

NBER WORKING PAPER SERIES

EXTERNAL ECONOMIES AND INTERNATIONAL TRADE REDUX

Gene M. Grossman
Esteban Rossi-Hansberg

Working Paper 14425
<http://www.nber.org/papers/w14425>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2008

We thank Arnaud Costinot, Angus Deaton, Elhanan Helpman, Marc Melitz and Giovanni Maggi for discussions and the National Science Foundation (under grant SES 0451712) and Sloan Foundation for research support. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation nor the National Bureau of Economic Research, or those of any other organization.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2008 by Gene M. Grossman and Esteban Rossi-Hansberg. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

External Economies and International Trade Redux
Gene M. Grossman and Esteban Rossi-Hansberg
NBER Working Paper No. 14425
October 2008
JEL No. F1,F11

ABSTRACT

We study a world with national external economies of scale at the industry level. In contrast to the standard treatment with perfect competition and two industries, we assume Bertrand competition in a continuum of industries. In this setting, many of the "pathologies" of the standard treatment disappear. There typically exists a unique equilibrium with trade guided by "natural" comparative advantage. And, when a country has CES preferences and any finite elasticity of substitution between goods, gains from trade are assured.

Gene M. Grossman
International Economics Section
Department of Economics
Princeton University
Princeton, NJ 08544
and NBER
grossman@princeton.edu

Esteban Rossi-Hansberg
Princeton University
Department of Economics
Fisher Hall
Princeton, NJ 08544-1021
and NBER
erossi@princeton.edu

1 Introduction

External economies hold a central—albeit somewhat uncomfortable—place in the theory of international trade. Since Marshall (1879, 1890) at least, economists have known that increasing returns can be an independent cause of trade and that the advantages that derive from large scale production need not be confined within the boundaries of a firm. Marshallian externalities arise when knowledge and other public inputs associated with a firm’s output spill over to the benefit of other industry participants. After Marshall’s initial explication of the idea, a lengthy debate ensued on whether an industry with external economies of scale could logically be modelled as perfectly competitive.¹ Even after the matter was finally resolved in the affirmative by Chipman (1970), trade economists continued to bemoan the “bewildering variety of equilibria [that provide] a taxonomy rather than a clear set of insights” (Krugman, 1995, p.1251), the apparent tension between the forces of external economies and other, better-accepted determinants of the trade pattern, and the “paradoxical” implication that trade motivated by the gains from concentrating production need not benefit the participating countries.

The modeling of external economies began at a time when game theoretic approaches to imperfect competition were much less developed than they are today. Consequently, the logical consistency between external economies and perfect competition was seen as a great advantage in allowing the application of familiar tools of general equilibrium analysis. Economists were comfortable in assuming that competitors take prices as given in industries with many small firms. By assuming that small producers contribute little to aggregate industry output, they felt similarly justified in assuming that firms treat industry scale as given when assessing productivity and making decisions about their own output. Firms might recognize the link between industry scale and productivity and nonetheless perceive their own costs to be constant. As Chipman (1970, p.349) noted, “there is no logical contradiction involved in the notion that economic agents do not perceive things as they actually are, and while the idea of treating production functions parametrically may be more subtle and unusual than that of treating market prices in this way, both ideas are of the same logical order.”

In application, the competitive models of external economies and trade yielded some discomfiting results.² A potential for multiple equilibria, noted already by Ohlin (1933, p.55), was shown more formally by Matthews (1949-1950), Kemp (1964), and Ethier (1982a). Graham (1923) argued that scale economies could reverse the trade pattern predicted by “natural” comparative advantage and Ethier (1982a) proved him correct in a competitive model with external economies. Markusen and Melvin (1982) and Ethier (1982a) observed that country size

¹For a recount of the debate and its protagonists, see Chipman (1965).

