Firms and Trade: the New Trade Theory

Carlos Góes
IFC, World Bank Group

November 13, 2025

Preliminaries Consider a world with two countries $i \in \{H, F\}$. The countries are identical in their populations $L_H = L_F$, preferences and production technologies. We will first describe these components of the economy and characterize the autarky equilibrium. We will then explore what happens if a country opens up to trade.

Demand Consumers in country i supply their labor inelastically and earn labor income $w_i L_i$. They have preferences over many goods $\varphi \in \Phi_i$ where $\Phi_i := \{1, 2, \cdots, N\}$ is the set of all goods available in the domestic economy.

Economists use a constant elasticity-of-substitution (CES) utility function to capture preferences in a flexible way. The key parameter, $\sigma > 1$ is the elasticity of substitution: the larger σ is, the more readily consumers switch between varieties when their relative prices change (think "Coke vs. Pepsi" with a high σ). When σ is close to 1, varieties are harder to substitute –each feels almost like a distinct necessity – so consumers tolerate bigger price gaps before adjusting their baskets.

$$\max_{\{q_i(\varphi)\}_{\varphi \in \Phi_i}} Q_i \equiv \left[\sum_{\varphi \in \Phi_i} q_i(\varphi)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \qquad s.t. \qquad P_i Q_i = \sum_{\varphi \in \Phi_i} p_i(\varphi) q_i(\varphi) \leq I_i = w_i L_i$$

First, a note on notation. $q_i(\varphi)$ and $p_i(\varphi)$ are the quantity demanded and price in country i of the good $\varphi \in \{1, 2, \cdots, N\}$. The summation notation simply iterates over the elements of set $\Phi_i := \{1, 2, \cdots, N\}$. For instance, total expenditure of consumers in country i is $\sum_{\varphi \in \Phi_i} p_i(\varphi) q_i(\varphi) = p_i(1)q_i(1) + p_i(2)q_i(2) + \cdots + p_i(N)q_i(N)$. Finally, a word on aggregation. We define Q_i to be the *composite consumption basket* defined as an aggregate of all goods. Implicitly defined is the "price index" P_i which can be seen as the "price" of the composite consumption basket. Intuitively, total expenditure across all goods must equal the cost of the consumption basket $P_i Q_i = \sum_{\varphi \in \Phi_i} p_i(\varphi) q_i(\varphi)$.

How do we solve this problem? This is a constrained maximization problem. The easiest way to approach it is to set up a Lagrangian and take first order conditions:

$$\mathcal{L} = \left[\sum_{\varphi \in \Phi_i} q_i(\varphi)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} + \lambda \left[I_i - \sum_{\varphi \in \Phi_i} p_i(\varphi) q_i(\varphi) \right]$$

where λ is our Lagrange multiplier. And what is λ? Lambda is a job is to keep track of how valuable the last dollar in the consumer's pocket is. Imagine slipping one more dollar into the consumer's wallet. Because the budget set expands by exactly \$1, utility rises by:

$$\frac{\partial \mathcal{L}}{\partial I_i} = \lambda$$

so λ literally measures "extra utils per extra dollar." A high λ says the consumer is desperate for another dollar (the budget is tight); a low λ says an extra dollar hardly moves the needle.

There are N first order conditions in this maximization problem, one for each $q_i(\varphi)$ satisfying:

$$\frac{\partial Q_i}{\partial q_i(\varphi)} - \lambda p_i(\varphi) = 0 \iff \underbrace{\frac{\partial Q_i}{\partial q_i(\varphi)}}_{\substack{\text{marginal benefit} \\ \text{in utils}}} = \underbrace{\lambda p_i(\varphi)}_{\substack{\text{marginal cost} \\ \text{in utils}}} \qquad \text{for each } \varphi \in \Phi_i$$

Note that the result of maximizing utility makes clear the "utils per extra dollar" interpretation of λ. The left hand side states shows the marginal utility from one more unit of good φ. The right hand side shows the utility cost of the dollars needed to buy it (λ utils per dollar $\times$ price). Setting them equal forces every good to deliver the same utility per dollar spent.

