Online Supplement to

The Extensive Margin of Exporting Products: A Firm-level Analysis

Costas Arkolakis, *Yale University, CESifo and NBER*
Sharat Ganapati, *Yale University*
Marc-Andreas Muendler, *UC San Diego, CESifo and NBER*

December 18, 2015

This Online Supplement is organized in five sections. Section S1 generalizes the Arkolakis, Ganapati and Muendler (2015, henceforth AGM) model to nested consumer preferences (as previewed in Appendix C to AGM). Each inner nest holds the products in a firm's product line with an elasticity of substitution that differs from that of the outer nest over product lines of different firms. In Section S2, we accommodate market penetration costs as in Arkolakis (2010) and show that our simulated methods of moments estimator is invariant to marketing costs at the level of product lines. Section S3 presents tabulations of the underlying Brazilian export data for 2000. In Section S4, we report Monte Carlo simulations that assess identification under our estimation routine. Finally, Section S5 offers a variety of robustness checks for our main estimates, using alternative assumptions and restricted samples.

S1 Nested Preferences with Different Elasticities

We analyze a generalized version of the AGM model of multi-product firms that allows for within-firm cannibalization effects. The main result is that the qualitative properties of the AGM model are retained: the size distribution of firm sales and the distribution of the firms' numbers of products is consistent with regularities in Brazilian exporter data as well as other data sets. More importantly, the general equilibrium properties of the model do not depend on the inner nests' elasticity (the elasticity across the products of a given firm's product composite) so that the general equilibrium of the model can be easily characterized using the tools of Dekle et al. (2007).

While the model is highly tractable, the introduction of one more demand elasticity adds a further degree of freedom. This degree of freedom can be disciplined using independent estimates for the outer and inner nests' elasticities, such as those of Broda and Weinstein (2006). Under an according parametrization, the model can be used for counterfactual exercises that simulate the impact of changes in trade costs on the firm size distribution and the distribution of the firms' numbers of products.

In the following subsection we present and solve the generalized model. We derive its aggregate properties in subsection S1.2. Subsection S1.3 concludes the presentation of the model with nested preferences.

S1.1 Model

There is a countable number of countries. We label the source country of an export shipment with s and the export destination with d.

We adopt a two-tier nested CES utility function for consumer preferences.[45] Each inner nest of consumer preferences aggregates a firm's products with a CES utility function and an elasticity of substitution ε. Using marketing terminology, the product composite of the inner nest can be called a firm's product line or product mix. The product lines of different firms are then aggregated using an outer CES utility nest with an elasticity σ. Each firm offers a countable number of products but there is a continuum of firms in the world. We assume that every product line is uniquely offered by a single firm, but a firm may ship different product lines to different destinations. Formally, the representative consumer's utility function at destination d is given by

$$U_d = \left(\sum_s \int_{\Omega_{sd}} \left(\sum_{g=1}^{G_{sd}(\omega)} q_{sdg}(\omega)^{\frac{\varepsilon-1}{\varepsilon}} \right)^{\frac{\sigma-1}{\sigma}\frac{\varepsilon}{\varepsilon-1}} \mathrm{d}\omega \right)^{\frac{\sigma}{\sigma-1}}$$

where $q_{sdg}(\omega)$ is the quantity consumed of the g-th product of firm ω, producing in country s. Ω_{sd} is the set of firms from source country s selling to country d.

The representative consumer's first-order conditions imply that demand for the g-th product of firm ω in market d is

$$q_{sdg}(\omega) = p_{sdg}(\omega)^{-\varepsilon} P_{sd}(\omega; G_{sd})^{\varepsilon-\sigma} P_d^{\sigma-1} T_d,$$

where $p_{sdg}(\omega)$ is the price of that product,

$$P_{sd}(\omega; G_{sd}) \equiv \left[\sum_{g=1}^{G_{sd}(\omega)} p_{sdg}(\omega)^{-(\varepsilon-1)} \right]^{-1/(\varepsilon-1)}$$

is the ideal price index for the product line of firm ω selling $G_{sd}(\omega)$ products in market d, and

$$P_d \equiv \left[\sum_s \int_{\Omega_{sd}} P_{sd}(\omega; G_{sd})^{-(\sigma-1)} \, \mathrm{d}\omega \right]^{-1/(\sigma-1)}$$

is the ideal consumer price index in market d. T_d is total consumption expenditure.

[45] Atkeson and Burstein (2008) use a similar nested CES form in a heterogeneous-firms model of trade but their outer nest refers to different industries and the inner nests to different firms within the industry. Eaton and Kortum (2010) present a stochastic model with nested CES preferences to characterize the firm size distribution and their products under Cournot competition. In our model, firms do not strategically interact with other firms. This property of the model allows us to characterize general equilibrium beyond the behavior of individual firms.

S1.1.1 Firm optimization

We assume that the firm has a linear production function for each product. A firm with overall productivity ϕ faces an efficiency $\phi/h(g)$ in producing its g'th product, where $h(g)$ is an increasing function with $h(1) = 1$. We call the firm's total number of products G_{sd} at destination d its *exporter scope* at d. Productivity is the only source of firm heterogeneity so that, under the model assumptions below, firms of the same type ϕ from country s face an identical optimization problem in every destination d. Since all firms with productivity ϕ will make identical decisions in equilibrium, it is convenient to name them by their common characteristic ϕ from now on.[46]

The firm also incurs local entry costs to sell its g-th product in market d: $f_{sd}(g) > 0$ for $g > 1$, with $f_{sd}(0) = 0$. These incremental product-specific fixed costs may increase or decrease with exporter scope. The overall entry cost for market d is denoted by $F_{sd}(G) \equiv \sum_{g=1}^{G} f_{sd}(g)$ and strictly increases in exporter scope by definition.

Profits of a firm with productivity ϕ from country s that sells products $g = 1, ..., G_{sd}$ in d at prices p_{sdg} are

$$\pi_{sd}(\phi) = \sum_{g=1}^{G_{sd}} \left(p_{sdg} - \frac{w_s}{\phi/h(g)} \tau_{sd} \right) p_{sdg}^{-\varepsilon} \cdot P_{sd}(\phi; G_{sd})^{\varepsilon-\sigma} P_d^{\sigma-1} T_d - F_{sd}(G_{sd}). \tag{S.1}$$

We consider the first-order conditions with respect to the prices p_{sdg} of each product g, consistent with an optimal product-line price $P_{sd}(\phi; G_{sd})$, and also with respect to exporter scope G_{sd}. As shown in Appendix S-A to this Online Supplement, the first-order conditions with respect to prices imply a constant markup over marginal cost for all products equal to $\tilde{\sigma} \equiv \sigma/(\sigma-1)$.

Using the constant markup rule in demand for the g-th product of a firm with exporter scope G_{sd} yields optimal sales of the product

$$p_{sdg}(\phi) q_{sdg}(\phi) = \left(\tilde{\sigma} \frac{w_s \tau_{sd}}{\phi/h(g)} \right)^{-(\varepsilon-1)} P_{sd}(\phi; G_{sd})^{\varepsilon-\sigma} P_d^{\sigma-1} T_d. \tag{S.2}$$

Using this result and the definition of $P_{sd}(\phi; G_{sd})$, we can rewrite profits that a firm generates at destination d selling G_{sd} products as

$$\begin{aligned} \pi_{sd}(\phi; G_{sd}) &= P_{sd}(\phi; G_{sd})^{-(\sigma-1)} \frac{P_d^{\sigma-1} T_d}{\sigma} - F_{sd}(G_{sd}) \\ &= H(G_{sd})^{-(\sigma-1)} (\tilde{\sigma} w_s \tau_{sd})^{-(\sigma-1)} \frac{\phi^{\sigma-1} P_d^{\sigma-1} T_d}{\sigma} - F_{sd}(G_{sd}), \end{aligned} \tag{S.3}$$

[46] To simplify the exposition, we assume here that firms face no other idiosyncratic cost components, whereas the AGM model also allows for a destination specific market access cost shock c_d so that a firm in that model is characterized by a pair of shocks (ϕ, c_d). The derivations here can be readily generalized to such idiosyncratic market access costs.

where

$$H(G_{sd}) \equiv \left[\sum_{g=1}^{G_{sd}} h(g)^{-(\varepsilon-1)}\right]^{-1/(\varepsilon-1)}$$

is the firm's product efficiency index.

Similar to Assumption 1 in AGM, we impose the following assumption, which is necessary for optimal exporter scope to be well defined.

Assumption S.1 *Parameters are such that* $Z_{sd}(G) = f_{sd}(G)/[H(G)^{-(\sigma-1)} - H(G-1)^{-(\sigma-1)}]$ *strictly increases in* G.

The expression for $Z_{sd}(G)$ reduces to $Z_{sd}(G) = f_{sd}(G)h(G)^{\sigma-1}$ when $\varepsilon = \sigma$. In that case, Assumption S.1 is identical to the one considered in AGM.

For a firm to enter a destination market, its productivity has to exceed a threshold ϕ^*_{sd}, where ϕ^*_{sd} is implicitly defined by zero profits for the first product:

$$P_d^{\sigma-1} T_d \left[P_{sd}\left(\phi^*_{sd}; 1\right)\right]^{-(\sigma-1)} = \sigma f_{sd}(1).$$

Using the convention $h(1) = 1$ for $G = 1$ in (S.3) yields

$$\left(\phi^*_{sd}\right)^{\sigma-1} = \sigma f_{sd}(1) \frac{\left(\tilde{\sigma} w_s \tau_{sd}\right)^{\sigma-1}}{P_d^{\sigma-1} T_d} . \tag{S.4}$$

Similarly, we can define the threshold productivity of selling G products in market d. The firm is indifferent between introducing a G-th product or stopping with an exporter scope of $G-1$ at the product-entry threshold $\phi^{*,G}_{sd}$ if

$$\pi_{sd}\left(\phi^{*,G}_{sd}; G\right) - \pi_{sd}\left(\phi^{*,G}_{sd}; G-1\right) = 0. \tag{S.5}$$

Using equations (S.3) and (S.4) in this profit equivalence condition, we can solve out for the implicitly defined product-entry threshold $\phi^{*,G}_{sd}$, at which the firm sells G_{sd} or more products,

$$\left(\phi^{*,G}_{sd}\right)^{\sigma-1} = \frac{\left(\phi^*_{sd}\right)^{\sigma-1}}{H(G_{sd})^{-(\sigma-1)} - H\left(G_{sd}-1\right)^{-(\sigma-1)}} \frac{f_{sd}(G_{sd})}{f_{sd}(1)} = \frac{\left(\phi^*_{sd}\right)^{\sigma-1}}{f_{sd}(1)} Z_{sd}(G_{sd}), \tag{S.6}$$

where we define $\phi^{*,1}_{sd} \equiv \phi^*_{sd}$. So, under Assumption S.1, the profit equivalence condition (S.5) implies that the product-entry thresholds $\phi^{*,G}_{sd}$ strictly increase with G and more productive firms will weakly raise exporter scope compared to less productive firms.

