage, fig.20 combines information on world variety coverage on the vertical scale with the measure of exporters on the horizontal scale that cover at least that fraction of world variety. In 1963 (resp. 2009) only the upper 20% of exporters cover more than 70% (90%) of the world bundle. Further, about 1/3 of all exporters cover less than 50% of world variety in 2009 while in 1963 this share is as high as 2/3. Hence, product coverage is highly fragmented, notwithstanding a substantial increase in exporter-specific bundle variety between 1963 and 2009.

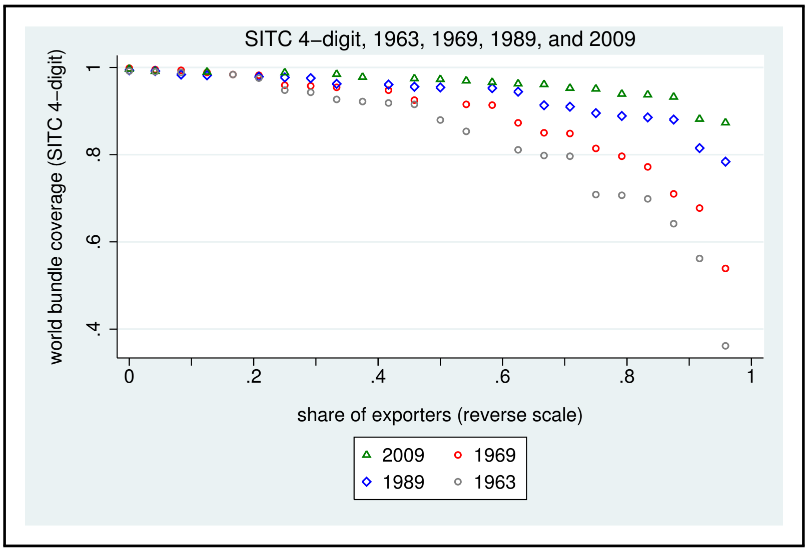
Figure 20: World bundle coverage: full sample



If the set of exporters is restricted to the square sample defined as the set of countries that export to every one of the other countries in each year of the sample, the extent of fragmentation in product coverage is strongly reduced (fig.21). In particular, exporter-specific bundle variety covers more than 80% of world variety in 1990-2009 in this 24-country sample.

Next, I document the extent of fragmentation in market coverage by exporter-specific bundle variety. Specifically, for each exporter I compute the fraction of goods which reach at least $1, 2, \ldots, N$ markets, and scale this by the total number of markets in which the exporter is active. I then compute the rate of decay in bundle variety associated with increasing the share of markets reached by the product mix. The average rate of decay in the pooled sample (all exporters and years) is 1.13: a one percentage point increase in the number of served markets is associated with a 1.13 percentage point reduction in bundle variety. Only

Figure 21: World bundle coverage: square sample

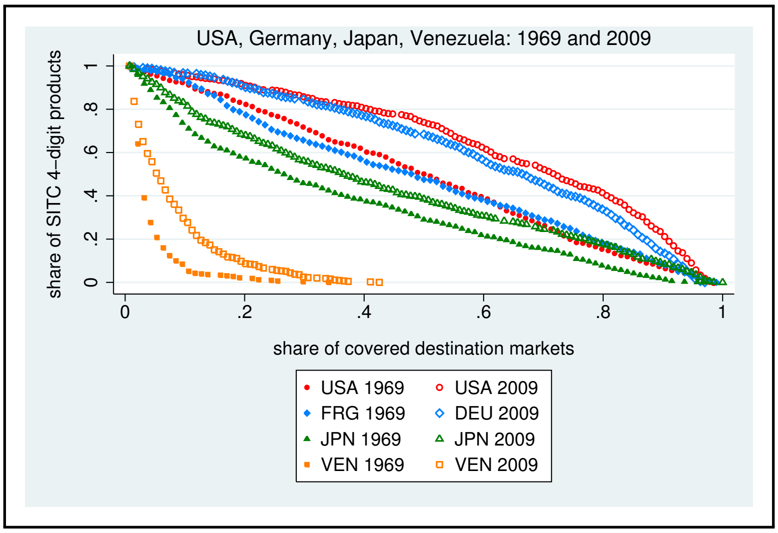


a small fraction of the bundle that the exporter provides to the world market reaches every destination market in which the exporter is active. Hence, zeros are prevalent in sectoral trade. Nonetheless, exporter ability to maintain bundle variety in relatively more markets improves overtime. Thus, the average rate of decay in the sample is reduced from 1.28 in 1963 to 1.05 in 2009.

Fig.22 illustrates these patterns for USA, Germany, Japan, and Venezuela in 1969 and 2009. Focusing on the cross-section, the rate of decay in bundle variety is lower for exporters that have the best coverage of world variety. Thus, USA and Germany have full coverage of world variety in 1969, and 80% of their product bundle reaches 20% of their destination markets. Japan has 96% coverage of world variety in 1969, and just 60% of this bundle reaches 20% of Japan's destination markets. Venezuela covers 54% of world variety in 1969, but less than 5% of this bundle reaches 20% of Venezuela's destination markets. There is a rightward shift of the distribution for each exporter between 1969 and 2009. This illustrates exporters' increased ability to maintain product variety within their set of active destination markets.

