Together at Last: Trade Costs, Demand Structure, and Welfare

MRAZOVA, Monika, NEARY, Peter


DOI : 10.1257/aer.104.5.298
International trade, like other branches of applied theory, has made enormous progress in recent decades by building on a central insight of Dixit and Stiglitz (1977): the easiest way to model a taste for variety, an essential foundation for a theory of monopolistic competition, is using a conventional utility function defined over the quantities of all potential commodities. To operationalize this, they considered two alternative specifications of the utility function: additively separable preferences and the CES special case. The former approach was used in Krugman (1979), one of the first applications of monopolistic competition to trade. However, he assumed that trade was unrestricted, and modeled trade liberalization only as an expansion of the global economy. When he and others turned to examine restrictions to trade, it became the norm to consider only the CES case: in the words of Krugman (1980, p. 953), "it seems worth sacrificing some realism to gain tractability." The result is paradoxical. We now have a clear understanding of many issues in trade under monopolistic competition, and, thanks to Arkolakis, Costinot, and Rodríguez-Clare (2012)—henceforth, ACRC—and others, a clear basis for quantifying the gains from trade, but only under CES assumptions, with their unsatisfactory implication that firms’ price-cost mark-ups are invariant to shocks.

A number of authors have considered particular alternatives to the CES. By contrast, the combination of trade costs and general demands has received little attention.1 In this paper we show that trade costs and additively separable preferences can be combined in a simple model, which is tractable without sacrificing too much realism. Section I sketches the model and introduces two key concepts: superconvex demand and superconcave utility. Section II compares the implications of integrated and segmented markets for prices and mark-ups. Section III shows how the pattern of sales across markets responds to globalization and trade-cost shocks. Section IV derives the implications for the gains from trade, while Section V discusses the problems of calibrating them. Technical details and additional references which for reasons of space have had to be omitted are given in the online version and in Mrázová and Neary (2013).

I. Preliminaries

Except for allowing trade costs, the setting is the same as in Krugman (1979). In each of \( \kappa + 1 \) identical countries, there is a single monopolistically competitive industry, with a measure \( n \) of identical firms, each producing a single symmetrically differentiated variety. International trade incurs symmetric iceberg trade costs \( \tau \), but no fixed costs. It follows that trade...
is all-or-nothing: except when trade costs are prohibitive, every consumer in the world consumes each of the \( N \equiv (\kappa + 1)n \) varieties produced in the world. Why do they bother? Because they have a taste for variety, modeled by expressing utility \( U \) as a monotonically increasing function of an integral of identical sub-utility functions. With symmetry, we need only distinguish between the consumption of a typical home and imported variety, \( x \) and \( x^* \), respectively:

\[
(1) \quad U = F[n\{u(x) + \kappa u(x^*)\}] \quad F', \ u' > 0, \ u'' < 0.
\]

Maximizing this, facing given income and prices, leads to inverse Frisch demands: \( p = \lambda^{-1}u(x) \) and \( p^* = \lambda^{-1}u(x^*) \); \( \lambda \) is the marginal utility of income, which firms take as given in choosing their optimal sales.

With so much symmetry assumed, the general-equilibrium structure of the model is straightforward. Goods-market clearing requires that each firm's output, denoted by \( y \), meet global demand for its product, with the proviso that \( \tau x^* \) units must be shipped abroad to ensure that \( x^* \) arrive:

\[
(2) \quad y = L(x + \kappa \tau x^*).
\]

Labor is the only factor of production, and the supply of identical worker-consumers in each country is fixed at \( L \). Technology follows the Dixit-Stiglitz specification, perhaps the simplest possible way of allowing for increasing returns. Each firm requires \( f \) workers to operate, and \( c \) workers to produce a unit of output. Labor-market clearing in every country therefore implies

\[
(3) \quad L = n(f + cy).
\]

We follow the Marx-Keynes-Krugman-Melitz convention of measuring nominal variables in labor units, so the wage is set equal to one by choice of numéraire.