Export sales can be written succinctly as

$$\begin{aligned} t_{sd}(\phi) &= \left(\tilde{\sigma}\frac{w_s \tau_{sd}}{\phi}\right)^{1-\varepsilon} P_d^{\sigma-1} T_d \sum_{g=1}^{G_{sd}} h\left(g\right)^{1-\varepsilon} \left[P_d\left(G_{sd}(\phi)\right)\right]^{\varepsilon-\sigma} \\ &= \sigma f_{sd}(1) \left(\frac{\phi}{\phi^*_{sd}}\right)^{\sigma-1} H\left(G_{sd}(\phi)\right)^{-(\sigma-1)} \end{aligned} \tag{S.7}$$

using equation (S.4). This sales relationship is similar in both models with $\varepsilon \neq \sigma$ and models with $\varepsilon = \sigma$. The only difference between the two types of models is that $H(G_{sd})$ depends on ε by (S.3). If the term $H(G_{sd})$ converges to a constant for $G_{sd} \to \infty$, then export sales are Pareto distributed in the upper tail if ϕ is Pareto distributed. Similar to Proposition 1 in the main text, we can state

Proposition S.1 *Suppose Assumption S.1 holds. Then for all s, d:*

- *exporter scope $G_{sd}(\phi)$ is positive and weakly increases in ϕ for $\phi \geq \phi^*_{sd}$;*
- *total firm exports $t_{sd}(\phi)$ are positive and strictly increase in ϕ for $\phi \geq \phi^*_{sd}$.*

Proof. The first statement follows directly from the discussion above. The second statement follows because $H(G_{sd}(\phi))^{-(\sigma-1)}$ strictly increases in $G_{sd}(\phi)$ and $G_{sd}(\phi)$ weakly increases in ϕ so that $t_{sd}(\phi)$ strictly increases in ϕ by (S.7). ■

Similar to AGM, we define *exporter scale* (an exporter's mean sales) in market d as

$$a_{sd}(\phi) = \sigma f_{sd}(1) \left(\frac{\phi}{\phi^*_{sd}} \right)^{\sigma-1} \frac{H\left(G_{sd}(\phi)\right)^{1-\sigma}}{G_{sd}(\phi)}$$

Under a mild condition, exporter scale $a_{sd}(\phi)$ increases with ϕ and thus with a firm's total sales $t_{sd}(\phi)$. The following sufficient condition ensures that exporter scale increases with total sales.

Case C1 *The function $Z_{sd}(g)$ strictly increases in g with an elasticity*

$$\frac{\partial \ln Z_{sd}(g)}{\partial \ln g} > 1.$$

Case S.1 is more restrictive than Assumption S.1 in that the condition not only requires Z_{sd} to increase with g but that the increase be more than proportional. We can formally state the following result.

Proposition S.2 *If $Z_{sd}(g)$ satisfies Case C1, then sales per export product $a_{sd}(\phi)$ strictly increase at the discrete points $\phi = \phi^*_{sd}$, $\phi^{*,2}_{sd}$, $\phi^{*,3}_{sd}$, ...*

Proof. Compared to AGM, $Z_{sd}(g)$ is defined in more general terms, but it enters the relevant relationships in the same way as in AGM before. Case C1 therefore also suffices in the nested-utility model, and the proposition holds (see the Appendix in AGM for details of the proof for non-nested utility). ■

S1.1.2 Within-firm sales distribution

We revisit optimal sales per product and their relationship to exporter scope and the product's rank in a firm's sales distribution. The relationship lends itself to estimation in micro data. Using the productivity thresholds for firm entry (S.4) and product entry (S.6) in optimal sales (S.2) and simplifying yields

$$\begin{aligned} p_{sdg}(\phi)x_{sdg}(\phi) &= \sigma\, Z_{sd}(G_{sd})H(G_{sd})^{\varepsilon-\sigma}\left(\frac{\phi}{\phi^{*,G}_{sd}}\right)^{\sigma-1} h(g)^{-(\varepsilon-1)} \qquad (S.8)\\ &= \sigma\,\frac{f_{sd}(G_{sd})H(G_{sd})^{\varepsilon-1}}{1-[1-h(G_{sd})^{-(\varepsilon-1)}/H(G_{sd})^{-(\varepsilon-1)}]^{\frac{\sigma-1}{\varepsilon-1}}}\left(\frac{\phi}{\phi^{*,G}_{sd}}\right)^{\sigma-1} h(g)^{-(\varepsilon-1)}. \end{aligned}$$

Note that $H(G)^{\varepsilon-\sigma}$ strictly falls in G if $\varepsilon > \sigma$. Under Case C1, the term $Z_{sd}(G_{sd})H(G_{sd})^{\varepsilon-\sigma}$ must strictly increase in G, however, because individual product sales strictly drop as the product index g increases and $h(g)^{-(\varepsilon-1)}$ falls. So, if $Z_{sd}(G_{sd})H(G_{sd})^{\varepsilon-\sigma}$ did not strictly increase in G, average sales per product would not strictly increase, contrary to Proposition S.2.

Compared to AGM, the relationship (S.8) is not log-linear if $\varepsilon \neq \sigma$ and requires a non-linear estimator, similar to the general case in continuous product space (Arkolakis and Muendler 2010). One possibility is a Simulated Method of Moments estimator, extending the one in AGM.

S1.2 Aggregation

To derive clear predictions for equilibrium we specify a Pareto distribution of firm productivity following Helpman et al. (2004) and Chaney (2008). A firm's productivity ϕ is drawn from a Pareto distribution with a source-country dependent location parameter b_s and a shape parameter θ over the support $[b_s, +\infty)$ for all destinations s. The cumulative distribution function of ϕ is $\Pr = 1 - (b_s)^\theta/\phi^\theta$ and the probability density function is $\theta(b_s)^\theta/\phi^{\theta+1}$, where more advanced countries are thought to have a higher location parameter b_s. Therefore the measure of firms selling to country d, that is the measure of firms with productivity above the threshold ϕ^*_{sd}, is

$$M_{sd} = J_s \frac{b_s^\theta}{(\phi^*_{sd})^\theta}. \qquad (S.9)$$

As a result, the probability density function of the conditional productivity distribution for entrants is given by

$$\mu_{sd}(\phi) = \begin{cases} \theta(\phi^*_{sd})^\theta/\phi^{\theta+1} & \text{if } \phi \geq \phi^*_{sd} \\ 0 & \text{otherwise.} \end{cases} \qquad (S.10)$$

We define the resulting Pareto shape parameter of the total sales distribution as $\tilde{\theta} \equiv \theta/(\sigma-1)$.

With these distributional assumptions we can compute a number of aggregate statistics from the model. We denote aggregate bilateral sales of firms from s to country d as T_{sd}.

The corresponding average sales are defined as $\bar{T}_{sd}$, so that $T_{sd} = M_{sd}\bar{T}_{sd}$ and

$$\bar{T}_{sd} \equiv \int_{\phi^*_{sd}} t_{sd}(\phi)\, \mu_{sd}(\phi)\, \mathrm{d}\phi. \tag{S.11}$$

Similarly, we define average local entry costs as

$$\bar{F}_{sd} \equiv \int_{\phi^*_{sd}} F_{sd}(G_{sd}(\phi))\, \mu_{sd}(\phi)\, \mathrm{d}\phi.$$

To compute $\bar{T}_{sd}$, we impose two additional assumptions mirroring Assumptions 2 and 3 in AGM.

Assumption S.2 *Parameters are such that* $\theta > \sigma - 1$ *.*

Assumption S.3 *Parameters are such that the mean market access cost*

$$\tilde{F}_{sd} \equiv \sum_{G=1}^{\infty} f_{sd}(G)^{1-\tilde{\theta}} \left[H\left(G\right)^{1-\sigma} - H\left(G-1\right)^{1-\sigma} \right]^{\tilde{\theta}}$$

is strictly positive and finite.

Then we can make the following statement.

Proposition S.3 *Suppose Assumptions S.1, S.2 and S.3 hold. Then average sales* $\bar{T}_{sd}$ *per firm are a constant multiple of average local entry costs* $\bar{F}_{sd}$

$$\bar{T}_{sd} = \frac{\tilde{\theta}\sigma}{\tilde{\theta}-1} \bar{F}_{sd} = f_{sd}(1)^{\tilde{\theta}} \tilde{F}_{sd}\ .$$

Proof. See Appendix S-C to this Online Supplement. ■

As a result, bilateral expenditure trade shares can be expressed as

$$\lambda_{sd} = \frac{M_{sd}\bar{T}_{sd}}{\sum_k M_{kd}\bar{T}_{kd}} = \frac{J_s(b_s)^{\theta}(w_s\tau_{sd})^{-\theta}\, f_{sd}(1)^{-\tilde{\theta}}\bar{F}_{sd}}{\sum_k J_k(b_k)^{\theta}(w_k\tau_{kd})^{-\theta}\, f_{kd}(1)^{-\tilde{\theta}}\bar{F}_{kd}}, \tag{S.12}$$

an expression that depends on the values of ε and σ only insofar as these parameters affect $\bar{F}_{sd}$ through $H(G)$.

We can also compute mean exporter scope at a destination:

$$\begin{aligned}
\bar{G}_{sd} &= \int_{\phi^*_{sd}} G_{sd}(\phi)\mu_{sd}(\phi)\mathrm{d}\phi \\
&= \left(\phi^*_{sd}\right)^{\theta}\theta \left[\int_{\phi^*_{sd}}^{\phi^{*,2}_{sd}} \phi^{-\theta-1}\mathrm{d}\phi + \int_{\phi^{*,2}_{sd}}^{\phi^{*,3}_{sd}} 2\phi^{-\theta-1}\mathrm{d}\phi + \ldots \right] \\
&= \frac{\left(\phi^{*,2}_{sd}\right)^{-\theta} - \left(\phi^*_{sd}\right)^{-\theta}}{\left(\phi^*_{sd}\right)^{-\theta}} + \frac{\left(\phi^{*,3}_{sd}\right)^{-\theta} - \left(\phi^{*,2}_{sd}\right)^{-\theta}}{\left(\phi^*_{sd}\right)^{-\theta}} + \ldots .
\end{aligned}$$

Completing the integration, rearranging terms and using equation (S.6), we obtain

$$\bar{G}_{sd} = f_{sd}(1)^{\tilde{\theta}} \sum_{g=1}^{\infty} Z_{sd}\left(g\right)^{-\tilde{\theta}}. \tag{S.13}$$

For the average number of products to be well defined and finite we require one more assumption:

Assumption S.4 *Parameters are such that* $\sum_{g=1}^{\infty} Z_{sd}(g)^{-\tilde{\theta}}$ *is strictly positive and finite.*

In Appendix S-C to this Online Supplement we show that the firms' fixed cost expense is a constant share of their total sales (where we denote means using a bar), as summarized in Proposition S.3:

$$\frac{\bar{F}_{sd}}{\bar{T}_{sd}} = \frac{\tilde{\theta}-1}{\tilde{\theta}\sigma}.$$

We derive aggregate welfare in Appendix S-D to this Online Supplement and demonstrate in Appendix S-E to this Supplement that wage income and profit income can be expressed as a constant share of total output y_s per capita:

$$\pi_s = \eta y_s, \quad w_s = \left(1-\eta\right) y_s,$$

where $\eta \equiv 1/(\tilde{\theta}\sigma)$. Since aggregates of the model do not depend on ε, the equilibrium definition is the same as in AGM.

S1.3 Summary

We have characterized an extension of the AGM model, in which the elasticity of substitution between a firm's individual products does not equal the elasticity of substitution across product lines of different firms. The extended model retains the main qualitative implications of the baseline AGM model, in which the two elasticities are the same. Future work using the structure of the generalized model to obtain estimates of the two elasticities may lead to a better understanding of the substitution effects within and across firms.

S2 Combination of Market Access and Market Penetration Cost Definitions

We turn to a generalization of AGM to nest both market access costs (as in AGM) and market penetration costs (as in Arkolakis 2010) as special cases.