Tab.1 illustrates the reduction in the share of zero trade flows (ztf) in terms of world variety coverage in (1), and destination-specific variety coverage in (2). Col.(3) shows that the reduction in the share of ztf has proceeded at a quicker

Figure 22: Evolution of product coverage in world markets



pace in the square sample. In (4)-(7) I investigate the evolution of bilateral product overlap. Product overlap is defined in (4)-(5) as the share of product categories supplied by the source as well as by the reference country (USA in (4), China in (5)) to each destination market. The degree of similarity in the composition of country-specific product bundles is found to have increased overtime independently of the choice of the reference country.

In (6)-(7) product overlap is defined in terms of value shares, i.e. the value share of trade in each destination which occurs in product categories supplied by the source as well as by the benchmark country (Brazil in (6), China in (7)). Product overlap is found to have increased more quickly for pairs in the square sample whenever the reference country is not itself in the square sample (i.e. has sufficiently low initial coverage of world variety).

To sum up, the salient features of the data are the prevalence of zeros in sectoral trade and the increasing product overlap in country-specific composite goods. The complementary finding is that exporters with better world variety coverage export a higher fraction of this bundle to any given market and hence export more goods on average to any given market. Average bundle variety has a one-to-one mapping to unobserved exporter ability in the model.

Table 1: **Bundle variety and bundle similarity in 1963-2009**

depvar:	Share of ZTF			Overlap: count		Overlap: value	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
year	$-0.0074^{\sharp}$	$-0.0025^{\sharp}$	$-0.0023^{\sharp}$	$0.0042^{\sharp}$	$0.0168^{\sharp}$	$0.0107^{\sharp}$	$0.0161^{\sharp}$
square			$-0.0034^{\sharp}$			$0.0039^{\sharp}$	$0.0016^{\sharp}$
Obs	9,381	645,659	645,659	631,656	636,207	634,307	636,202
R^2	0.94	0.61	0.64	0.39	0.54	0.41	0.42
Ref.				USA	CHN	BRA	CHN

Notes: The share of ZTF is computed at the SITC 4-digit level for the world market in (1), separately for each destination market in (2)-(3). Standard errors are clustered by source in (1), by pair in (2)-(9). Source fixed effects are always included. Destination fixed effects are included in (2)-(9). $\sharp$: p<0.01. 'square' gives the change in the slope for pairs in the square sample.

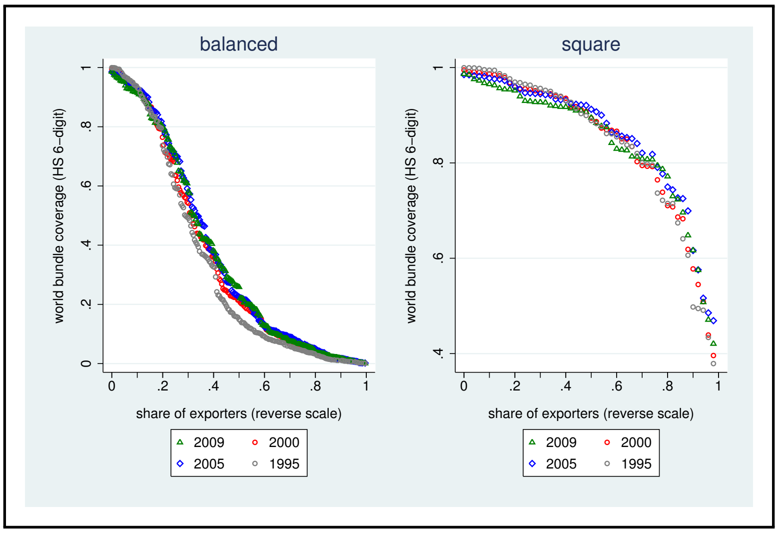
A.2 Product coverage in BACI (HS6, 1995-2009)

I investigate whether the extent of fragmentation is classification-specific by reporting the same sequence of stylized facts at the HS6 level (about 5000 products) in 1995-2009.

Fig.23 documents the extent of fragmentation in world variety coverage in the balanced and square samples in the BACI dataset. The balanced sample comprises 13651 pairs that trade in each year between 1995 and 2009. The square sample comprises 50 countries that trade with each of the other 49 countries in each year between 1995 and 2009. Product coverage is again found to be highly fragmented, with 50% of exporters covering less than 25% of world variety. In the square sample fragmentation is reduced to 40-100% coverage of the world bundle, but there is still substantial variation in product coverage. The distribution of product coverage is largely unchanged in 1995-2009.

Next I document the extent of fragmentation in market coverage by exporter-specific bundle variety. I compute the fraction of HS 6-digit products that reach at least $1, 2, \ldots, N$ markets, and scale this by the total number of markets in which the exporter is active. I then compute the rate of decay in bundle variety

Figure 23: World bundle coverage in BACI



associated with increasing the share of active markets. The average rate of decay in the pooled balanced sample is 1.12. This is virtually identical to the rate obtained at the SITC 4-digit level. Exporter ability to maintain bundle variety in relatively more markets is found to have improved overtime: the average rate of decay is reduced from 1.23 in 1995 to 1.10 in 2009.

Fig.24 illustrates these patterns for USA, Germany, Japan, and Venezuela on the left pane, and for the BRICs on the right pane. Focusing on the cross-section, I find that the rate of decay in bundle variety is lower for exporters that have the best coverage of world variety (USA, Germany, and China in 2009).