Now let us work out $\frac{\partial Q_i}{\partial q_i(\varphi)}$ using the chain rule:

$$\frac{\partial Q_i}{\partial q_i(\varphi)} = \underbrace{\frac{\sigma}{\sigma-1}\left[\sum_{\varphi\in\Phi_i} q_i(\varphi)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}-1}}_{\text{derivative of the outer function}} \times \underbrace{\frac{\sigma-1}{\sigma} q_i(\varphi)^{\frac{\sigma-1}{\sigma}-1}}_{\text{derivative of the inner}} = \lambda p_i(\varphi)$$

$$\underbrace{\left[\sum_{\varphi\in\Phi_i} q_i(\varphi)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{1}{\sigma-1}}}_{=Q_i^{1/\sigma}} \times q_i(\varphi)^{-\frac{1}{\sigma}} = \lambda p_i(\varphi)$$

Solving for $q_i(\varphi)$ we find that:

$$\boxed{q_i(\varphi) = \lambda^{-\sigma} p_i(\varphi)^{-\sigma} Q_i \iff p_i(\varphi) q_i(\varphi) = \lambda^{-\sigma} p_i(\varphi)^{1-\sigma} Q_i}$$

Using the the definition of Q_i allows us to solve for $\lambda^{-\sigma}$:

$$\begin{aligned} Q_i &\equiv \left[\sum_{\varphi\in\Phi_i} q_i(\varphi)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} = \left[\sum_{\varphi\in\Phi_i} \left(\lambda^{-\sigma} p_i(\varphi)^{-\sigma} Q_i\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} = \lambda^{-\sigma} Q_i \left[\sum_{\varphi\in\Phi_i} p_i(\varphi)^{1-\sigma}\right]^{-\frac{\sigma}{1-\sigma}} \\ \iff \lambda^{-\sigma} &= \frac{1}{\left[\sum_{\varphi\in\Phi_i} p_i(\varphi)^{1-\sigma}\right]^{-\frac{\sigma}{1-\sigma}}} \end{aligned}$$

Using the budget constraint allows us to solve for the price level P_i:

$$\begin{aligned}
P_i Q_i &= \sum_{\varphi \in \Phi_i} p_i(\varphi) q_i(\varphi) = \sum_{\varphi \in \Phi_i} \lambda^{-\sigma} p_i(\varphi)^{1-\sigma} Q_i = Q_i \lambda^{-\sigma} \sum_{\varphi \in \Phi_i} p_i(\varphi)^{1-\sigma} \\
\iff P_i &= \frac{1}{\left[\sum_{\varphi \in \Phi_i} p_i(\varphi)^{1-\sigma}\right]^{-\frac{\sigma}{1-\sigma}}} \times \sum_{\varphi \in \Phi_i} p_i(\varphi)^{1-\sigma} = \left[\sum_{\varphi \in \Phi_i} p_i(\varphi)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}
\end{aligned}$$

So $\lambda^{-\sigma} = P_i^{\sigma}$. Finally, using the fact that $Q_i = w_i L_i / P_i$, we can write the demand functions as:

$$\boxed{q_i(\varphi) = \underbrace{\left(\frac{p_i(\varphi)}{P_i}\right)^{-\sigma}}_{\text{relative price}} \times \underbrace{\frac{w_i L_i}{P_i}}_{\text{real income}}}$$

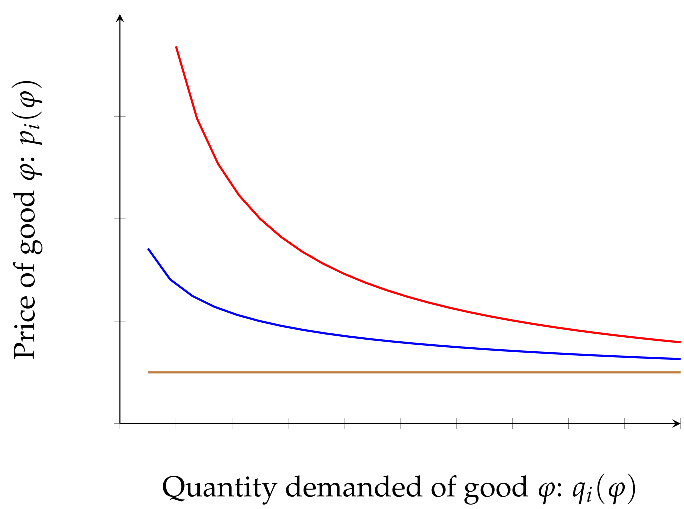

Figure 1: Demand curve with different elasticities: $\sigma = 1.5$, $\sigma = 3$, $\sigma = \infty$