S2.1 Restatement of AGM market access costs

We retain from AGM the specification that a firm draws not only a productivity parameter ϕ but also a destination specific market access cost shock c_d with well defined moments (and possibly a non-unitary mean). Suppose any two firms from source country s happen to draw identical productivity ϕ and happen to draw an identical market access cost parameter $c_d \in (0, \infty)$; those two firms face an identical optimization problem in every destination d at the time of their product access decision. The pair of shocks (ϕ, c_d) therefore completely characterizes a firm's market access decision.

To accommodate market penetration costs as in Arkolakis (2010), we extend the market access cost definition (from AGM) and postulate that a firm's incremental market access cost also depends on its optimal choice of market penetration: a firm from country s decides the fraction n_{sd} of the L_d consumers who the firms wants to reach with its product composite (product line or mix) shipped to destination d. Consistent with the treatment of a firm in Arkolakis (2010) as the seller of a single product line (brand), we adopt the convention that a firm picks a common penetration rate for all its $g = 1, \ldots, G_{sd}$ products shipped to a destination ($n_{sdg} = n_{sd}$ for all g).[47]

As in AGM, a firm (ϕ, c_d) faces a product-destination specific *incremental market access cost* $c_d \, \bar{f}_{sd}(g; n_{sd})$, where $c_d \in (0, \infty)$ is a stochastic firm-specific market access cost shock. A firm that adopts an exporter scope of G_{sd} at destination d therefore incurs a total market access cost of

$$F_{sd}\left(G_{sd}, c_d; n_{sd}\right) = \textstyle\sum_{g=1}^{G_{sd}} c_d \bar{f}_{sd}(g; n_{sd}). \tag{S.14}$$

For any positive market penetration choice $n_{sd} > 0$, the firm's market access cost is zero at zero scope and strictly positive otherwise:

$$\bar{f}_{sd}(0; n_{sd}) = 0 \quad \text{and} \quad \bar{f}_{sd}(g; n_{sd}) > 0 \quad \text{for all } g = 1, 2, \ldots, G_{sd},$$

where $\bar{f}_{sd}(g; n_{sd})$ is a continuous function in $[1, +\infty) \times [0, +\infty)$.

Arkolakis (2010) uses specific functional forms for market penetration costs, derived from primitives on consumer demand and product marketing. We discuss the generalized market access and market penetration cost definition also in terms of those specific functional forms. Extending (7) in AGM, we specify

$$\begin{aligned} \bar{f}_{sd}(g; n_{sd}) &= f_{sd}(n_{sd}) \cdot g^{\delta_{sd}} && \text{for } \delta_{sd} \in (-\infty, +\infty) \qquad \text{and} \\ h(g) &= g^{\alpha} && \text{for } \rho \in [0, +\infty). \end{aligned} \tag{S.15}$$

The market access cost parameter $f_{sd}(n_{sd})$ is zero at zero penetration and strictly positive otherwise:

$$f_{sd}(0) = 0 \quad \text{and} \quad f_{sd}(n_{sd}) > 0 \quad \text{for all } n_{sd} > 0$$

where $f_{sd}(n_{sd})$ is a continuous function in $[0, +\infty)$.

Recall from AGM that a firm also faces a multiplicative i.i.d. shock ξ_{sdg} to its g-th product's appeal at a destination d (with mean $\mathbb{E}\left[\xi_{sdg}(\omega)\right] = 1$, positive support and known realization at the time of consumer choice). Under CES consumer demand, it

[47] A further generalization that allows for firm-product specific optimal choices of n_{sdg} would result in interesting novel relations between core competency in production and market penetration choices: a firm's efficiency schedule $\phi_g \equiv \phi/h(g)$ would interact with product-specific market penetration costs in the product adoption decisions. We leave this generalization for future work.

is irrelevant whether the firm sets optimal price before or after the firm observes the product's appeal realization (see footnote 15 in AGM); price is a proportional markup over the firm-product's marginal production cost irrespective of the size of demand.

However, consistent with AGM and the deterministic setup of Arkolakis (2010), the firm has to take both the product entry (exporter scope G_{sd}) and the market penetration decision (consumer fraction n_{sd}) prior to observing any product appeal shock. It follows that, for a firm with an optimal and strictly positive market penetration at a destination d ($n_{sd} > 0$), the first-order conditions for a firm (ϕ, c_d) and therefore its optimal exporter scope G_{sd} and individual product sales are identical to those presented in AGM—with only two differences in interpretation:[48] we replace the product-invariant part of incremental market access costs with $f_{sd} \equiv f_{sd}(n_{sd})$ and we replace the revenue shifter (equation (8) in AGM) with $D_{sd} = D_{sd}(n_{sd})$ using

$$D_{sd}(n_{sd}) \equiv n_{sd} \cdot \bar{D}_{sd} \qquad \text{for } \bar{D}_{sd} \equiv \left(\frac{P_d}{\tilde{\sigma} \tau_{sd} w_s} \right)^{\sigma-1} \frac{T_d}{\sigma} \tag{S.16}$$

under a given (optimal) market penetration rate $n_{sd} \in (0, 1]$. Our original AGM model is the special case with $n_{sd} = 1$.

S2.2 Generalization of market penetration costs from Arkolakis (2010)

We now show that a specific functional form for $f_{sd}(n_{sd})$ accommodates Arkolakis (2010) market penetration costs as a special case and preserves a firm's relevant optimality conditions from Arkolakis (2010). The wage bill required to reach n_{sd} consumers in a market of size L_d is $F_{sd}(\cdot, \cdot; n_{sd})$, where L_d is a parameter for the firm and n_{sd} is a decision variable.[49] For a firm with given optimal exporter scope G_{sd} and market access cost draw c_d, define the firm's market penetration cost function $F_{sd}(\cdot, \cdot; n_{sd})$ to be equal to its *total market access cost* from (S.14) with

$$F_{sd}(\cdot, \cdot; n_{sd}) \equiv \sum_{g=1}^{G_{sd}} c_d \, \bar{f}_{sd}(g; n_{sd}) = \frac{(L_d)^{\rho}}{\psi_{sd}(\cdot, \cdot)} \frac{1 - (1 - n_{sd})^{1-\beta}}{1 - \beta} \quad \text{for } \beta \in (0, 1) \cup (1, +\infty), \tag{S.17}$$

where

$$\psi_{sd}(G_{sd}, c_d) \equiv \frac{\bar{\psi}}{(w_s)^{\gamma} (w_d)^{1-\gamma}} \frac{1}{c_d \cdot \sum_{g=1}^{G_{sd}} g^{\delta_{sd}}}$$

[48] In AGM, f_{sd} is sometimes also stated as $f_{sd}(1)$ for the first product.

[49] Arkolakis (2010) includes market size L_d as an argument in the market penetration cost function and formally states a functional form for $F_{sd}(\cdot, \cdot; n_{sd}) \equiv f(n_{sd}; L_d)$ in equation (2). Arkolakis (2010) treats $f(n_{sd}; L_d)$ as the labor requirement needed to reach $n_{sd} L_d$ consumers and uses a factor of proportionality ψ to standardize the requirement given the composite wage payment $(w_s)^{\gamma} (w_d)^{1-\gamma}$. For comparability to AGM, we think of $F_{sd}(\cdot, \cdot; n_{sd})$ in wage bill equivalents and standardize with an accordingly scaled factor of proportionality ψ_{sd}.

and $\bar{\psi}$ is a positive scalar (similar to the original ψ from Arkolakis 2010). Importantly, the generalization of the market penetration cost from Arkolakis (2010) to $F_{sd}(\cdot,\cdot;n_{sd})$ in equation (S.17) preserves the four relevant properties of the market penetration cost function: (i) the market penetration cost vanishes at zero penetration since $F_{sd}(\cdot,\cdot;0)=0$, (ii) it strictly increases in n since $\partial F_{sd}(\cdot,\cdot;n)/\partial n > 0$ for $n \in [0,1]$, (iii) it is convex in n since $\partial^2 F_{sd}(\cdot,\cdot;n)/(\partial n)^2 > 0$ for $n \in [0,1]$, and (iv) it is unbounded since $\lim_{n\to\infty} F_{sd}(\cdot,\cdot;n) = +\infty$.

Equivalently, for consistency with AGM and $c_d\,\bar{f}_{sd}(g;n_{sd}) = c_d\,f_{sd}(n_{sd})\cdot g^{\delta_{sd}}$ by equation (S.15), we define

$$f_{sd}(n_{sd}) \equiv \frac{(w_s)^{\gamma}(w_d)^{1-\gamma}(L_d)^{\rho}}{\bar{\psi}}\,\frac{1-(1-n_{sd})^{1-\beta}}{1-\beta} \quad \text{for } \beta \in (0,1)\cup(1,+\infty). \tag{S.18}$$

These mutually consistent but alternative fixed cost definitions allow us to now switch perspective between a firm's optimality conditions for the market penetration rate n_{sd} (given the optimal exporter scope G_{sd} and the market access cost draw c_d) on the one hand side, and a firm's optimality conditions for exporter scope G_{sd} (given the optimal market penetration rate n_{sd} and the market access cost draw c_d) on the other hand side. Both the optimal exporter scope G_{sd} and the market penetration rate n_{sd} decisions need to be made prior to observing the products' appeal shocks (ξ_{sdg}), so a firm takes the two decisions simultaneously.

We already pointed out above that the optimality conditions on exporter scope G_{sd} are the same as those in AGM, merely replacing $f_{sd} \equiv f_{sd}(n_{sd})$ and $D_{sd} \equiv D_{sd}(n_{sd})$ for given optimal n_{sd}. We now turn to showing that the new optimality conditions for the market penetration rate n_{sd}, given optimal exporter scope G_{sd} and the market access cost draw c_d, are a straightforward restatement of the related optimality conditions from Arkolakis (2010), simply generalizing the cost scalar to $\psi_{sd}(G_{sd},c_d)$. The original Arkolakis (2010) model is the special case with $G_{sd}=c_d=1$. The original Melitz (2003) model is a special case with both $G_{sd}=c_d=1$ and $\beta=0$.

S2.3 Optimal market penetration costs given optimal exporter scope

Given the simultaneous choice of exporter scope and the market penetration rate, we can solve without loss of generality for the market penetration rate n_{sd} presuming that exporter scope G_{sd} is optimal. Suppose, conditional on destination market access, a type (ϕ, c_d) firm is setting optimal individual product prices (facing a fraction n_{sd} of consumer demand under monopolistic competition) and optimal exporter scope (given market penetration n_{sd}). The resulting first-order conditions from the profit maximizing equation imply identical markups over marginal cost $\tilde{\sigma} \equiv \sigma/(\sigma-1) > 1$ for each firm-product under

$\sigma > 1$.[50] Given optimal exporter scope $G_{sd}(\phi, c_d)$, and using the optimal pricing decision in the firm's profit function, we obtain the firm's expected profits (prior to product appeal shock realizations) at destination d:

$$\pi_{sd}(\phi, c_d) = \max_{n_{sd}} D_{sd}(n_{sd})\, \phi^{\sigma-1}\, \bar{H}\big(G_{sd}\big)^{-(\sigma-1)} - F_{sd}\left(G_{sd}, c_d; n_{sd}\right),$$

with $F_{sd}\left(G_{sd}, c_d; n_{sd}\right)$ given by (S.17), the penetration dependent revenue shifter given by

$$D_{sd}(n_{sd}) \equiv n_{sd} \cdot \bar{D}_{sd} \qquad \text{for } \bar{D}_{sd} \equiv \left(\frac{P_d}{\tilde{\sigma}\tau_{sd}w_s}\right)^{\sigma-1} \frac{T_d}{\sigma}$$

and the average product efficiency index in destination d for a firm with exporter scope G_{sd} given by $\bar{H}\big(G_{sd}\big)^{-(\sigma-1)} \equiv \sum_{g=1}^{G_{sd}} h(g)^{-(\sigma-1)}$.