Turning to the evolution in product coverage between 1995 and 2009, the graph illustrates strong heterogeneity in terms of improvement in exporter ability to maintain product variety in the set of active destination markets. The most spectacular improvement is experienced by China for which the magnitude of the slope decreases from nearly vertical to about 1.2. By 2009, roughly 40% of products that China exports to at least one destination reaches 50% of markets in which China is active.

Tab.2 illustrates the reduction in the share of zero trade flows (ztf) in terms of world variety coverage in (1), and destination-specific variety coverage in (2). Col.(3) shows that the reduction in the share of ztf has proceeded at a quicker pace in the square sample. In (4)-(7) I investigate the evolution of bilateral

Figure 24: Evolution of product coverage (BACI)

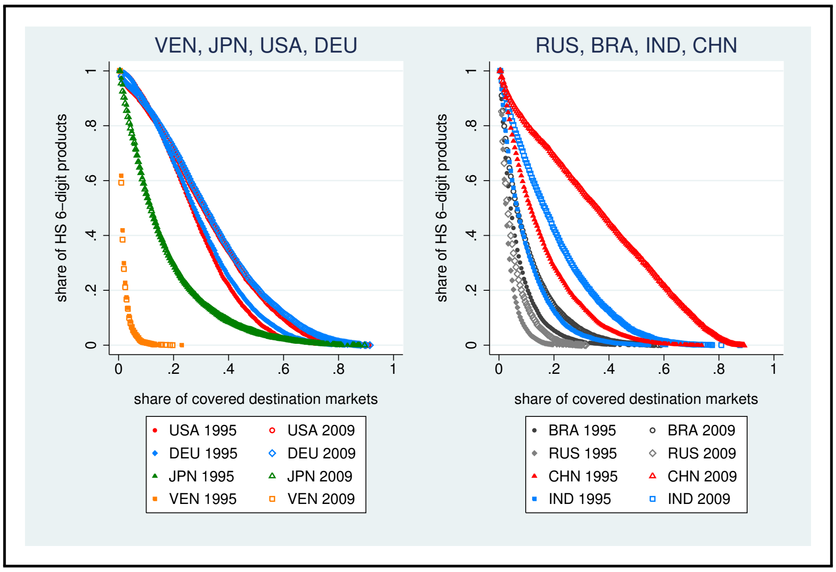


Table 2: **Bundle variety and bundle similarity in 1995-2009 (BACI)**

depvar:	Share of ZTF			Overlap: count		Overlap: value	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
year	$-0.0039^{\sharp}$	$-0.0024^{\sharp}$	$-0.0020^{\sharp}$	$0.0054^{\sharp}$	$0.0258^{\sharp}$	$0.0053^{\sharp}$	$0.0236^{\sharp}$
square			$-0.0026^{\sharp}$			$-0.0031^{\sharp}$	$-0.0032^{\sharp}$
Obs	3,090	204,732	204,732	198,762	199,752	198,762	199,752
R^2	0.99	0.53	0.56	0.80	0.73	0.53	0.49
Ref.				USA	CHN	USA	CHN

Notes: The share of ZTF is computed at the HS 6-digit level for the world market in (1), separately for each destination market in (2)-(3). Standard errors are clustered by source in (1), by pair in (2)-(9). Source fixed effects are always included. Destination fixed effects are included in (2)-(9). ♯: p<0.01. 'square' gives the change in the slope for pairs in the square sample.

product overlap. Product overlap is defined in (4)-(5) as the share of product categories supplied by the source as well as by the reference country (USA in

(4), China in (5)) to each destination market. The degree of similarity in the composition of country-specific product bundles has increased independently of the choice of the reference country.

In (6)-(7) product overlap is defined in terms of the value share of trade in each destination which occurs in product categories supplied by the source as well as by the benchmark country (USA in (6), China in (7)). Product overlap is found to have increased less quickly for pairs in the square sample relatively to each of these two reference countries. This illustrates that variation in the extent of fragmentation within the square sample remains relatively strong between 1995 and 2009.

B Full, balanced, and square samples

B.1 The UN COMTRADE dataset: 1963-2009

The full sample contains between 80 and 169 reporters and between 152 and 219 partners.[51] The stable sample is defined as the set of 132 exporters which export to at least 10 destination markets in each year between 1963 and 2009. The stable sample includes the full set of destination markets in which these 132 exporters are active and effectively covers the full set of destination markets in each year. Trade coverage in the stable sample decreases from 98 to 96% of total trade from 1963 to 1999, and further decreases to 94% of total trade between 2000 and 2009.