We plot demand curve for different elasticities in Figure 1. The red curve ($\sigma = 1.5$) is steep: quantity must fall a lot to raise price because substitution is hard. The blue curve ($\sigma = 3$) is flatter: with easy substitution, a small price hike loses many buyers. As $\sigma \to \infty$ the curve becomes horizontal – market power goes to zero and any small deviation from competitive prices would drive quantity to zero.

Production To produce a given quantity $q_i(\varphi)$, firms use the following amount of labor:

$$\ell = \bar{f} + a^* q_i(\varphi) \iff q_i(\varphi) = \frac{1}{a^*}(\ell - \bar{f})$$

where a^* denotes how many workers are necessary to produce a single unit of good φ. These are the **unit labor requirements**, as we have seen in the Ricardian model. By analyzing the right hand side of the equation above, we see that labor only contributes to output if they hire more than $\bar{f}$ workers. To produce $q_i(\varphi)$ units, then, the firm needs to hire $\bar{f}$ workers just to set up shop and ℓ workers to attain that level of production.

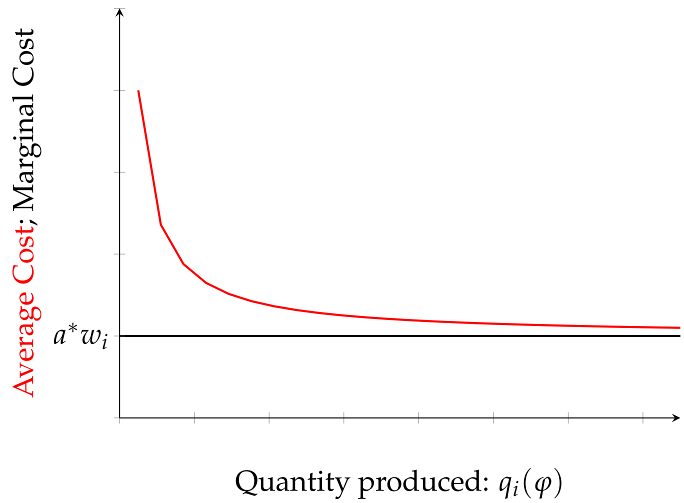

Figure 2: Average cost as a function of output

Another way of thinking of $\bar{f}$ is the "cost of entering the market". Every new firm must pay the same up-front cost for product design, advertising, setting up a factory, etc. Once that hurdle is cleared, producing an extra unit only costs labor at a constant marginal cost.

Firms have a monopoly over the production of their goods. This means that they have market power. **Instead of taking prices as given, they take demand as given and choose prices that will maximize profits**. If a firm enters the market, they maximize:

$$\begin{aligned} & \max_{p_i(\varphi)} \quad \pi_i(\varphi) \equiv p_i(\varphi) q_i(\varphi) - w_i a^* q_i(\varphi) - w_i \bar{f} \qquad s.t. \qquad q_i(\varphi) = \left(\frac{p_i(\varphi)}{P_i}\right)^{-\sigma} \times \frac{w_i L_i}{P_i} \\ \iff & \max_{p_i(\varphi)} \quad \pi_i(\varphi) \equiv p_i(\varphi)^{1-\sigma} P_i^{\sigma} \times \frac{w_i L_i}{P_i} - w_i a^* p_i(\varphi)^{-\sigma} P_i^{\sigma} \times \frac{w_i L_i}{P_i} - w_i \bar{f} \end{aligned}$$

This is a simple concave maximization problem that you know how to solve[1]. Optimal prices satisfy:

$$\begin{aligned} 0 &= (1-\sigma) \times p_i(\varphi)^{-\sigma} P_i^{\sigma} \times \frac{w_i L_i}{P_i} + \sigma \times w_i a^* p_i(\varphi)^{-\sigma-1} P_i^{\sigma} \times \frac{w_i L_i}{P_i} \\ 0 &= (1-\sigma) + \sigma \times w_i a^* p_i(\varphi)^{-1} = -p_i(\varphi)(\sigma - 1) + \sigma \times w_i a^* \end{aligned}$$