The first-order condition for maximizing profit $\pi_{sd}(\phi, c_d)$ with respect to the market penetration rate n_{sd} is equivalent to

$$\frac{\bar{D}_{sd}}{L_d}\, \phi^{\sigma-1}\, \bar{H}(G_{sd})^{-(\sigma-1)} = \frac{1}{\psi_{sd}(G_{sd}, c_d)\, (L_d)^{1-\rho}} \frac{1}{(1-n_{sd})^{\beta}}, \tag{S.19}$$

given $\beta \in (0,1) \cup (1, +\infty)$ for a type (ϕ, c_d) firm with optimal exporter scope $G_{sd}(\phi, c_d)$. Similar to the first-order condition in Arkolakis (2010, equation (8)), the left-hand side of the condition shows the marginal revenue of a firm's product line (net of labor production cost) per consumer and the right-hand side the marginal cost per consumer of bringing the product line to destination d.

The zero-consumer threshold of minimal productivity for a firm to start penetrating a market can be found by setting $n_{sd} = 0$ in (S.19) and solving out for productivity. The zero-consumer threshold for productivity is

$$\phi_{sd}^{*,n=0}(c_d)^{\sigma-1} \equiv \frac{(L_d)^{\rho-1}}{\psi_{sd}(G_{sd}, c_d)} \frac{\bar{H}(G_{sd})^{\sigma-1}}{\bar{D}_{sd}/L_d}. \tag{S.20}$$

A firm compares the marginal per-consumer revenue from reaching an infinitesimally small mass of consumers (the left-hand side of (S.19)) to the marginal per-consumer cost of reaching that infinitesimally small mass (the right-hand side of (S.19)). Given elastic CES

[50] A firm selling an optimal number of products G_{sd} to destination market d has an expected profit of

$$\pi_{sd}(\phi, c_d) = \max_{G_{sd}} \sum_{g=1}^{G_{sd}} \mathbb{E}\left[\max_{\{p_{sdg}\}_{g=1}^{G_{sd}}} \left(p_{sdg} - \tau_{sd}\frac{w_s}{\phi/h(g)}\right)\left(\frac{p_{sdg}}{P_d}\right)^{-\sigma} \xi_{sdg}\frac{T_d}{P_d}\right] - F_{sd}\left(G_{sd}, c_d; n_{sd}\right).$$

The firm's first-order conditions with respect to every individual price p_{sdg} imply an optimal product price

$$p_{sdg}(\phi) = \tilde{\sigma}\, \tau_{sd}\, w_s\, h(g)/\phi$$

with an identical markup over marginal cost $\tilde{\sigma} \equiv \sigma/(\sigma-1) > 1$ for $\sigma > 1$. Product price does not depend on the appeal shock realization because the shock enters profits multiplicatively, so it is irrelevant whether a firm is setting price before or after the appeal shock is observed.

demand, more productive firms extract higher marginal per-consumer revenue, so they choose higher rates of market penetration. Similar to Arkolakis (2010, Proposition 1), for $\beta > 0$, a type (ϕ, c_d) firm will choose to stay out of destination d and set $n_{sd}(\phi, c_d) = 0$ if $\phi < \phi_{sd}^{*,n=0}(c_d)$. Conversely, two firms of types (ϕ_1, c_d) and (ϕ_2, c_d) will enter and set $n_{sd}(\phi_2, c_d) > n_{sd}(\phi_1, c_d) \geq 0$ if $\phi_2 > \phi_1 \geq \phi_{sd}^{*,n=0}(c_d)$.

Inverting the first-order condition (S.19) to solve for the optimal $n_{sd}(\phi, c_d)$, and using (S.20), yields the optimal market penetration rate for a firm's product line

$$n_{sd}(\phi, c_d) = 1 - \left(\frac{\phi_{sd}^{*,n=0}(c_d)}{\phi}\right)^{(\sigma-1)/\beta} \qquad \text{if } \phi \geq \phi_{sd}^{*,n=0}(c_d). \tag{S.21}$$

S2.4 Equilibrium properties

In the combined model with both AGM market access costs and Arkolakis (2010) market penetration costs, the optimal exporter scope choice implies that the productivity threshold $\phi_{sd}^{*,1}(c_d)$ for exporting at all from s to d is $\phi_{sd}^{*,1}(c_d)$, while the zero-consumer threshold is $\phi_{sd}^{*,n=0}$. Both need to be satisfied, so the *effective entry threshold* is $\phi_{sd}^{*}(c_d) = \max[\phi_{sd}^{*,1}(c_d), \phi_{sd}^{*,n=0}]$.[51]

Note that $\psi_{sd}(G_{sd}, c_d)$ strictly monotonically decreases in c_d by (S.17), so the zero-consumer threshold $\phi_{sd}^{*,n=0}(c_d)$ strictly monotonically increases in c_d by (S.20).[52] Similarly, by AGM's equation (10) the productivity threshold for exporting at all ($G_{sd} \geq 1$) strictly increases in c_d. We conclude that, also in the combined model with both AGM market access costs and Arkolakis (2010) market penetration costs, a higher market access cost draw c_d strictly raises the effective entry threshold $\phi_{sd}^{*}(c_d)$.

However, the effect of a higher market access cost draw c_d on realized total market access cost $F_{sd}(G_{sd}, c_d; n_{sd})$ is ambiguous by (S.17). The reason is that $\psi_{sd}(G_{sd}, c_d)$ strictly monotonically decreases in c_d with a unitary elasticity, thus raising F_{sd} with a unitary

[51] By AGM's equation (10), the productivity threshold for exporting at all ($G_{sd} = 1$) is implicitly given by

$$\begin{aligned} \phi_{sd}^{*,1}(c_d)^{\sigma-1} &\equiv \frac{c_d\, f_{sd}\big(n_{sd}(\phi_{sd}^{*,1}, c_d)\big)}{D_{sd}\big(n_{sd}(\phi_{sd}^{*,1}, c_d)\big)} \\ &= \frac{(L_d)^{\rho}}{\psi_{sd}(1, c_d)\,\bar{D}_{sd}} \frac{1 - \left(\frac{\phi_{sd}^{*,n=0}(c_d)}{\phi_{sd}^{*,1}(c_d)}\right)^{(\sigma-1)(1-\beta)/\beta}}{1-\beta} \frac{1}{1 - \left(\frac{\phi_{sd}^{*,n=0}(c_d)}{\phi_{sd}^{*,1}(c_d)}\right)^{(\sigma-1)/\beta}}, \end{aligned} \tag{S.22}$$

where the latter equality follows from (S.17), (S.18) and (S.21) under the condition that $\phi_{sd}^{*,1}(c_d) \geq \phi_{sd}^{*,n=0}(c_d)$. Restating (S.20), the zero-consumer threshold for productivity is

$$\phi_{sd}^{*,n=0}(c_d)^{\sigma-1} \equiv \frac{(L_d)^{\rho}}{\psi_{sd}(1, c_d)\,\bar{D}_{sd}}\, \bar{H}(1)^{\sigma-1}$$

for a firm's first product ($G_{sd} = 1$), where $\bar{H}(1) = 1$. Using the latter condition in (S.22), it follows that $\phi_{sd}^{*,1}(c_d) \geq \phi_{sd}^{*,n=0}(c_d)$ need not hold for $\beta < 1$ or $\beta > 1$.

[52] The optimal market penetration rate $n_{sd}(\phi, c_d)$ therefore strictly decreases in c_d by (S.21) for $\sigma > 1$ and $\beta > 0$.

elasticity, while $n_{sd}(\phi, c_d)$ strictly decreases in c_d with a non-unitary elasticity by (S.21), thus lowering F_{sd} with a non-unitary elasticity. The net effect of a c_d shock on a firm's realized $F_{sd}(G_{sd}, c_d; n_{sd})$ is therefore ambiguous.

S2.5 Implications for estimation

A firm's optimal market penetration rate $n_{sd}(\phi, c_d)$ for its product line shifts the product-invariant part of incremental market access costs $c_d f_{sd}(n_{sd})$. Our estimator flexibly allows for a firm-destination specific market access cost shock c_d, which also shifts the product-invariant part of incremental market access costs $c_d f_{sd}(n_{sd})$. Our estimator therefore subsumes within the stochastic market access cost parameter c_d any firm-destination specific variation in the market penetration rate, and fully accounts for the possibility that firms optimally set their market penetration rate n_{sd}.

Similarly, our estimator flexibly allows for a non-unitary firm-destination specific average product appeal shock $\bar{\xi}_{sd\cdot} = \sum_{g=1}^{G_{sd}} \xi_{sdg}/G_{sd}$. Our estimator therefore subsumes within the average product appeal shock $\bar{\xi}_{sd\cdot}$ any firm-destination specific variation in the revenue shifter $D_{sd}(n_{sd}) = n_{sd}\bar{D}_{sd}$, and fully accounts for the possibility that firms optimally set their market penetration rate n_{sd}.

In summary, our existing estimation framework in AGM flexibly allows for the consequences of market penetration costs as in Arkolakis (2010).

S3 Export Products and Export Destinations

Tables S.1 and S.2 show Brazil's top ten export destinations by number of exporters and the top ten exported HS 6-digit product codes by total value in 2000. In Table S.1, Argentina is the most common export destination and the United States receives most Brazilian exports in value. In Table S.2, medium-sized aircraft is the leading export product in value, followed by wood pulp and biofuel products.

Table S.1: **Top Brazilian Export Destinations**

Destination	# Exporters	Export Value (USD)
Argentina	4,590	5,472,333,618
Uruguay	3,251	504,642,201
USA	3,083	9,772,577,557
Chile	2,342	1,145,161,210
Paraguay	2,319	561,065,104
Bolivia	1,799	282,543,791
Mexico	1,336	1,554,452,204
Venezuela	1,333	658,281,591
Germany	1,217	1,364,610,059
Peru	1,191	329,896,577

Source: SECEX 2000, manufacturing firms and their manufactured products.

Table S.2: **Top Brazilian Exported Items**

HS Code	Description	Value (USD)
880230	Airplanes between 2 and 15 tons	2,618,856,983
470329	Bleached non-coniferous chemical wood pulp	1,523,403,942
230400	Soybean oil-cake and other solid residues	1,245,752,048
870323	Passenger vehicles between 1,500 and 3,000 cc	1,197,222,859
852520	Transmission apparatus incorporating reception apparatus	926,618,451
640399	Footwear, with outer soles	854,950,667
720712	Semifinished products of iron or nonalloy steel	802,801,270
760110	Unwrought aluminum, not alloyed	765,195,563
200911	Orange juice, frozen	561,103,666
170111	Raw solid cane sugar	520,544,094

Source: SECEX 2000, manufacturing firms and their manufactured products.

S4 Monte Carlo Simulations

To document identification under our simulated method of moments estimator, we run Monte Carlo tests with generated data. We generate 333,000 Brazilian firms under the initial parameters Θ, where

$$\Theta = \left\{\delta_1, \delta_2, , \tilde{\alpha}, \tilde{\theta}, \sigma_\xi, \sigma_c\right\} = \{-1.17, -0.90, 1.73, 1.84, 1.89, 0.53\}$$

reflects the baseline estimates from Table 2 in the main text. The generated data have approximately 10,000 exporters. We then apply our simulated method of moments routine to the generated data and find the optimum, recovering an estimate of the parameter vector $\hat{\Theta}$. We repeat the data generation and estimation procedure 30 times and report in Table S.3 the mean and standard deviation of the elements of $\hat{\Theta}$.