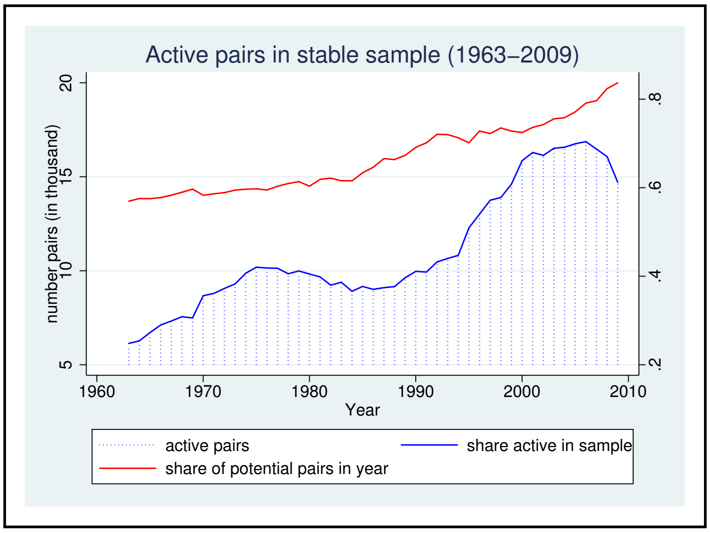
The square sample contains the set of 24 countries which trade with every one of the other 23 countries in each year. In alphabetical order these are: Australia, Austria, Belgium-Luxembourg, Canada, Denmark, Finland, France, Germany, Greece, Hong Kong, Ireland, Israel, Italy, Japan, Malaysia, Mexico, Netherlands, the Philippines, Portugal, Spain, Sweden, Switzerland, United Kingdom, USA. Trade coverage in the square sample fluctuates between 50 and 60% in 1963-2002, and is reduced to 38% of total trade by 2009.

By definition, the square sample contains 552 pairs in each year. The evolution of the number of active pairs and of the share of active pairs out of the total number of possible trading pairs is shown in fig.25. The number of active pairs

[51] The reader is referred to Archanskaia and Daudin (2012) for a detailed investigation of sample effects.

has more than doubled in this stable sample between 1963 and 2009 (in dash, left scale). Active pairs make up between 60% and 80% of the total number of possible trade relationships in any given year, with a clear upward trend since the mid-1980s (in red, right scale). If the focus is placed on the set of pairs that trade a positive amount in at least one year of the sample, the share of active pairs is found to increase by 35 percentage points between 1963 and 2009 (in blue, right scale).[52] The maximum is reached in 2006 when about 2/3 of pairs which trade at least once between 1963 and 2009 are reporting non-zero trade.

Figure 25: Active pairs in the stable sample (1963-2009)



B.2 The BACI dataset: 1995-2009

The full sample contains 212 countries that make up 34566 pairs active at least once. The balanced sample contains 206 partners and 209 reporters that make up the 13651 pairs present in each year between 1995 and 2009. The balanced sample covers between 96 and 100% of trade in the full sample.

The square sample contains 2450 pairs that correspond to the set of 50 countries that trade with every one of the other 49 countries in each year. Trade coverage in the square sample decreases from 80 to 70% of total trade between 1995 and 2009.

[52] Country creation in the early 1990s explains the two-hump shape of the graph.

The countries of the square sample are: Argentina, Australia, Austria, Belgium-Luxembourg, Bulgaria, Brazil, Canada, Chile, China, Colombia, Costa Rica, Croatia, Cyprus, the Czech Republic, Denmark, Ecuador, Egypt, Finland, France, Germany, Greece, Hong Kong, Hungary, India, Ireland, Israel, Italy, Japan, Korea, Malaysia, Mexico, Morocco, Netherlands, Norway, the Philippines, Poland, Portugal, Romania, Singapore, the Slovak Republic, Slovenia, South Africa, Spain, Sweden, Switzerland, Thailand, Turkey, United Kingdom, Uruguay, USA.

C Consistent price aggregation under truncation

This appendix shows that the approximation of the ideal price index that consists in aggregating sectoral prices using observed expenditure weights allows obtaining unbiased estimates of the upper-tier elasticity only when the full product range is observed. In the presence of truncation, the approximation distorts information on relative prices of exporter-specific bundles because the severity of truncation is itself a function of underlying exporter ability.

C.1 Consistent price aggregation for non-truncated bundles

From the model (11) we know that σ' enters the expression of the ideal price index through the argument $\gamma = (\theta + 1 - \sigma')/\theta$ of the Gamma function and through the exponent to which $\Gamma(\gamma)$ is raised. Hence, even if σ' is unknown, the true theoretical price index $P_i(\sigma')$ can be recovered by appropriate rescaling of any other price index $P_i(\tilde{\sigma}')$:

$$P_i(\sigma') \quad = \quad \frac{\kappa(\sigma')}{\kappa(\tilde{\sigma}')} P_i(\tilde{\sigma}') \tag{42}$$

where $\tilde{\sigma}' \in]1, \theta + 1[$ and $\kappa(\sigma') = \{\Gamma(\gamma)\}^{1/(1-\sigma')}$.[53] The scalar $\delta(\tilde{\sigma}') = \kappa(\sigma')/\kappa(\tilde{\sigma}')$ is invariant across exporters. Hence, information on relative prices of exporter-specific product bundles is preserved if price indices are computed using a value of σ' that differs from the true value.

This property is used to prove that the price index $\tilde{P}_i$ obtained by using observed expenditure weights to aggregate sectoral prices instead of the true value

[53] It can be shown that $\kappa(\sigma')$ is increasing in σ' whereby the level of the price index is increasing in σ'.

of σ' can also be used to back out the true theoretical price index $P_i(\sigma')$ by appropriate rescaling.