Solving for prices:

$$\boxed{p_i(\varphi) = \frac{\sigma}{\sigma - 1} \times w_i a^*}$$

[1]Note $d\pi/dp > 0, d^2\pi/dp^2 < 0$

Three important things emerge from this price. First, $w_i a^*$ is the **marginal cost** to produce one additional unit. Second, since $\sigma > 1$, then $\frac{\sigma}{\sigma-1} > 1$ is a **mark-up** that producers charge on top of marginal cost. Third, since we assumed that a^* is the same for every firm in the economy, the price **does not depend on the particular good** φ. So we can summarize pricing under monopolistic competition as:

$$\boxed{p_i(\varphi) = p^* = \text{mark up} \times \text{marginal cost} \qquad \text{for all goods } \varphi \in \Phi_i}$$

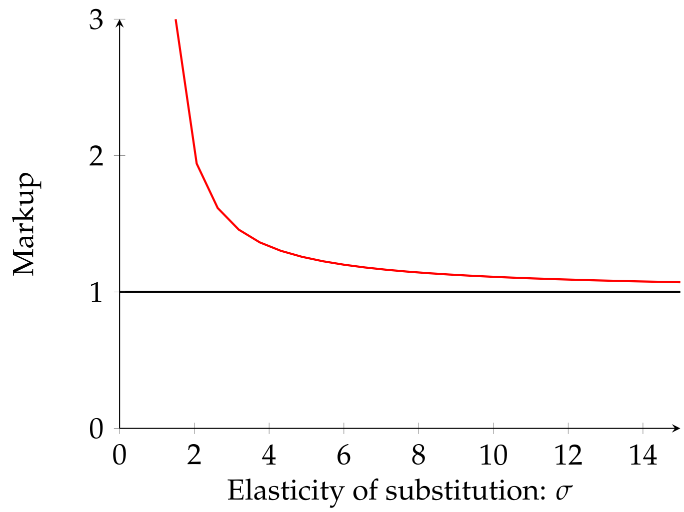

Figure 3: Markup as a function of the elasticity of substitution

Autarky Equilibrium We still have some questions unanswered. What is the quantity demanded from each firm? How many firms N will actually enter the market in equilibrium? What are the wages w_i and the price levels P_i?

Firms will only enter the market if they expect their profit $\pi_i(\varphi) \geq 0$ to be nonnegative (at least zero). If not, they would make a loss, so it would be rational to exit the market. But if profits are (strictly) positive ($\pi_i(\varphi) > 0$), then new entrants would have an incentive to pay the fixed cost $w_i \bar{f}$, set up a new shop for a new product, charge the markup over marginal cost, and make a profit. Entry continues until the last comer finds that her expected profit is exactly zero.

In other words, firms will enter the market up to the point in which there is no additional expected profit to be made and $\pi_i(\varphi) = 0$.

$$\pi_i(\varphi) = (p^* - MC)q^* - w_i \bar{f} = \left(\frac{\sigma}{\sigma - 1} a^* w_i - a^* w_i \right) q^* - w_i \bar{f} = 0 \iff q^* = (\sigma - 1) \times \frac{\bar{f}}{a^*}$$

In equilibrium, the quantity produced by each firm depends on the elasticity of substitution (if σ is high, markups will be smaller, and quantity sold per firm will be higher) and on the fixed cost $\bar{f}$, which controls the returns to scale in this model.

A surprising result is that, since all firms are identical, none of them will make positive profits in equilibrium! Consumer spending $w_i L_i$ is fixed in the short run. Each additional good gives shoppers another option, so demand for every incumbent variety falls. Because the markup is unchanged, the contribution margin per unit stays the same, but the number of units sold per firm shrinks.

Knowing the optimal size of firms q^*, we can now calculate total utility:

$$Q_i = \left[\sum_{\varphi \in \Phi_i} q_i(\varphi)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} = \left[N^*(q^*)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} = (N^*)^{\frac{\sigma}{\sigma-1}} q^*$$

where N^* is the total number of goods offered in equilibrium.