The Monte Carlo results in Table S.3 document that our procedure accurately pinpoints all parameters of interest. In particular $(\hat{\tilde{\alpha}}, \hat{\sigma}_\xi, \hat{\sigma}_c)$ are precisely estimated, with point estimates close to the initial parameters behind the generated data and with standard errors less than 2 percent of the true value. Similarly, $\hat{\delta}$ and $\hat{\tilde{\theta}}$ are estimated close to their true values, their standard errors are under 4 percent of their true values. The proximity of our parameter estimates to the initial parameters underlying the data generation, and their precision, substantiate the hypothesis that our simulated method of moments estimator identifies the AGM model's parameters of interest.

S5 Sensitivity Analysis

To assess the robustness of our baseline estimates in AGM, we perform a number of modifications to our main specification. Overall, we find that our baseline results are

Table S.3: **Monte Carlo Results**

Θ	δ_1	δ_2	$\tilde{\alpha}$	$\tilde{\theta}$	σ_ξ	σ_c	$\delta_1 - \delta_2$
Parameter of generated data	-1.17	-0.90	1.73	1.84	1.89	0.53	-0.27
Estimate (mean)	-1.21	-0.93	1.71	1.91	1.88	0.51	-0.28
(s.e.)	(0.03)	(0.03)	(0.02)	(0.05)	(0.03)	(0.01)	(0.01)

remarkably robust to sample restrictions and alternative variable definitions.

S5.1 Adjusted sales

Our baseline estimates imply a large and statistically significant difference between δ_{LAC} and δ_{ROW}. In a first robustness check, we strive to rule out that this difference could be driven by different typical sales across sets of products that Brazilian firms ship to LAC and non-LAC countries. We therefore correct sales and control for the mean sales of product groups at the HS 2-digit level. Concretely, we take the upward or downward deviation of a firm's HS 6-digit product sales to a destination $\log y^p_{\omega d}$ from the worldwide product-group sales mean of Brazilian exporters:

$$\tilde{y}_{\omega dg} = \exp\left\{\log y_{\omega dg} - \frac{1}{M}\sum_{\omega}\frac{1}{N}\sum_{d}\sum_{g\in \text{HS } 2}\log y_{\omega dg}\right\}.$$

This adjustment does not reduce the sample size. We report the results in the row *1. Adjusted sales* of Table S.4. The estimates are broadly consistent with the baseline, but the estimated scope elasticities of market access costs δ and of product efficiency $\tilde{\alpha}$ are lower in absolute magnitude, and so is the estimated Pareto shape parameter $\tilde{\theta}$. These estimates imply that both the within-firm product distribution is more concentrated in the top product and the between-firm sales distribution has more firms in the upper tail with extremely high sales. An intuitive explanation is that demeaning sales by the average exporter's typical sale in a product group exacerbates sales deviations of specific products, thus making distributions appear more extreme. However, signs of our estimates stay the same and broad magnitudes remain qualitatively similar to the baseline estimates.

S5.2 Advanced manufacturing

A related robustness concern is that estimation could be driven by different feasible exporter scopes across product groups that Brazilian firms ship to LAC and non-LAC destination. For example, more differentiated industries, or more technology driven industries, might allow for the export of more potential varieties, or the HS classification system might simply provide more individual HS 6-digit products within more refined HS 2-digit product groups. In a second robustness exercise we therefore split the sample into firms that are active in relatively "advanced manufacturing" industries. We present results from our definition of advanced manufacturing as three top-level NACE sectors "Manufacture of machinery and equipment", "Manufacture of electrical and optical equipment" and "Manufacture of transport equipment" (codes DK, DL and DM).

Under this sectoral restriction, a markedly reduced sample size of only 2,539 Brazilian manufacturing exporters remains. Despite the considerable drop in sample size, however, results in the row labeled *2. Advanced manufacturing* in Table S.4 are broadly consistent with the baseline. In advanced manufacturing industries, the difference in scope elasticities of market access costs δ is larger between LAC and non-LAC countries. While market access costs drop off even faster with scope in LAC countries in advanced manufacturing industries, in non-LAC destination the converse is the case: market access costs remain more elevated at higher scopes than in the average manufacturing industry. The gap between the LAC and non-LAC scope elasticities of market access costs is almost double as wide in advanced manufacturing industries as it is in the average industry. For our counterfactual exercise, this wider difference of market access cost elasticities between LAC and non-LAC countries implies even more pronounced benefits of harmonizing market access costs across the world. Other parameter estimates are similar to the baseline, and specially the scope elasticity of product efficiency $\tilde{\alpha}$ is not statistically significantly different in advanced industries compared to the average industry. Overall, every sign remains the same and estimates that are statistically different from the baseline remain comparable in their qualitative economic implications.

S5.3 Eight-digit NCM product categories

To make our results closely comparable to evidence from other countries, in our main text we define a product as a Harmonized System (HS) 6-digit code, which is internationally comparable by requirement of the World Customs Organization (WCO) across its 200 member countries. To query the potential sensitivity of our results to a refined product classification, we use the Mercosur 8-digit level (*Nomenclatura Comum do Mercosul* NCM8), which roughly corresponds to the 8-digit HS level by the World Customs Organization. As the row *3. NCM 8-digit manufacturing* in Table S.4 shows, our results are hardly sensitive at all to the change in level of disaggregation. No single estimate is statistically different from our baseline estimates.

S5.4 Sensitivity to destinations Argentina and United States

Two destination markets dominate Brazilian manufacturing exports: Argentina (the top destination in terms of exporter counts) and the United States (the top destination in terms of export value). To assure ourselves that the estimates are not driven by potential outlier behavior of export flows to those two destinations, we remove them from the sample. As Table S.4 in the row labeled *4. Dropping ARG, USA* shows, only the estimate of the firm size distribution's shape parameter $\tilde{\theta}$ becomes statistically significantly different from the baseline estimate. All other estimates are statistically indistinguishable from the baseline estimates. When we omit Argentina and the United States, the higher estimate for the

Table S.4: **Robustness for Select Subsamples**

estimate (s.e.)	δ_{LAC}	δ_{ROW}	$\tilde{\alpha}$	$\tilde{\theta}$	σ_{ξ}	σ_{c}	$\delta_{\text{LAC}}-\delta_{\text{ROW}}$
Baseline	-1.17	-0.90	1.73	1.84	1.89	0.53	-0.27
(all manufacturing)	(0.09)	(0.11)	(0.09)	(0.09)	(0.04)	(0.03)	(0.05)
1. Adjusted	-0.89	-0.70	1.53	1.63	1.92	0.60	-0.18
sales	(0.06)	(0.06)	(0.06)	(0.07)	(0.03)	(0.02)	(0.03)
2. Advanced	-1.27	-0.78	1.77	1.64	2.01	0.58	-0.50
manufacturing	(0.14)	(0.17)	(0.12)	(0.23)	(0.08)	(0.06)	(0.13)
3. NCM 8-digit	-1.16	-0.84	1.72	1.76	1.82	0.55	-0.32
manufacturing	(0.06)	(0.08)	(0.05)	(0.10)	(0.04)	(0.03)	(0.05)
4. Dropping ARG, USA	-1.22	-0.86	1.75	2.04	1.73	0.55	-0.36
(all manufacturing)	(0.08)	(0.07)	(0.07)	(0.09)	(0.04)	(0.02)	(0.07)
5. Exporter share	-1.43	-1.19	1.98	1.86	1.81	0.54	-0.24
10 percent	(0.11)	(0.14)	(0.10)	(0.11)	(0.04)	(0.03)	(0.05)

Source: SECEX 2000, manufacturing firms and their manufactured products.
Note: Products at the HS 6-digit level. Estimates of δ_{LAC} indicates the scope elasticity for incremental product access costs for Brazilian firms shipping to other LAC destinations. Similarly δ_{ROW} perform the same role for exports to non-LAC destinations. See text for full description of various specifications.

Pareto tail index $\tilde{\theta}$ implies a lower probability mass in the upper tail of firms with extremely high sales. Even though Argentina and the United States attract a large number of export entrants from Brazil, these markets also exhibit a stronger concentration of exports among just a few top-selling firms than the average Brazilian export destination.

S5.5 Exporter share

An arguably important moment for our simulated method of moments is the share of formally established Brazilian manufacturing firms that export. Among the universe of Brazilian firms with at least one employee, only three percent of firms are exporters in 2000. This share is similar to that observed in other countries, for which data on the universe of firms with at least one employee is available. However, censuses and surveys in most developing and some industrialized countries truncate their target population of firms from below with thresholds up to 20 employees. To query sensitivity of our estimates to the share of exporters, we hypothetically consider an alternative share of 10 percent of Brazilian firms exporting. This exercise serves two purposes. First, comparisons of our findings to future results in other countries may depend on using a hypothetically truncated target population of firms from below. Second, we can check how our results might depend on a hypothetically more export oriented manufacturing sector such as, for instance, the U.S. manufacturing sector.

Table S.4 reports the results in the row *5. Exporter share 10 percent*. Compared to the baseline, the scope elasticities of market access costs δ and of product efficiency $\tilde{\alpha}$ increase

Table S.5: **Alternative Regional Aggregates**

δ_1	δ_2	$\delta_1 - \delta_2$
Baseline Latin America and Caribbean (LAC)	Baseline Rest of World	-0.27 (0.05)
1. Mercosur	Rest of LAC (Non-Mercosur)	-0.03 (0.03)
2. Mercosur	Rest of World (Non-Mercosur)	-0.18 (0.03)

Source: SECEX 2000, manufacturing firms and their manufactured products.
Note: Products at the HS 6-digit level.

in absolute magnitude. Intuitively, the estimator tries to "explain" the hypothetically higher share of exporters with relatively faster declines in market access costs as exporter scope increases but to offset those access cost reductions with relatively steeper declines in product efficiency away from core competency, so as to keep matching the overall pattern of exporter scopes across destinations. The other three parameter estimates remain similar to the baseline estimates. This final robustness exercise therefore clarifies how the firm entry margin influences identification: if firm entry with the first product were hypothetically more prevalent, then for a given common market access cost component $f_{sd}(1)$ the access cost schedule would need to decline faster with scope, leading to wider exporters scopes everywhere, unless production efficiency also declines faster with scope.

S5.6 Sensitivity to Mercosur

In a final set of robustness exercises, we alternate the pairings of regional aggregates. In the baseline, we split the world into LAC (Latin American and the Caribbean) and the Rest of the World (non-LAC). In a first alteration, we drop destination countries outside of LAC from our sample and split LAC into Mercosur destinations in 2000 (Argentina, Paraguay, Uruguay) and non-Mercosur destinations. In the row labelled *1. Mercosur–Rest of LAC*, Table S.5 reports the results for the difference in the scope elasticities of market access costs between the two sub-regions within LAC, and the difference is negative as in the baseline but small (and not statistically different from zero). This finding justifies our treatment of LAC in the baseline as a relatively homogeneous region for Brazilian exporters. In a second alteration of the regional split, we discern between Mercosur destinations in 2000 and the Rest of the World, where the Rest of the World includes LAC countries outside Mercosur as well as non-LAC destinations. Expectedly, given the earlier results in the baseline and in the first alteration, the difference is negative but not quite as pronounced in magnitude as the difference between LAC and the Rest of the World. We therefore conclude that LAC countries outside Mercosur are more similar to Mercosur than to the Rest of the World and consider our baseline split of destinations into LAC and non-LAC an adequate country grouping.