To see how the approximation works, rearrange the lower-tier demand equation (5) to express the landed price of the non-truncated bundle in terms of any product k:

$$P_i = P_{ik}\left[\frac{X_{ik}}{X_i}\right]^{1/(\sigma'-1)} \tag{43}$$

Denote the total number of categories in the product classification by N and work with the discrete version of the ideal price index by summing (43) across these N categories:

$$P_i = \frac{1}{N}\sum_{k=1}^{N} P_{ik}\left[\frac{X_{ik}}{X_i}\right]^{1/(\sigma'-1)} \tag{44}$$

The approximation to the ideal price index is obtained by replacing the theoretical weights $[X_{ik}/X_i]^{1/(\sigma'-1)}$ with observed expenditure weights:

$$\tilde{P}_i = \sum_{k=1}^{N} P_{ik}\left[\frac{X_{ik}}{X_i}\right] \tag{45}$$

The ordering of bundle prices must be unchanged since (45) is a monotonic transformation of (44). The approximation also preserves the cardinal ranking. This is proved by showing that relative bundle prices are unchanged when expenditure weights generated by the true but unknown lower-tier elasticity are used together with the assumption $\sigma' = 2$ in aggregation.

Denote sectoral expenditure associated with any given value of σ' by $X_{ik}(\sigma')$ and use the associated sectoral demand equation $X_{ik}(\sigma') = [P_{ik}/P_i(\sigma')]^{(1-\sigma')}X_i$ to rewrite observed sectoral weights in terms of weights associated with $\sigma' = 2$.

$$X_{ik}(\sigma') = \frac{[P_{ik}/P_i(\sigma')]^{(1-\sigma')}}{[P_{ik}/P_i(2)]^{(1-2)}} X_{ik}(2) \tag{46}$$

Use the result that $P_i(\sigma') = \delta(2)P_i(2)$ and simplify (46) to get:

$$X_{ik}(\sigma') = \left[\frac{P_{ik}}{P_i(\sigma')}\right]^{(2-\sigma')} \delta(2)X_{ik}(2) \tag{47}$$

Plug (47) in (45) to get:

$$\tilde{P}_i = \sum_{k=1}^{N} P_{ik} \left[\frac{P_{ik}}{P_i(\sigma')}\right]^{(2-\sigma')} \delta(2) \left[\frac{X_{ik}(2)}{X_i}\right] \qquad (48)$$

Use the fact that $[X_{ik}(2)/X_i] = [P_{ik}/P_i(2)]^{(-1)}$ and $P_i(\sigma') = \delta(2)P_i(2)$ to rewrite (48):

$$\tilde{P}_i = \sum_{k=1}^{N} P_{ik} \left[\frac{P_{ik}}{P_i(\sigma')}\right]^{(2-\sigma')} \delta(2) \left[\frac{P_{ik}}{\delta(2)^{-1}P_i(\sigma')}\right]^{-1} \qquad (49)$$

Simplifying (49) leaves two components: $\tilde{P}_i = P_i(\sigma')^{\sigma'-1}\sum_{k=1}^{N} P_{ik}^{(2-\sigma')}$. The second component is linked to the price index $P_i(\tilde{\sigma}')$, with $\tilde{\sigma}' = \sigma' - 1$. This component can be written $\sum_{k=1}^{N} P_{ik}^{(2-\sigma')} = \sum_{k=1}^{N} P_{ik}^{(1-\tilde{\sigma}')} = P_i(\tilde{\sigma}')^{1-\tilde{\sigma}'}$. Combining all of the above gives:

$$\tilde{P}_i = P_i(\sigma')^{\sigma'-1} P_i(\tilde{\sigma}')^{1-\tilde{\sigma}'} \qquad (50)$$

The latter is linked to the true price index: $P_i(\sigma') = \delta(\tilde{\sigma}')P_i(\tilde{\sigma}')$ whereby:

$$\tilde{P}_i = \delta(\tilde{\sigma}')^{\sigma'-2} P_i(\sigma') \qquad (51)$$

The approximation $\tilde{P}_i$ is obtained by appropriate rescaling of the true underlying price index. The scalar $\delta(\tilde{\sigma}')^{\sigma'-2}$ is invariant across exporters.[54] The approximation preserves information on relative prices. Hence, an unbiased estimate of the upper-tier elasticity can be obtained for non-truncated product bundles without prior knowledge of the lower-tier elasticity.

C.2 Consistent price aggregation for truncated bundles

The price of the truncated bundle exceeds the price of the non-truncated bundle for any exporter i.[55] To establish this, consider the partial derivative of

[54] The scalar exceeds 1 whenever $\sigma' > 2$. This is proved by showing that $\partial\kappa(\sigma')/\partial\sigma' > 0$ for $1 < \sigma' < \theta + 1$. The approximation overestimates the price level whenever $\sigma' > 2$ but preserves information on relative prices.