The elasticity of substitution is a sufficient statistic for comparative statics with CES preferences. With general additive preferences, we need to know two statistics to understand the positive effects of exogenous shocks: the elasticity \( \varepsilon(x) \equiv -\frac{p(x)}{xp'(x)} \) and convexity \( \rho(x) \equiv -\frac{xp''(x)}{p'(x)} \) of demand. In addition, to understand normative implications, we need to know the elasticity of the sub-utility function \( \xi(x) \equiv \frac{xu'(x)}{u(x)} \); this is an inverse measure of consumers' taste for diversity, and must lie between zero and one. We write \( \varepsilon^* = \varepsilon(x^*) \) and so on for parameters pertaining to imports.\(^2\)

Many implications of these parameters can be summarized using two key properties. The first we call superconvexity of demand: a demand function is superconvex at a point if it is more convex than a CES demand function with the same elasticity, \( \rho \geq \frac{\varepsilon + 1}{\varepsilon} \), otherwise it is subconvex. With superconvexity, the elasticity of demand rises as per capita consumption increases, so it is crucial for the difference between a firm’s mark-ups on its home and foreign sales. The second property we call superconcavity of utility: a sub-utility function is superconvex at a point if it is more concave than a CES sub-utility function with the same elasticity, \( \xi \geq \frac{\varepsilon - 1}{\varepsilon} \), otherwise it is subconcave.\(^3\) With superconcavity, the elasticity of utility falls, i.e., taste for diversity rises, as per capita consumption increases, so it is crucial in determining how consumers trade off changes at the intensive and extensive margins of consumption.

II. Integrated or Segmented Markets?

The first issue we must address is whether home and foreign markets are integrated or segmented. Integrated markets imply that prices are equalized, \( p^* = \tau p \), segmented markets that marginal revenues are equalized, \( r^*_x = \tau r_x \).\(^4\) With trade costs, these coincide if and only if demands are CES, since otherwise the ratio of price to marginal revenue differs across markets:

\[
(4) \quad p = \frac{\varepsilon}{\varepsilon - 1} r_x, \quad p^* = \frac{\varepsilon^*}{\varepsilon^* - 1} r_x^*.
\]

What does this imply for the pattern of price-cost mark-ups across markets? When markets are integrated, mark-ups are the same

\(^2\) In the CES case, only the elasticity of substitution \( \sigma \) matters: \( \{\xi^*, \varepsilon^*, \rho^*\} = \{\xi, \varepsilon, \rho\} = \{\frac{\sigma - 1}{\sigma}, \sigma, \frac{\sigma + 1}{\sigma}\} \).

\(^3\) The concavity of an arbitrary sub-utility function is \( -\frac{u''}{u'} = \frac{1}{\xi} \); while, from footnote 2, the concavity of a CES sub-utility function with elasticity \( \xi \) is \( 1 - \xi \).

\(^4\) \( r(x) \equiv xp(x) \) and \( r(x^*) \equiv x^* p(x^*) \) denote sales revenue at home and abroad, respectively.
at home and abroad: \( \frac{\tau}{c} = \frac{\tau^*}{c^*} \). However, when markets are segmented, mark-ups differ in a way that depends on the convexity of demand. With subconvexity, the elasticity is higher in export markets and so from (4) the mark-up is lower. Moreover, the price charged abroad is lower than the trade-cost-inclusive home price: \( p^* < \tau p \). Segmented markets thus exhibit reciprocal dumping but without oligopoly as in Brander and Krugman (1983). All these statements are reversed if demands are superconvex: margins are higher abroad; prices there exceed the trade-cost-inclusive home price; and markets exhibit reciprocal antidumping. Which is the more likely case? Though there is no clear consensus, the balance of empirical and other evidence suggests that subconvex demands are more realistic than superconvex, implying that reciprocal dumping is the norm.

III. Globalization or Colder Icebergs?

To save space, we focus on the more realistic case of segmented markets. We consider two kinds of trade liberalization: increases in the number of countries \( \kappa \) (“globalization”), and reductions in trade costs \( \tau \) (“colder icebergs”). These have very different effects on the conditions for firm and industry equilibrium.