Appendix to the Online Supplement

S-A Optimal Product Prices

We characterize the first-order conditions for the firm's optimal pricing rules at every destination d. There are $G_{sd}(\phi)$ first-order conditions with respect to p_{sdg}. For any $G_{sd}(\phi)$, taking the first derivative of profits $\pi_{sd}(\phi)$ from (S.1) with respect to p_{sdg} and dividing by $p_{sdg}^{-\varepsilon} P_{sd}(\phi; G_{sd})^{\varepsilon-\sigma} P_d^{\sigma-1} T_d$ yields

$$\begin{aligned}\frac{\partial \pi_{sd}(\phi)}{\partial p_{sdg}} &= P_d^{\sigma-1} T_d \cdot P_{sd}(\phi; G_{sd})^{\varepsilon-\sigma}\, p_{sdg}^{-\varepsilon} \Bigg\{ 1 - \varepsilon \left(1 - \frac{w_s}{\phi/h(g)}\, \tau_{sd}\, p_{sdg}^{-1} \right) \\ &\quad + (\varepsilon-\sigma) P_{sd}(\phi; G_{sd})^{\varepsilon-1} \sum_{k=1}^{G_{sd}(\phi)} \left(p_{sdk} - \frac{w_s}{\phi/h(k)}\, \tau_{sd} \right) p_{sdk}^{-\varepsilon} \Bigg\}. \end{aligned} \tag{S-A.1}$$

The first-order conditions require that (S-A.1) is equal to zero for all products $g = 1, \ldots, G_{sd}(\phi)$. Use the first-order conditions for any two products g and g' and reformulate to find

$$p_{sdg}/p_{sdg'} = h(g)/h(g').$$

So the firm must optimally charge an identical markup over the marginal costs for all products. Define this optimal markup as $\bar{m}$. To solve out for $\bar{m}$ in terms of primitives, use $p_{sdg} = \bar{m}\, w_s \tau_{sd}/[\phi/h(g)]$ in the first-order condition above and simplify:

$$1 - \varepsilon \frac{1}{\bar{m}} + (\varepsilon-\sigma) P_{sd}(\phi; G_{sd})^{\varepsilon-1}\, \frac{\bar{m}-1}{\bar{m}} \sum_{k=1}^{G_{sd}(\phi)} p_{sdk}^{-(\varepsilon-1)} = 0.$$

Note that $\sum_{k=1}^{G_{sd}(\phi)} p_{sdk}^{-(\varepsilon-1)} = P_{sd}(\phi; G_{sd})^{-(\varepsilon-1)}$. Solving the first-order condition for $\bar{m}$, we find the optimal markup over each product g's marginal cost

$$\bar{m} = \tilde{\sigma} \equiv \sigma/(\sigma-1).$$

A firm with productivity ϕ optimally charges a price

$$p_{sdg}(\phi) = \tilde{\sigma}\, w_s \tau_{sd}/[\phi/h(g)] \tag{S-A.2}$$

for its products $g = 1, ..., G_{sd}(\phi)$.

S-B Second-order Conditions

We now turn to the second-order conditions for price choice. To find the entries along the diagonal of the Hessian matrix, take the first derivative of condition (S-A.1) with respect to the own price p_{sdg} and then replace $w_s \tau_{sd}/[\phi/h(g)] = p_{sdg}(\phi)/\tilde{\sigma}$ by the first-order condition to find

$$
\begin{aligned}
\frac{\partial^2 \pi_{sd}(\phi)}{(\partial p_{sdg})^2} &= P_d^{\sigma-1}\, T_d \cdot P_{sd}\,(\phi; G_{sd})^{\varepsilon-\sigma}\; p_{sdg}^{-\varepsilon} \Big\{ -\frac{\varepsilon}{\tilde{\sigma}} p_{sdg}^{-1} \\
&\qquad + (\varepsilon-\sigma) P_{sd}\,(\phi; G_{sd})^{\varepsilon-1} \left[-(\varepsilon-1) + \varepsilon/\tilde{\sigma}\right] p_{sdg}^{-\varepsilon} \\
&\qquad + (\varepsilon-\sigma)(\varepsilon-1) P_{sd}\,(\phi; G_{sd})^{2(\varepsilon-1)}\, p_{sdg}^{-\varepsilon} \cdot \sum_{k=1}^{G_{sd}} (1 - 1/\tilde{\sigma}) p_{sdk}^{-(\varepsilon-1)} \Big\} \\
&= P_{sd}\,(\phi; G_{sd})^{\varepsilon-\sigma}\, P_d^{\sigma-1}\, T_d \cdot \Big\{ -\varepsilon p_{sdg}^{-\varepsilon-1} + (\varepsilon-\sigma) P_{sd}\,(\phi; G_{sd})^{\varepsilon-1}\, p_{sdg}^{-2\varepsilon} \Big\} / \tilde{\sigma}.
\end{aligned} \tag{S-B.3}
$$

This term is strictly negative if and only if

$$
(\varepsilon-\sigma) P_{sd}\,(\phi; G_{sd})^{\varepsilon-1}\, p_{sdg}^{-(\varepsilon-1)} < \varepsilon.
$$

If $\varepsilon \leq \sigma$, this last condition is satisfied because the left-hand side is weakly negative and $\varepsilon > 0$. If $\varepsilon > \sigma$, then we can rewrite the condition as $p_{sdg}^{-(\varepsilon-1)} / [\sum_{k=1}^{G_{sd}} p_{sdg}^{-(\varepsilon-1)}] < 1 < \varepsilon/(\varepsilon-\sigma)$ so that the condition is satisfied. The diagonal entries of the Hessian matrix are therefore strictly negative for any demand elasticity configuration across nests.

To derive the entries off the diagonal of the Hessian matrix, we take the derivative of condition (S-A.1) for product g with respect to any other price $p_{sdg'}$ and then replace $w_s \tau_{sd} / [\phi / h(g')] = p_{sdg'}(\phi)/\tilde{\sigma}$ by the first-order condition to find

$$
\begin{aligned}
\frac{\partial^2 \pi_{sd}(\phi)}{\partial p_{sdg}\, \partial p_{sdg'}} &= P_d^{\sigma-1}\, T_d \cdot P_{sd}\,(\phi; G_{sd})^{\varepsilon-\sigma}\; p_{sdg}^{-\varepsilon} \Big\{ (\varepsilon-\sigma) P_{sd}\,(\phi; G_{sd})^{\varepsilon-1} \left[-(\varepsilon-1) + \varepsilon/\tilde{\sigma}\right] p_{sdg'}^{-\varepsilon} \\
&\qquad + (\varepsilon-\sigma)(\varepsilon-1) P_{sd}\,(\phi; G_{sd})^{2(\varepsilon-1)}\, p_{sdg'}^{-\varepsilon} \sum_{k=1}^{G_{sd}} (1 - 1/\tilde{\sigma}) p_{sdk}^{-(\varepsilon-1)} \Big\} \\
&= P_d^{\sigma-1}\, T_d \cdot (\varepsilon-\sigma) P_{sd}\,(\phi; G_{sd})^{\varepsilon-\sigma+\varepsilon-1}\; p_{sdg}^{-\varepsilon} p_{sdg'}^{-\varepsilon} / \tilde{\sigma}.
\end{aligned} \tag{S-B.4}
$$

This term is strictly positive if and only if $\varepsilon > \sigma$.

Having derived the entries of the Hessian matrix, it remains to establish the conditions under which the Hessian is negative definite. We discern two cases. First the case of $\varepsilon \leq \sigma$ and then the case $\varepsilon > \sigma$.

S-B.1 Negative definiteness of Hessian if $\varepsilon \leq \sigma$

By (S-B.3) and (S-B.4), the Hessian matrix can be written as

$$
\mathbf{H} = P_{sd}\,(\phi; G_{sd})^{\varepsilon-\sigma}\, P_d^{\sigma-1}\, T_d \left[\mathbf{H}_A + (\varepsilon-\sigma) P_{sd}\,(\phi; G_{sd})^{\varepsilon-1}\, \mathbf{H}_B \right],
$$

where

$$\mathbf{H}_A \equiv \begin{pmatrix} -\varepsilon p_{sd1}^{-\varepsilon-1} & & & \\ 0 & -\varepsilon p_{sd2}^{-\varepsilon-1} & & \\ 0 & 0 & -\varepsilon p_{sd3}^{-\varepsilon-1} & \\ \dots & & & \dots \end{pmatrix} \text{ and } \mathbf{H}_B \equiv \begin{pmatrix} p_{sd1}^{-\varepsilon} p_{sd1}^{-\varepsilon} & & & \\ p_{sd2}^{-\varepsilon} p_{sd1}^{-\varepsilon} & p_{sd2}^{-\varepsilon} p_{sd2}^{-\varepsilon} & & \\ p_{sd3}^{-\varepsilon} p_{sd1}^{-\varepsilon} & p_{sd3}^{-\varepsilon} p_{sd2}^{-\varepsilon} & p_{sd3}^{-\varepsilon} p_{sd3}^{-\varepsilon} & \\ \dots & & & \dots \end{pmatrix}.$$

The Hessian matrix $\mathbf{H}$ is negative definite if and only if the negative Hessian

$$-\mathbf{H} = P_{sd}\left(\phi; G_{sd}\right)^{\varepsilon-\sigma} P_d^{\sigma-1} T_d \left[-\mathbf{H}_A + (\sigma-\varepsilon) P_{sd}\left(\phi; G_{sd}\right)^{\varepsilon-1} \mathbf{H}_B\right]$$

is positive definite. Note that the sum of one positive definite matrix and any number of positive semidefinite matrices is positive definite. Hence if $-\mathbf{H}_A$ and $\mathbf{H}_B$ are positive semidefinite and at least one of the two matrices is positive definite (given $\varepsilon \leq \sigma$), then the Hessian is negative definite.

A necessary and sufficient condition for a matrix to be positive definite is that the leading principal minors of the matrix are positive. The leading principal minors of $-\mathbf{H}_A$ are positive, so $-\mathbf{H}_A$ is positive definite. For $\mathbf{H}_B$, the first leading principal minor is positive, and all remaining principal minors are equal to zero. So $\mathbf{H}_B$ is positive semidefinite. Therefore the Hessian matrix $\mathbf{H}$ is negative definite.

S-B.2 Negative definiteness of Hessian if $\varepsilon > \sigma$

Another necessary and sufficient condition for the Hessian matrix $\mathbf{H}$ to be negative definite is that the leading principal minors alternate sign, with the first principal minor being negative. The first diagonal entry is strictly negative as is any diagonal entry by (S-B.3). An application of the leading principal minor test in our case requires a recursive computation of the determinants of $G_{sd}(\phi)$ submatrices (a solution of polynomials with order up to $G_{sd}(\phi)$). We choose to check for negative definiteness of the Hessian in two separate ways when $\varepsilon > \sigma$. First, we derive a sufficient (but not necessary) condition for negative definiteness of the Hessian and query its empirical validity. Second, we present a necessary (but not sufficient) condition for negative definiteness of the Hessian for any pair of two products.

Sufficiency. A sufficient condition for the Hessian to be negative definite is due to McKenzie (1960): a symmetric diagonally dominant matrix with strictly negative diagonal entries is negative definite. A matrix is diagonally dominant if, in every row, the absolute value of the diagonal entry strictly exceeds the sum of the absolute values of all off-diagonal entries. By our derivations above, all diagonal entries of the Hessian are strictly negative.

For $\varepsilon > \sigma$, the condition for the Hessian to be diagonally dominant is

$$\sum_{k \neq g}^{G_{sd}} (\varepsilon-\sigma) P_{sd}\left(\phi; G_{sd}\right)^{\varepsilon-1} p_{sdk}^{-\varepsilon} p_{sdg}^{-\varepsilon} \;<\; \varepsilon\, p_{sdg}^{-\varepsilon-1} - (\varepsilon-\sigma) P_{sd}\left(\phi; G_{sd}\right)^{\varepsilon-1} p_{sdg}^{-2\varepsilon}$$

for all of a firm ϕ's products (rows of its Hessian), where we cancelled the strictly positive terms $P_d^{\sigma-1} T_d P_{sd}(\phi; G_{sd})^{\varepsilon-\sigma} / \tilde{\sigma}$ from the inequality.