[55] To focus on the incidence of exporter ability, all expressions in this subsection are obtained while conditioning prices and expenditure on destination-specific characteristics and bilateral trade frictions.

the ideal price index for the truncated bundle with respect to the cost threshold $\bar{w}$. The sign of this derivative is determined by the sign of the outer exponent $1/(1-\sigma')$:

$$\frac{\partial \bar{P}_i}{\partial \bar{w}} = \frac{\partial}{\partial \bar{w}} \left\{ \left[\int_0^{\bar{w}} p_i^{1-\sigma'} f(p_i) \mathrm{dp_i} \right]^{1/(1-\sigma')} \right\} < 0 \tag{52}$$

The solution to the integral in (52) is given by (18). Using (18) it is immediate that the second derivative is positive and that the first derivative tends to 0 in the limit:

$$\lim_{\bar{w}\to\infty} \partial \bar{P}_i / \partial \bar{w} = 0$$

The price of the truncated bundle is monotonically increasing in the extent of truncation. Moreover, the gap between the truncated and the non-truncated price is monotonically decreasing in exporter ability. To see this, form the ratio of the two prices using (12) and (22) and consider the partial derivative with respect to $(\tilde{z}_i/c_i)$:

$$\frac{\partial(\bar{P}_i/P_i)}{\partial(\tilde{z}_i/c_i)} = \frac{\partial}{\partial(\tilde{z}_i/c_i)} \left\{ \left[\frac{\upsilon}{\Gamma(\gamma)} \right]^{\frac{1}{(1-\sigma')}} \left[\frac{\tilde{z}_i}{c_i} \right]^{\frac{(\theta-\sigma'+1)-\alpha}{(\sigma'-1)}} \right\} \tag{53}$$

The exponent on $(\tilde{z}_i/c_i)$ can be rearranged as $(1-\sigma)(\theta-\sigma'+1)/(\sigma'-1)^2 < 0$ establishing that the derivative is negative. For any two exporters i' and i such that $(\tilde{z}_{i'}/c_{i'})/(\tilde{z}_i/c_i) < 1$, the relative price of the truncated bundle exceeds the relative price of the non-truncated bundle for the less able exporter: $\bar{P}_{i'}/\bar{P}_i > P_{i'}/P_i$. Hence, truncation increases price dispersion.

Next, consider the approximation that uses observed expenditure weights to compute the price of the truncated product bundle:

$$\tilde{\bar{P}}_i(\sigma', \bar{w}) = \sum_{\{k|P_{ik} \leq \bar{w}\}} P_{ik} \left[\frac{X_{ik}(\sigma')}{\bar{X}_i(\bar{w})} \right] \tag{54}$$

To relate expenditure weights to underlying price indices, rearrange the sectoral demand equation and sum across all prices below the cost threshold:

$$\sum_{\{k|P_{ik}\leq\bar{w}\}} P_{ik}^{1-\sigma'} = \sum_{\{k|P_{ik}\leq\bar{w}\}} \frac{X_{ik}(\sigma')}{X_i} P_i(\sigma')^{1-\sigma'} = \frac{\bar{X}_i(\bar{w})}{X_i} P_i(\sigma')^{1-\sigma'} \tag{55}$$

Truncated expenditure $\bar{X}_i(\bar{w})$ is written as a function of truncated and non-truncated prices:

$$\bar{X}_i(\bar{w}) \quad = \quad \bar{P}_i(\sigma')^{1-\sigma'} P_i(\sigma')^{\sigma'-1} X_i \tag{56}$$

Plugging (56) and (47) in (54) gives:

$$\tilde{\bar{P}}_i(\sigma',\bar{w}) \quad = \quad \sum_{\{k|P_{ik}\leq\bar{w}\}} P_{ik}^{3-\sigma'} P_i(\sigma')^{-1} \delta(2) \bar{P}_i(\sigma')^{\sigma'-1} \left[\frac{X_{ik}(2)}{X_i}\right] \tag{57}$$

To simplify (57), use the fact that $[X_{ik}(2)/X_i] = [P_{ik}/P_i(2)]^{(-1)}$ and $P_i(\sigma') = \delta(2)P_i(2)$:

$$\tilde{\bar{P}}_i(\sigma',\bar{w}) \quad = \quad \sum_{\{k|P_{ik}\leq\bar{w}\}} P_{ik}^{2-\sigma'} \bar{P}_i(\sigma')^{\sigma'-1} \tag{58}$$

Define $\tilde{\sigma}' = \sigma' - 1$ and use the fact that $\bar{P}_i(\tilde{\sigma}')^{1-\tilde{\sigma}'} = \sum_{\{k|P_{ik}\leq\bar{w}\}} P_{ik}^{1-\tilde{\sigma}'}$ to get:

$$\tilde{\bar{P}}_i(\sigma',\bar{w}) = \bar{P}_i(\tilde{\sigma}')^{1-(\sigma'-1)} \bar{P}_i(\sigma')^{\sigma'-1} = \left[\frac{\bar{P}_i(\sigma')}{\bar{P}_i(\tilde{\sigma}')}\right]^{\sigma'-1} \bar{P}_i(\tilde{\sigma}') \tag{59}$$

The approximation delivers the truncated price index for $(\sigma'-1)$ instead of σ'. This index is rescaled by an exporter-specific ratio of truncated price indices $\bar{P}_i(\sigma')/\bar{P}_i(\tilde{\sigma}')$.