At the firm level, profit maximization equalizes \( \tau \)-inclusive marginal revenues across markets, as we have seen. Totally differentiating, using “hats” to denote proportional changes (\( \hat{x} \equiv d \log x, x \neq 0 \)):

(5) \[ \hat{r} + \hat{\tau} = \hat{r}^* + \hat{\tau}, \]

where \( \eta \equiv -\frac{x r x}{\bar{r} x} = \frac{2 - \rho}{\varepsilon - 1} \) is the elasticity of marginal revenue at home. This implies a positive relationship between home and foreign sales, as illustrated by the curves labeled “MR = MC” in Figure 1. These are shifted upward by reductions in trade costs, but are unaffected by changes in the number of countries.

At the industry level, free entry requires that operating profits, \( \pi + \kappa \pi^* \), equal fixed costs \( f \). From the first-order condition, operating profits in an export market are: \( \pi^* = (p^* - \tau c) L x^* = \frac{\tau c}{\varepsilon - 1}, \) and analogously at home. Totally differentiating the free-entry condition:

(6) \[ \omega \varepsilon \eta \hat{x} + (1 - \omega \varepsilon) \varepsilon^* \eta^* \hat{x}^* = -(1 - \omega \varepsilon)(\hat{\kappa} + \hat{\tau}), \]

where \( \omega \varepsilon \equiv \frac{x}{\pi + \kappa \pi^*} \) is the home-market share in operating profits. This implies a negative relationship between home and foreign sales, as illustrated by the curves labeled “\( \Pi = 0 \)” in Figure 1. These are affected in the same way by increases in \( \tau \) and \( \kappa \): both pivot the curve anti-clockwise around the autarky point \( A \) (the zero-profit point conditional on not exporting).

Combining these results we can deduce the effects on home and export sales:

(7) \[ \bar{e} \varepsilon \eta \hat{x} = (1 - \omega \varepsilon)[-\hat{\kappa} + (\varepsilon^* - 1)\hat{\tau}] \]

(8) \[ \bar{e} \varepsilon \eta^* \hat{x} = -(1 - \omega \varepsilon)\hat{\kappa} - [1 + \omega \varepsilon(\varepsilon - 1)]\hat{\tau}, \]

where \( \bar{e} \varepsilon \) is an aggregate elasticity weighted by profit shares: \( \bar{e} \varepsilon \equiv \omega \varepsilon \varepsilon + (1 - \omega \varepsilon)\varepsilon^* \). Borrowing terminology from the analysis of devaluation, globalization leads to expenditure-reduction (more varieties in the world reduces spending on each individual variety) whereas a fall in trade costs leads to expenditure-switching (consumption of

---

5 We owe this insight to Sergey Kokovin.
6 The firms’ first- and second-order conditions require \( \varepsilon > 1 \) and \( \rho < 2 \) respectively, so \( \eta \) must be positive.
imported varieties rises at the expense of domestic ones, as shown in Figure 1).

The two shocks clearly have very different effects on home and foreign sales. Moreover, their effects cannot be aggregated, because the elasticities of marginal revenue differ. There are two exceptions to this rule. One is free trade: the elasticities of marginal revenue are now the same on both home and export sales, and the $MR = MC$ and $\Pi = 0$ loci are straight lines with slopes of 1 and $-\frac{1}{\tau}$ respectively. The other is CES demands: both loci are now straight lines, and sales adjust to changes in trade costs along a smooth locus $ACF$, mirroring their smooth adjustment to changes in the number of countries along a straight-line $MR = MC$ locus. This explains why, as noted by ACRC, the effects of both shocks on aggregate welfare in the CES case are isomorphic: both can be summarized in terms of their effects on the home-market share in total output. This isomorphism breaks down when demands are not CES.