Using the optimal price (S-A.2) of product g from the first-order condition and rearranging terms yields the following condition

$$\frac{\sum_{k=1}^{G_{sd}} h(k)^{-\varepsilon}}{\sum_{k=1}^{G_{sd}} h(k)^{-(\varepsilon-1)}} \quad < \quad \frac{\varepsilon}{\varepsilon-\sigma}\, h(g)^{-1} \tag{S-B.5}$$

for the Hessian to be a diagonally dominant matrix at the optimum.

By convention and without loss of generality $h(1) = 1$ for a firm with productivity ϕ. So the product efficiency schedule $h(g)$ strictly exceeds unity for the second product and subsequent products. As a result, the left-hand side of the inequality is bounded above for an exporter with a scope of at least two products at a destination:

$$\frac{\sum_{k=1}^{G_{sd}} h(k)^{-\varepsilon}}{\sum_{k=1}^{G_{sd}} h(k)^{-(\varepsilon-1)}} < \frac{\sum_{k=1}^{G_{sd}} h(k)^{-(\varepsilon-1)}}{\sum_{k=1}^{G_{sd}} h(k)^{-(\varepsilon-1)}} = 1.$$

A sufficient (but not necessary) condition for the Hessian to be negative definite is therefore

$$1 \leq h(g) < \frac{\varepsilon}{\varepsilon-\sigma}$$

for all of the firm's products. However, the Hessian can still be negative definite even if this condition fails. Clearly, the Hessian becomes negative definite the closer is ε to σ because then the off-diagonal entries approach zero and the Hessian is trivially diagonally dominant. Moreover, the Hessian can be negative definite even if it is not a diagonally dominant matrix.

To query the empirical validity of the sufficient condition $h(g) < \varepsilon/(\varepsilon-\sigma)$, consider evidence on products and brands in Broda and Weinstein (2006). Their preferred estimates for ε and σ within and across domestic U.S. brand modules are 11.5 and 7.5. Estimates in AGM suggest that $\alpha(\varepsilon-1)$ is around 1.84 under the specification that $h(g) = g^{\alpha}$. These parameters imply that the condition $h(g) < \varepsilon/(\varepsilon-\sigma)$ is satisfied for Hessians with up to 414 products. In the AGM data, no firm-country observations involve 415 or more products in a market (with a median of one product and a mean of 3.52). Even if additional products individually violate the sufficient condition, Hessians with more products may still be negative definite.

Necessity. Consider any two products g and g'. Negative definiteness of the Hessian must be independent of the ordering of products, so these two products can be assigned the first and second row in the Hessian without loss of generality. As stated before, a necessary and sufficient condition for the Hessian to be negative definite is that the leading principal minors of the Hessian alternate sign, with the first principal minor being negative. A necessary condition for the Hessian to be negative definite is therefore that the principal minors of any two products (first and second in the Hessian) alternate sign, with the first principal minor negative and the second positive.

The first principle minor is strictly negative because all diagonal entries are strictly negative by (S-B.3). The second principal minor is strictly positive if and only if the determinant satisfies

$$2\,\varepsilon^2 P_{sd}\left(\phi;G_{sd}\right)^{-(\varepsilon-1)}-\varepsilon(\varepsilon-\sigma)\left(p_{sdg}^{-(\varepsilon-1)}+p_{sdg'}^{-(\varepsilon-1)}\right)-(\varepsilon-\sigma)^2\left(p_{sdg}p_{sdg'}\right)^{-(\varepsilon-1)}P_{sd}\left(\phi;G_{sd}\right)^{\varepsilon-1}>0, \tag{S-B.6}$$

where we cancelled the strictly positive terms $P_d^{2(\sigma-1)}T_d P_{sd}\left(\phi;G_{sd}\right)^{2(\varepsilon-\sigma)}/\tilde{\sigma}^2$ from the inequality and multiplied both sides by $p_{sdg}^{\varepsilon-1}p_{sdg'}^{\varepsilon-1}P_{sd}\left(\phi;G_{sd}\right)^{-(\varepsilon-1)}$.

To build intuition, consider the dual-product case with $G_{sd}(\phi)=2$. Then condition (S-B.6) simplifies to

$$\frac{h(g)^{-(\varepsilon-1)}}{\sum_{k=1}^{G_{sd}}h(k)^{-(\varepsilon-1)}}\cdot\frac{h(g')^{-(\varepsilon-1)}}{\sum_{k=1}^{G_{sd}}h(k)^{-(\varepsilon-1)}}<\frac{\varepsilon}{\varepsilon-\sigma}\frac{\varepsilon+\sigma}{\varepsilon-\sigma}.$$

For $\varepsilon>\sigma$, both terms in the product on the right-hand side strictly exceed unity while the terms in the product on the left-hand side are strictly less than one, and the condition is satisfied.

In the multi-product case with $G_{sd}(\phi)>2$, replace $p_{sdg}^{-(\varepsilon-1)}+p_{sdg'}^{-(\varepsilon-1)}=P_{sd}\left(\phi;G_{sd}\right)^{-(\varepsilon-1)}-\sum_{k\neq g,g'}p_{sdk}^{-(\varepsilon-1)}$ in condition (S-B.6) and simplify to find

$$\frac{h(g)^{-(\varepsilon-1)}}{\sum_{k=1}^{G_{sd}}h(k)^{-(\varepsilon-1)}}\cdot\frac{h(g')^{-(\varepsilon-1)}}{\sum_{k=1}^{G_{sd}}h(k)^{-(\varepsilon-1)}}<\frac{\varepsilon}{\varepsilon-\sigma}\frac{\varepsilon+\sigma}{\varepsilon-\sigma}+\frac{\varepsilon}{\varepsilon-\sigma}\frac{\sum_{k\neq g,g'}h(k)^{-(\varepsilon-1)}}{\sum_{k=1}^{G_{sd}}h(k)^{-(\varepsilon-1)}}.$$

For $\varepsilon>\sigma$, the necessary condition on any two products of a multi-product firm is trivially satisfied by the above derivations because the additional additive term on the right-hand side is strictly positive.

In summary, parameters of our model are such that, for any two products of a multi-product firm, the second-order condition is satisfied.

S-C Proof of Proposition S.3

Average sales from s to d are

$$\bar{T}_{sd}=\int_{\phi_{sd}^*}y_{sd}(G_{sd})\frac{\theta\left(\phi_{sd}^*\right)^{\theta}}{\phi^{\theta+1}}\,\mathrm{d}\phi=\sigma f_{sd}(1)\theta\int_{\phi_{sd}^*}\frac{\phi^{\sigma-2-\theta}/\left(\phi_{sd}^*\right)^{\sigma-1-\theta}}{H\left(G_{sd}(\phi)\right)^{\sigma-1}}\,\mathrm{d}\phi.$$

The proof of the proposition follows from the following Lemma.

Lemma 2 *Suppose Assumptions S.1, S.2 and S.3 hold. Then*

$$\int_{\phi_{sd}^*}\frac{\phi^{\sigma-2-\theta}/\left(\phi_{sd}^*\right)^{\sigma-1-\theta}}{H\left(G_{sd}(\phi)\right)^{\sigma-1}}\,d\phi=\frac{f_{sd}(1)^{\tilde{\theta}-1}}{\theta-(\sigma-1)}\tilde{F}_{sd},$$

where

$$\tilde{F}_{sd} \equiv \sum_{\upsilon=1}^{\infty} \frac{\left[f_{sd}\left(\upsilon\right)\right]^{1-\tilde{\theta}}}{\left[H\left(\upsilon\right)^{1-\sigma} - H\left(\upsilon-1\right)^{1-\sigma}\right]^{-\tilde{\theta}}}.$$

Proof. Note that

$$\begin{aligned}
\int_{\phi_{sd}^{*}} \frac{\phi^{\sigma-2-\theta} / \left(\phi_{sd}^{*}\right)^{\sigma-1-\theta}}{H\left(G_{sd}(\phi)\right)^{\sigma-1}} \mathrm{d}\phi &= H(1)^{1-\sigma} \int_{\phi_{sd}^{*}}^{\phi_{sd}^{*,2}} \phi^{\sigma-2-\theta} \mathrm{d}\phi + H(2)^{1-\sigma} \int_{\phi_{sd}^{*,2}}^{\phi_{sd}^{*,3}} \phi^{\sigma-2-\theta} \mathrm{d}\phi + \ldots \\
&= H(1)^{1-\sigma} \left[\frac{\left(\phi_{sd}^{*,2}\right)^{\sigma-1-\theta} - \left(\phi_{sd}^{*}\right)^{\sigma-1-\theta}}{\left[\theta - (\sigma-1)\right] \left(\phi_{sd}^{*}\right)^{\sigma-1-\theta}} \right] \\
&\quad + H(2)^{1-\sigma} \left[\frac{\left(\phi_{sd}^{*,3}\right)^{\sigma-1-\theta} - \left(\phi_{sd}^{*,2}\right)^{\sigma-1-\theta}}{\left[\theta - (\sigma-1)\right] \left(\phi_{sd}^{*}\right)^{\sigma-1-\theta}} \right] + \ldots .
\end{aligned}$$

Also note that, using equations (S.4) and (S.6), the ratio

$$\left[\left(\phi_{sd}^{*,2}\right)^{\sigma-1-\theta} - \left(\phi_{sd}^{*}\right)^{\sigma-1-\theta}\right] / \left(\phi_{sd}^{*}\right)^{\sigma-1-\theta}$$

can be rewritten as

$$\begin{aligned}
&\frac{\left(\phi_{sd}^{*,G}\right)^{\sigma-1-\theta} - \left(\phi_{sd}^{*,G-1}\right)^{\sigma-1-\theta}}{\left(\phi_{sd}^{*}\right)^{\sigma-1-\theta}} = \\
&= \frac{\left[\frac{\left(\phi_{sd}^{*}\right)^{\sigma-1}}{H(g)^{1-\sigma} - H(g-1)^{1-\sigma}} \frac{f_{sd}(g)}{f_{sd}(1)}\right]^{\frac{\sigma-1-\theta}{\sigma-1}} - \left[\frac{\left(\phi_{sd}^{*}\right)^{\sigma-1}}{H(g-1)^{1-\sigma} - H(g-2)^{1-\sigma}} \frac{f_{sd}(g-1)}{f_{sd}(1)}\right]^{\frac{\sigma-1-\theta}{\sigma-1}}}{\left[\left(\phi_{sd}^{*}\right)^{\sigma-1}\right]^{\frac{\sigma-1-\theta}{\sigma-1}}} \\
&= f_{sd}(1)^{\tilde{\theta}-1} \left\{ \frac{f_{sd}\left(g\right)^{1-\tilde{\theta}}}{\left[H\left(g\right)^{1-\sigma} - H\left(g-1\right)^{1-\sigma}\right]^{1-\tilde{\theta}}} - \frac{f_{sd}\left(g-1\right)^{1-\tilde{\theta}}}{\left[H\left(g-1\right)^{1-\sigma} - H\left(g-2\right)^{1-\sigma}\right]^{1-\tilde{\theta}}} \right\}.
\end{aligned}$$