Using (22) and defining $\zeta(\sigma') = (\theta - \alpha(\sigma'))/(\sigma'-1)$, the ratio is given by:

$$\frac{\bar{P}_i(\sigma')}{\bar{P}_i(\tilde{\sigma}')} \quad = \quad \left\{\frac{\upsilon(\sigma')^{1/(1-\sigma')}}{\upsilon(\tilde{\sigma}')^{1/(1-\tilde{\sigma}')}}\right\} \left\{\frac{\tilde{z}_i}{c_i}\right\}^{\zeta(\tilde{\sigma}')-\zeta(\sigma')} \tag{60}$$

The first term on the RHS is a positive scalar invariant across exporters. The sign of the exponent in the second term determines the relationship of this ratio with exporter ability $(\tilde{z}_i/c_i)$. Using the definition $\alpha(\sigma') = (\sigma'-\sigma)(\theta-\sigma'+$

$1)/(\sigma' - 1)$, the derivative of $\zeta(\sigma')$ is:

$$\frac{d\zeta(\sigma')}{d\sigma'} = -\frac{(\sigma'-1)\left[2(\theta-\sigma')+\sigma+1\right]+2(\sigma'-\sigma)(\theta-\sigma'+1)}{(\sigma'-1)^3} < 0 \quad (61)$$

As $\tilde{\sigma}' = \sigma' - 1 < \sigma'$, (61) establishes that the exponent is positive: $(\zeta(\tilde{\sigma}') - \zeta(\sigma')) > 0$. Hence, the ratio is increasing in exporter ability: $\partial(\bar{P}_i(\sigma')/\bar{P}_i(\tilde{\sigma}'))/\partial(\tilde{z}_i/c_i) > 0$. The approximation underestimates price dispersion of truncated product bundles: $\tilde{\bar{P}}_{i'}/\tilde{\bar{P}}_i < \bar{P}_{i'}/\bar{P}_i$. It leads to an upward bias in the estimate of bundle substitutability because, as shown in sec.2.2.2, truncation increases dispersion in bundle expenditure and dispersion in bundle prices.

The neutrality of the approximation with respect to relative prices is not preserved under truncation. Consequently, knowledge of the lower-tier elasticity is a necessary prerequisite for obtaining an unbiased estimate of the upper-tier elasticity.

D Aggregate productivity and upper-tier substitutability

This appendix presents estimates of upper-tier substitutability obtained in BACI and in UN COMTRADE when prices of truncated product bundles are instrumented using a proxy of aggregate productivity together with information on physical capital stocks.

As explained in sec.3.1, estimates of sectoral productivity are obtained for each exporter by exploiting variation in average sectoral expenditure across the set of markets in which the exporter is active. In UN COMTRADE, exporter-sector fixed effects $\hat{f}_i(k)$ are estimated by implementing (32). Sectoral productivity $(\sigma' - 1)\ln(z_i(k))$ is identified relatively to a benchmark country and sector. In BACI, exporter-sector fixed effects $\hat{f}_i(k)$ are estimated by implementing (35). Sectoral productivity $(\sigma' - 1)\ln(z_i(k))$ is identified relatively to the best exporter-specific productivity draw.

Estimates of σ'_f are used to compute a normalized index of aggregate productivity as the CES aggregate of estimated sectoral productivity draws after trimming extreme values:

$$\hat{z}_i = \left[\sum_k z_i(k)^{(\sigma'-1)}\right]^{1/(\sigma'-1)} \tag{62}$$

In BACI this normalized index may underestimate variation in exporter ability because the best draw and aggregate TFP are both likely to be higher for more able exporters. In UN COMTRADE the index may fail to pick up variation in exporter ability altogether because the estimate of aggregate productivity is normalized relatively to the same sector for all exporters while the quality of the draw in the benchmark sector is exporter-specific.

I check whether normalized indices of aggregate productivity obtained in BACI and in UN COMTRADE capture the same type of variation in the data as exporter ability measured through variation in bundle variety. Tab.3 summarizes the results.

Aggregate productivity and stocks of physical capital together explain 59% of variation in exporter-specific bundle variety in BACI (col.2). In UN COMTRADE both variables are significant and have the correct sign in the square sample at the 3-digit level although their explanatory power is strongly reduced (col.4). Together they explain just 26% of variation in bundle variety in 1963-2009. At the 2-digit level TFP estimates and bundle variety are negatively correlated in UN COMTRADE once I control for stocks of country-specific physical capital. These results indicate that the estimation strategy performs well in the BACI balanced sample at the 4-digit level but becomes weaker at the 3-digit level in UN COMTRADE. Results are no longer exploitable in UN COMTRADE at the 2-digit level.[56]

The price of the truncated product bundle is instrumented with the normalized index of aggregate TFP and physical capital stocks in BACI at the 4-digit level and in UN COMTRADE at the 3-digit level. Upper-tier substitutability is estimated by regressing bilateral expenditure on instrumented prices (sec.3.2). This is a robustness check for results presented in sec.3.2.3. I construct two graphs for each sample. The first graph reports first-stage coefficients. The second graph reports the coefficient estimated on instrumented prices.

[56] The magnitude of the coefficient on TFP (>1) is plausible only in BACI. If TFP is correctly measured, the magnitude of the coefficient should equal $\theta\gamma > 1$.

Table 3: **Bundle variety and estimated TFP in BACI and UN COMTRADE**

depvar:	BACI 4dgt		UNC 3dgt		UNC 2dgt	
	(1)	(2)	(3)	(4)	(5)	(6)
TFP	2.132***	0.936***	0.129***	0.082**	0.114***	-0.028**
	(0.096)	(0.066)	(0.037)	(0.035)	(0.018)	(0.013)
CAP		0.144***		0.046***		0.184***
		(0.005)		(0.005)		(0.002)
Obs	1,853	1,853	1,081	1,081	5,421	5,248
R^2	0.320	0.592	0.205	0.257	0.093	0.614

"TFP" is log of normalized aggregate productivity; "CAP" is log of physical capital stock.
Bundle variety is computed at HS 6-digit level in (1)-(2); sectoral tfp draws are obtained at 4-digit level.
Bundle variety is computed at SITC 4-digit level in (3)-(6); sectoral tfp draws are obtained at 3(2)-digit level.
Year fixed effects are included: 1995-2009 in (1)-(2); 1963-2009 in (3)-(6). *** p<0.01, ** p<0.05.