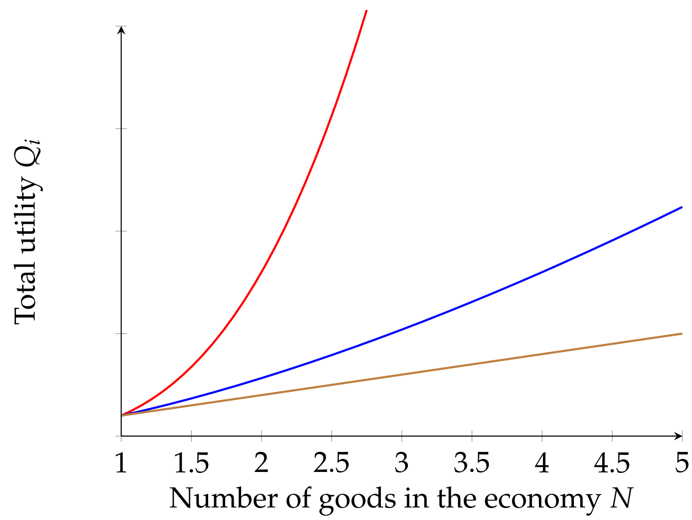

Figure 4: Love of variety with different elasticities: $\sigma = 1.5$, $\sigma = 3$, $\sigma = \infty$

The factor $\frac{\sigma}{\sigma-1} > 1$ captures the "love-of-variety" effect: when the elasticity of substitution is finite, adding new goods raises aggregate utility more than proportionally to their count because the composite good rewards diversity. We can also calculate the price level:

$$P_i = \left[\sum_{\varphi \in \Phi_i} p_i(\varphi)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} = \left[N^*(p^*)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} = \frac{p^*}{(N^*)^{\frac{1}{\sigma-1}}}$$

so price levels are decreasing in the number of goods available in the economy.

Finally, we can use the labor market clearing condition to pin down the number of firms. In equilibrium labor demand must equal labor supply L_i:

$$L_i = \sum_{\varphi \in \Phi_i} \left(\bar{f} + a^* q_i(\varphi)\right) = N\left(\bar{f} + a^* q^*\right) = N\sigma\bar{f} \iff N^* = \frac{L_i}{\sigma \bar{f}}$$

So the equilibrium number of firms N^* depends on the size of the market L_i, the fixed cost $\bar{f}$ and the elasticity of substitution σ. The larger the market size L_i, the more varieties it can accommodate in equilibrium. The larger the fixed cost $\bar{f}$, the larger the degree of increasing returns to scale, so firms must be larger and there will be fewer firms in equilibrium. The larger σ, the more substitutable goods are, so markups are smaller and surviving firms have to sell more units to be able to pay for fixed costs, so there are fewer firms in equilibrium.

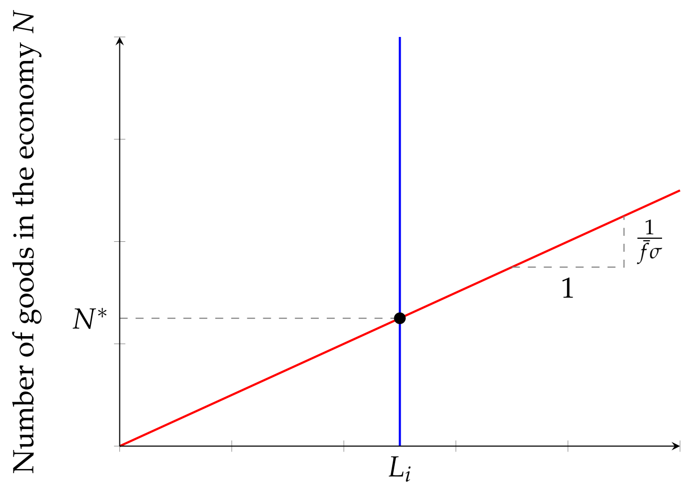

Figure 5: Labor market and number of firms equilibrium

Krugman model in autarky:

$$\boxed{\begin{array}{ll} p^* = \dfrac{\sigma}{\sigma-1} a^* w, & q^* = (\sigma-1)\dfrac{\bar{f}}{a^*}, \\ N^* = \dfrac{L_i}{\sigma \bar{f}}, & P = \dfrac{p^*}{(N^*)^{1/(\sigma-1)}}, \\ Q = \dfrac{I_i}{P} = (N^*)^{\frac{\sigma}{\sigma-1}} q^*. & w = 1 \end{array}}$$

Notes: a^* is the unit labor requirement; $\bar{f}$ fixed labor cost; $\sigma > 1$ CES elasticity.