We define[53]

$$\begin{aligned}
\tilde{F}_{sd} &\equiv \sum_{\upsilon=1} H(\upsilon)^{1-\sigma}\left[\frac{[f_{sd}(\upsilon+1)]^{1-\tilde{\theta}}}{\left[H(\upsilon+1)^{1-\sigma}-H(\upsilon)^{1-\sigma}\right]^{1-\tilde{\theta}}}-\frac{[f_{sd}(\upsilon)]^{1-\tilde{\theta}}}{\left[H(\upsilon)^{1-\sigma}-H(\upsilon-1)^{1-\sigma}\right]^{1-\tilde{\theta}}}\right] \\
&= \sum_{\upsilon=1}\left[\left[H(\upsilon)^{1-\sigma}-H(\upsilon-1)^{1-\sigma}\right]\frac{[f_{sd}(\upsilon)]^{1-\tilde{\theta}}}{\left[H(\upsilon)^{1-\sigma}-H(\upsilon-1)^{1-\sigma}\right]^{1-\tilde{\theta}}}\right] \\
&= \sum_{\upsilon=1}\frac{[f_{sd}(\upsilon)]^{1-\tilde{\theta}}}{\left[H(\upsilon)^{1-\sigma}-H(\upsilon-1)^{1-\sigma}\right]^{-\tilde{\theta}}}.
\end{aligned}$$

With this definition we obtain

$$\int_{\phi^*_{sd}}\frac{\phi^{\sigma-2-\theta}/\left(\phi^*_{sd}\right)^{\sigma-1-\theta}}{H\left(G_{sd}(\phi)\right)^{\sigma-1}}\mathrm{d}\phi=\frac{f_{sd}(1)^{\tilde{\theta}-1}}{\theta-(\sigma-1)}\tilde{F}_{sd}.$$

■

S-D Welfare

We have that

$$\begin{aligned}
P_d^{1-\sigma} &= \sum_s\int_{\phi^*_{sd}}\left[P_{sd}(\phi)\right]^{1-\sigma}\mu(\phi)\mathrm{d}\phi \\
&= \sum_s\int_{\phi^*_{sd}} M_{sd}\left[\sum_{\upsilon=1}^{G_{sd}(\phi)}\left(\tilde{\sigma}\frac{w_s}{\phi/h(g)}\tau_{sd}\right)^{1-\varepsilon}\right]^{\frac{1-\sigma}{1-\varepsilon}}\frac{\theta\left(\phi^*_{sd}\right)^{\theta}}{\phi^{\theta+1}}\mathrm{d}\phi \\
&= \sum_s\left(\tilde{\sigma}w_s\tau_{sd}\right)^{1-\sigma}b_s^{\theta}\theta\left[H(1)^{1-\sigma}\left(\frac{\left(\phi^{*,2}_{sd}\right)^{\sigma-1-\theta}-\left(\phi^{*,1}_{sd}\right)^{\sigma-1-\theta}}{\theta-(\sigma-1)}\right)+\ldots\right] \\
&= \sum_s\left(\tilde{\sigma}w_s\tau_{sd}\right)^{-\theta}b_s^{\theta}\theta\left(\frac{f_{sd}(1)}{\frac{1}{\sigma}T_d}\right)^{1-\tilde{\theta}}\left[H(1)^{1-\sigma}\left(\frac{\left(\phi^{*,2}_{sd}\right)^{\sigma-1-\theta}-\left(\phi^{*,1}_{sd}\right)^{\sigma-1-\theta}}{\left(\phi^{*,1}_{sd}\right)^{\sigma-1-\theta}}\right)+\ldots\right],
\end{aligned}$$

[53] In the special case with $\varepsilon=\sigma$, we can rearrange the terms and find

$$\tilde{F}_{sd}=\sum_{\upsilon=1}^{\infty}\frac{[f_{sd}(\upsilon)]^{1-\tilde{\theta}}}{\left[h(\upsilon)^{\sigma-1}\right]^{-\tilde{\theta}}}=\sum_{\upsilon=1}^{\infty}\frac{[f_{sd}(\upsilon)]^{1-\tilde{\theta}}}{h(\upsilon)^{-\theta}}.$$

where we use the definition of $\phi_{sd}^{*,1}$ for the last step. The final term in parentheses equals $\tilde{F}_{sd}$ so

$$P_d^{-\theta} = \frac{\theta \left(\tilde{\sigma}\right)^{-\theta}}{\left(\frac{1}{\sigma}\right)^{1-\theta/(\sigma-1)} T_d^{1-\tilde{\theta}}} \sum_s b_s^{\theta} \left(w_s \tau_{sd}\right)^{-\theta} \tilde{F}_{sd}.$$

Using this relationship in equation (S.12), we obtain

$$\left(\frac{T_d}{P_d}\right)^{\theta} = \left(\frac{T_d}{w_d}\right)^{\theta} \frac{\theta \left(\tilde{\sigma}\right)^{-\theta}}{\left(\sigma\right)^{\tilde{\theta}-1}} \frac{b_d^{\theta}}{\lambda_{dd}^{-\theta}} \frac{\tilde{F}_{dd}(1)}{T_d^{1-\tilde{\theta}}}.$$

If trade is balanced then $T_d = Y_d$, where T_d is consumption expenditure and Y_d is output. By the definition of $\tilde{F}_{dd}(1)$, this variable is homogeneous of degree $1-\tilde{\theta}$ in wages, and the wage bill share in output $w_d L_d / Y_d$ is constant in all equilibria (see proof below). We therefore arrive at the same welfare expression as in Arkolakis et al. (2012): the share of domestic sales in consumption expenditure λ_{dd} and the coefficient of the Pareto distribution are sufficient statistics to characterize aggregate welfare in the case of balanced trade.

The final step is to verify that the wage w_d is a constant fraction of per-capita output y_d so that the first ratio on the right-hand side is constant. We demonstrate this next.

S-E Constant Wage Share in Output per Capita

We show that the ratio w_d/y_d is a constant number. We first look at the share of fixed costs in bilateral sales. Average fixed costs incurred by firms from s selling to d are

$$\begin{aligned} \bar{F}_{sd} &= \int_{\phi_{sd}^{*}}^{\phi_{sd}^{*,2}} F_{sd}\left(1\right) \theta \frac{\left(\phi_{sd}^{*}\right)^{\theta}}{\phi^{\theta+1}} \mathrm{d}\phi + \int_{\phi_{sd}^{*,2}}^{\phi_{sd}^{*,3}} F_{sd}(2) \theta \frac{\left(\phi_{sd}^{*}\right)^{\theta}}{\phi^{\theta}} \mathrm{d}\phi + \\ &= -F_{sd}(1) \left(\phi_{sd}^{*}\right)^{\theta} \left[\left(\phi_{sd}^{*,2}\right)^{-\theta} - \left(\phi_{sd}^{*}\right)^{-\theta}\right] - F_{sd}(2) \left(\phi_{sd}^{*}\right)^{\theta} \left[\left(\phi_{sd}^{*,3}\right)^{-\theta} - \left(\phi_{sd}^{*,2}\right)^{-\theta}\right] - \ldots \end{aligned}$$

Using the definition $F_{sd}(G_{sd}) \equiv \sum_{g=1}^{G_{sd}} f_{sd}(g)$ and collecting terms with respect to $\phi_{sd}^{*,G}$ we can write the above expression as

$$\bar{F}_{sd} = f_{sd}(1) + \left(\phi_{sd}^{*,2}\right)^{-\theta} \left(\phi_{sd}^{*}\right)^{\theta} f_{sd}(2) + \left(\phi_{sd}^{*,3}\right)^{-\theta} \left(\phi_{sd}^{*}\right)^{\theta} f_{sd}(3) + \ldots .$$

Using the definition of $\phi_{sd}^{*,G}$ from equation (S.6) to replace terms in the above equation, we obtain

$$\left(\phi_{sd}^{*,G}\right)^{\sigma-1} = \frac{\left(\phi_{sd}^{*}\right)^{\sigma-1}}{H(G_{sd})^{-(\sigma-1)} - H\left(G_{sd}-1\right)^{-(\sigma-1)}} \frac{f_{sd}(G_{sd})}{f_{sd}(1)}.$$

Therefore

$$
\begin{aligned}
\bar{F}_{sd} &= f_{sd}(1) + \left(\frac{f_{sd}(2)^{1/(\sigma-1)} \left[H(2)^{-(\sigma-1)} - H(1)^{-(\sigma-1)}\right]^{-1/(\sigma-1)}}{f_{sd}(1)^{1/(\sigma-1)} \left[H(1)^{-(\sigma-1)}\right]^{-1/(\sigma-1)}} \right)^{-\theta} f_{sd}(2) + \dots \\
&= \left[f_{sd}(1) + f_{sd}(1)^{\tilde{\theta}} \left(f_{sd}(2)^{1/(\sigma-1)} \left[H(2)^{-(\sigma-1)} - H(1)^{-(\sigma-1)}\right]^{-1/(\sigma-1)} \right)^{-\theta} f_{sd}(2) + \dots \right] \\
&= \left[f_{sd}(1)\right]^{\tilde{\theta}} \left[f_{sd}(1)^{1-\tilde{\theta}} + \frac{f_{sd}(2)^{1-\tilde{\theta}}}{\left[H(2)^{-(\sigma-1)} - H(1)^{-(\sigma-1)}\right]^{-\tilde{\theta}}} + \frac{f_{sd}(3)^{1-\tilde{\theta}}}{\left[H(3)^{-(\sigma-1)} - H(2)^{-(\sigma-1)}\right]^{-\tilde{\theta}}} \dots \right] \\
&= \left[f_{sd}(1)\right]^{\tilde{\theta}} \tilde{F}_{sd}
\end{aligned}
$$

and hence

$$
\frac{\bar{F}_{sd}}{\bar{T}_{sd}} = \frac{f_{sd}(1)^{\tilde{\theta}} \left[f_{sd}(1)^{1-\tilde{\theta}} + \frac{f_{sd}(2)^{1-\tilde{\theta}}}{h(2)^{\theta}} + \dots \right]}{\frac{f_{sd}(1)^{\tilde{\theta}} \theta\sigma}{\theta-(\sigma-1)} \sum_{g=1}^{\infty} \frac{f_{sd}(g)^{1-\tilde{\theta}}}{h(g)^{\theta}}} = \frac{\theta - (\sigma-1)}{\theta\sigma} \frac{\sum_{g=1}^{\infty} \frac{f_{sd}(g)^{1-\tilde{\theta}}}{h(g)^{\theta}}}{\sum_{g=1}^{\infty} \frac{f_{sd}(g)^{1-\tilde{\theta}}}{h(g)^{\theta}}} < 1.
$$

Finally, the share of profits generated by the corresponding bilateral sales is the share of variable profits in total sales $(1/\sigma)$ minus the average fixed costs paid, as derived above. So

$$
\frac{\bar{\pi}_{sd}}{\bar{T}_{sd}} = \frac{1}{\sigma} - \frac{\theta - (\sigma-1)}{\theta\sigma} = \frac{\sigma-1}{\theta\sigma} = \frac{1}{\tilde{\theta}\sigma} \equiv \eta.
$$

This finding implies that the wage is a constant fraction of per capita income. The reason is that total profits for country s are $\pi_s L_s = \sum_k \lambda_{sk} T_k/(\tilde{\theta}\sigma)$, where $\sum_k \lambda_{sk} T_k$ is the country's total income because total manufacturing sales of a country s equal its total sales across all destinations. So profit income and wage income can be expressed as constant shares of total income:

$$
\pi_s L_s = \frac{1}{\tilde{\theta}\sigma} Y_s \quad \text{and} \quad w_s L_s = \frac{\tilde{\theta}\sigma - 1}{\tilde{\theta}\sigma} Y_s.
$$