Given changes in sales, it is easy to deduce the changes in prices, output, and firm numbers. Prices are directly linked to sales by the firms’ first-order conditions:

\[
\begin{align*}
\hat{p} &= \frac{\varepsilon + 1 - \varepsilon \rho}{\varepsilon (1 - \varepsilon)} \hat{x}, \\
\hat{p}^* &= \frac{\varepsilon^* + 1 - \varepsilon^* \rho^*}{\varepsilon^*(1 - \varepsilon^*)} \hat{x}^* + \hat{\tau}.
\end{align*}
\]

Both are increasing with sales if and only if demands are subconvex: as consumption rises, the elasticity of demand falls, and so mark-ups increase. Hence both globalization and lower trade costs reduce all prices.\(^7\)

As for adjustment at the intensive and extensive margins, these follow directly from the market-clearing conditions (2) and (3). From (2), the change in firm output is a weighted average of the changes in sales:

\[
\begin{align*}
\hat{y} = \omega_x \hat{x} + (1 - \omega_x)(\hat{k} + \hat{\tau} + \hat{x}),
\end{align*}
\]

where $\omega_x \equiv \frac{x}{x + \kappa \tau x}$ is the home-market share in total output. This in turn inversely determines the number of active firms and so of produced varieties in each country from the full-employment condition (3):

\[
\hat{n} = -\psi \hat{y}, \quad \psi = \frac{\varepsilon_h - 1}{\varepsilon_h}.
\]

Here, $\psi \equiv \frac{\varepsilon_v}{\varepsilon + \varepsilon_v}$ is the share of variable costs in total costs, which is an inverse measure of returns to scale. It is increasing in the aggregate elasticity $\varepsilon_h$, which is an output-weighted harmonic mean of the home and foreign demand elasticities: $\varepsilon_h \equiv \left[\omega_x \varepsilon^{-1} + (1 - \omega_x)(\varepsilon^*)^{-1}\right]^{-1}$.

Just as the changes in mark-ups in (9) hinge on subconvexity of demand, so too do those of output and firm numbers in (10) and (11). In the CES case, output is fixed, and exogenous shocks merely reallocate sales: globalization encourages firms to sell to more markets, but less in each; higher trade costs induce a reduction in production for exports which exactly offsets the increase in home sales. If instead demand is subconvex, mark-ups fall as per capita sales fall. Hence, with globalization, the negative effect on profits of lower sales per market exceeds the positive effect of a rise in the number of markets; so, to keep overall profits equal to zero, total output must rise. As for higher trade costs, to keep profits constant requires sales to fall by less in declining markets than they rise in expanding markets; so, here too total output must rise, at least in the neighborhood of free trade, where (7), (8), and (10) imply

\[
\begin{align*}
\hat{y} \bigg|_{\tau=1} &= (1 - \omega) \left(1 - \frac{1}{\varepsilon \eta}\right) (\hat{k} + \hat{\tau}).
\end{align*}
\]

The key expression on the right-hand side is positive if and only if the elasticity of marginal revenue $\eta$ exceeds the elasticity of inverse demand $\frac{1}{\varepsilon}$, which is equivalent to demand being

\(^7\) The $MR = MC$ locus reduces to $x^* = \tau^{-\gamma} x$, and the $\Pi = 0$ locus to $x + \kappa \tau x^* = y/L$. Eliminating $\tau$ gives the $ACF$ locus: $x^* = (\gamma y/L - x/\kappa)^{\tau^{-\gamma} - 1} x^{-\gamma-1}$.

\(^8\) Lower trade costs always reduce import prices: though the mark-up on foreign sales may rise or fall, the direct effect of the change in trade costs always dominates: $\hat{p}^* \tau \hat{\tau} = 1 - \frac{\varepsilon^* + 1 - \varepsilon^* \rho^*}{\varepsilon^*(2 - \rho^*)} \frac{1 + \omega_x (\varepsilon - 1)}{\varepsilon_x} > 0$.