Trade Equilibrium Assume the two symmetric countries (H and F), disregarding all trade barriers and shipping costs. Consumers now have preferences over domestic and foreign goods:

$$\max_{\{q_i(\varphi)\}_{\varphi \in \Phi_H \cup \Phi_F}} Q_i \equiv \left[\sum_{\varphi \in \Phi_H} q_i(\varphi)^{\frac{\sigma-1}{\sigma}} + \sum_{\varphi \in \Phi_F} q_i(\varphi)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \qquad s.t. \qquad P_i Q_i = \sum_{\varphi \in \Phi_H \cup \Phi_F} p_i(\varphi) q_i(\varphi) \leq I_i = w_i L_i$$

where $\Phi_H := \{H_1, H_2, \cdots, H_{N_H}\}$ is the set of domestic goods and $\Phi_F := \{F_1, F_2, \cdots, F_{N_F}\}$ is the set of foreign goods.

Everything else stays the same. Labor forces, and fixed costs are identical. Since preferences are identical in each country, each firms face the same demand curve. The optimal price they will choose will be the same:

$$p^* = \text{markup} \times \text{MC} = \frac{\sigma}{\sigma - 1} \times MC$$

Since fixed costs and marginal costs as the same as before, the free entry condition $\pi_i(\varphi) = 0$ pins down the same optimal quantity per firm:

$$q^* = (\sigma - 1) \times \frac{\bar{f}}{a^*}$$

How many firms exist in each country? Once again we use the labor market clearing condition to find that:

$$N^* = \frac{L_i}{\sigma \bar{f}}$$

The number of active firms in each country is still the same as in autarky. So what has changed? Integration changes only the *size of the market* in terms of available goods. That becomes clear once we evaluate what happens with welfare/utility:

$$\begin{aligned} Q_i &\equiv \left[\sum_{\varphi \in \Phi_H} q_i(\varphi)^{\frac{\sigma-1}{\sigma}} + \sum_{\varphi \in \Phi_F} q_i(\varphi)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \\ &= \left[N^*(q^*)^{\frac{\sigma-1}{\sigma}} + N^*(q^*)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \\ &= (2N^*)^{\frac{\sigma}{\sigma-1}} q^* \end{aligned}$$

which is higher than in the autarky equilibrium by a factor $2^{\frac{\sigma}{\sigma-1}} > 2$. So consumers are better off after trade openness:

$$Q_i^{\text{Trade}} = (2N^*)^{\frac{\sigma}{\sigma-1}} q^* > (N^*)^{\frac{\sigma}{\sigma-1}} q^* = Q_i^{\text{Autarky}}$$

Because the two countries are mirrors of each other, half of the new entrants locate in H and half in F. Hence each country *hosts* the same number of firms as before ($N^* = L/(\sigma \bar{f})$), but consumers in both countries can now purchase *all* $2N^*$ varieties.

Every firm sells one identical good to *both* markets. Country H exports the set of varieties it produces and imports the set produced in F, and vice-versa. Because the two sets have equal value, bilateral trade is balanced:

$$\text{Exports}_{H \to F} = \text{Imports}_{F \to H}$$

This two-way or *horizontal* intra-industry trade arises even though the countries are identical. In Ricardian or Heckscher–Ohlin (HO) models, identical countries would have *no* reason to trade—gains there hinge on cross-country differences. Here, gains come purely from the **love-of-variety** under CES preferences and the ability of increasing-returns firms to cover fixed costs in a larger market.

Trade in the Krugman model is fundamentally different from comparative advantage trade: *it is horizontal and survives even when countries look exactly alike*. Integration doubles the menu of goods, lowers the price index, and raises welfare without changing firm size, wages, or the markup.