Both instruments have the sign predicted by the model in BACI (fig.26) whereby observed prices are decreasing in fundamental exporter ability. This is no longer the case in UN COMTRADE where bundle prices are weakly increasing in estimated TFP (fig.28). Estimated TFP is a weak instrument but its inclusion in the first stage together with physical capital stocks helps stabilize the coefficient estimated in the instrumented specification. The magnitude of upper-tier substitutability fluctuates between 3-3.5 (resp. 4-4.5) in BACI (resp. UN COMTRADE). The elasticity is best described as stable in 1963-1995. The elasticity increases by 12.3% in BACI (fig.27) and by 16.5% in UN COMTRADE between 1995 and 2009 (fig.29).

Figure 26: BACI balanced 4-digit: bundle price and underlying ability

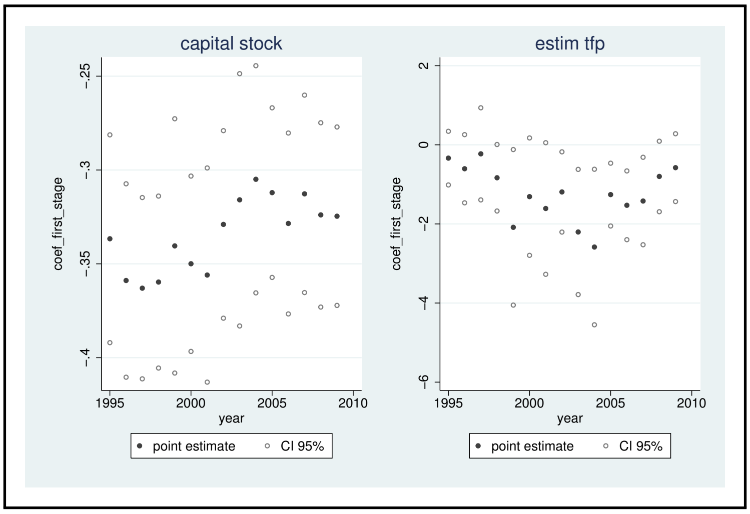


Figure 27: BACI balanced 4-digit: upper-tier substitutability

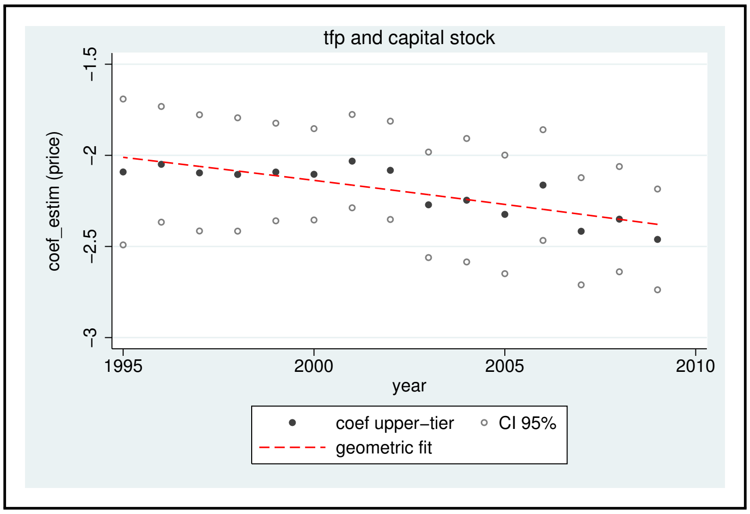


Figure 28: UNC square 3-digit: bundle price and underlying ability

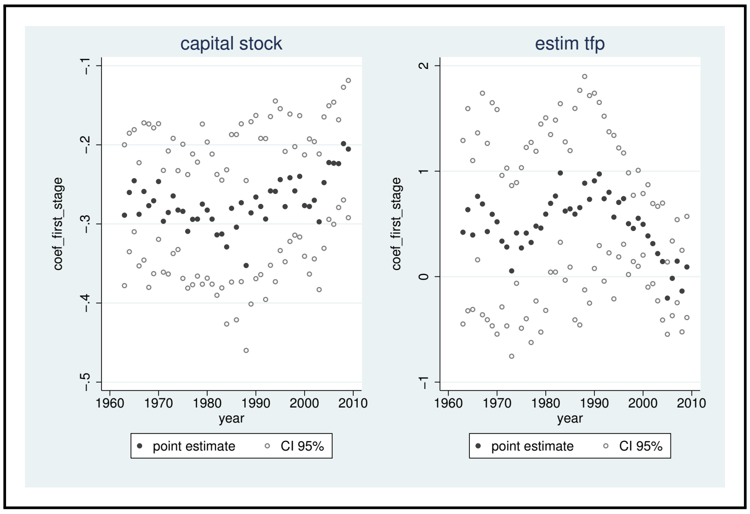


Figure 29: UNC square 3-digit: upper-tier substitutability