\(^9\) Here, both weights reduce to: $\omega_x = \omega_v = \frac{1}{\kappa + 1}$.
This implies that trade liberalization has opposite effects on total output, and so, from (11), on the number of firms per country, depending on whether it involves an increase in $\kappa$ or a reduction in $\tau$. Paradoxically, lower trade costs in the neighborhood of free trade reduce firm output and increase the number of domestic firms if demand is subconvex. Total sales always increase as $\tau$ falls, but when demand is subconvex they increase by less than the fall in trade costs.\footnote{Márzová and Neary (2013) note sufficient condition for gains is that $\psi > \xi$, so the initial equilibrium is efficient: for given values of the exogenous variables, no change in $n$ can raise welfare. This obtains either if preferences are CES (a familiar result from Dixit and Stiglitz 1977), or if a global anti-trust policy continually adjusts firm numbers to ensure efficiency. A second sufficient condition for gains is that $\psi - \xi$ and $\hat{n}$ have the same sign. For example, both are positive when utility is subconcave ($\psi > \xi$, so consumers desire more variety) and demand is subconvex (so, from Section III, trade liberalization increases the number of varieties).}

### IV. Gains from Trade

We measure welfare changes by the change in equivalent income, $Y$, needed to keep consumers at their initial utility level:

\[
(13) \quad \hat{Y} = \left( \frac{\bar{\varepsilon}}{\bar{\varepsilon}_u} \frac{1}{\xi_u} - 1 \right) \hat{N}_Y - \omega_Y \hat{p} - (1 - \omega_Y) \hat{p}^*,
\]

where $\hat{N}_Y$ is a composite change in the number of varieties.\footnote{Consider first the home-market shares. These are all equal in two cases. In free trade:}

V. Calibrating the Gains

Qualitative results such as those in the previous section are valuable for giving intuition, but the complexity of the general expressions when trade costs are initially positive means that we have to resort to calibration. Space constraints preclude our presenting detailed results, so instead we note some general considerations relating to calibrating the gains from trade.

Our results depend on relatively few parameters: various home-market shares, plus the elasticity and convexity of utility and demand for home and imported varieties. While this is not bad news for calibrationists, “relatively few” is more than two, the number that ACRC showed is needed to calibrate the gains from trade in CES-based models. Here, the same parameters arise as in their case—the home-market share in output and the elasticity of demand—but many variants of each are required.

Consider first the home-market shares. These are all equal in two cases. In free trade, they equal
the share of each country in world GNP, \( \frac{1}{1 + \kappa^r} \), with CES preferences, they equal \( \frac{1}{1 + \kappa^s} \).

More generally, they differ from each other, as shown in Table 1. For example, \( \omega_i > \omega_z > \omega_k \) if and only if demands are subconvex: home sales have higher markups, so they contribute more to profits than to sales value, and more to sales value than to production.\[15\]

Consider next the average elasticities. These too can be ranked, both relative to each other and relative to the elasticities of demand for home and imported varieties. Some rankings are independent of subconvexity: it is always true that \( \bar{\varepsilon}_i > \bar{\varepsilon}_z > \bar{\varepsilon}_k \).\[16\]

By contrast, ranking the average demand elasticities relative to the elasticities for both kinds of varieties hinges on subconvexity: \( \varepsilon < \bar{\varepsilon}_i < \varepsilon^* \) for all \( i \) if and only if demand is subconvex. In that case, calibration exercises that use elasticities estimated from import data will overestimate the true weighted elasticities. Higher elasticities typically reduce the gains from trade, so using import demand elasticities in calibration exercises will typically underestimate the gains from trade.

### VI. Conclusion

In this paper we have used the approach of Márzová and Neary (2013) to explore the implications of combining two real-world features typically studied in isolation in general-equilibrium trade models: variable demand elasticities, and barriers to international trade. Even in our simple setting, relaxing the assumption of CES preferences in monopolistic competition has surprising implications when trade is restricted. Integrated and segmented markets behave differently, the latter typically exhibiting reciprocal dumping. Globalization and lower trade costs have very different effects: the former reduces spending on all existing varieties, the latter switches spending from home to imported varieties; when demands are subconvex, globalization raises firm output but lower trade costs reduce it. Finally, calibrating gains from trade is harder. Many more parameters are needed, while import demand elasticities are likely to overestimate the true elasticities, and so underestimate the gains from trade.

### REFERENCES