²There are several good surveys of this literature, beginning with Chipman’s (1965) treatment of the earliest contributions. Subsequent overviews include Helpman (1984), Helpman and Krugman (1985, ch.3), Krugman (1995), and Choi and Yu (2003).

might play an independent role in determining the pattern of specialization. Finally, Graham (1923) advanced the possibility of losses from trade, an idea that has been further addressed by Matthews (1949-1950), Melvin (1969), Markusen and Melvin (1982), Ethier (1982a), and others. These “pathologies” are perhaps responsible for the diminished role that external economies now play in thinking about trade, even as such externalities are still considered to be empirically relevant in many contexts.³

The pathological outcomes arise in a canonical model of external economies and trade, one that has become well ensconced in the literature and that by now breeds little questioning of its underlying assumptions. Typically, there are two countries and two industries, one with constant returns to scale and the other with increasing returns to scale (IRS) due to national spillovers from production. Often, labor is the only productive factor. Firms in both sectors are price takers and, more importantly, those in the IRS sector take aggregate national output and hence their own potential productivity as given.

In this familiar setting, suppose that the relationship between productivity and national scale is such that zero output implies zero productivity. Then the existence of multiple equilibria is immediate. Firms that take aggregate output as given perceive zero productivity when no other national producers are active in their country. If output in a country is nil, no producer perceives a profit opportunity from entry, even when an absence of local competitors might imply a potential for profit for a firm that evaluates correctly the implications of entry at non-negligible scale. Thus, zero output in any location is self-reinforcing and production of the IRS good can be concentrated in either country for any set of parameter values.⁴

What about the pattern of trade? When production of the IRS good is concentrated in a country, the trade pattern necessarily accords with the ranking of relative productivities when measured at the equilibrium scales of production. That is, each country specializes according to *observed* comparative advantage, where the “observed” productivity of a country that does not produce the IRS good is taken to be zero. However, the pattern of specialization need not conform to the ranking of relative productivities when measured at a common scale of production. The exporters of the IRS good might be less productive than firms in the other country would be, were the latter to produce at a comparable aggregate scale. In this sense, the trade pattern may run counter to the dictates of *natural* comparative advantage. Notice that such a trade pattern violates the requirements for global efficiency in circumstances in which the country with comparative advantage in the IRS industry has sufficient resources to satisfy the first-best level of world demand for that good.

³See, for example, Caballero and Lyons (1989, 1990, 1992), Chan, Chen and Cheung (1995), Segoura (1996), Henriksen, Midlefart-Knarvik, and Steen (2001) and Antweiler and Treffer (2002).

⁴If productivity is bounded from below even when national output is zero, multiple trade equilibria will exist for a range of parameter values.

In the canonical model, country size can be an independent source of comparative advantage. Suppose, for example, that productivity in the IRS sector is modest when output is small, but not equal to zero. If a small country were to specialize in producing this good, the equilibrium price might be high in light of its meager resource base and a robust level of world demand. Such a high price might invite entry by producers in the large country. In such circumstances, there may be no equilibrium with production of the IRS concentrated in the small country. The large country gains a comparative advantage in the IRS industry by dint of having sufficient resources to produce at large scale.

The canonical model readily yields examples with losses from trade (see, e.g., Helpman and Krugman, 1985, p.55). The autarky equilibrium features underproduction of the IRS good, because firms fail to take into account the spillover benefit from increasing their output on the productivity of others. As is well known, trade can bring harm if it exacerbates a pre-existing market distortion (Bhagwati, 1971). It stands to reason that trade may reduce welfare in a country that shifts substantial resources from the single industry with increasing returns to one with constant returns.

In short, the pathologies that emerge from the canonical model rest on two critical assumptions. First, firms contemplate entry only at such small scale that they believe they can have negligible impact on national output and their own productivity. Second, industries are large in relation to countries' resource endowments, so that a small country may lack sufficient resources to meet world demands for the IRS good at a reasonable price. In this paper, we modify these two familiar assumptions and show how this dramatically narrows the scope for pathological outcomes. We assume in what follows that firms engage in Bertrand competition and that they correctly anticipate their productivity when they quote a price that would yield non-negligible sales. And we assume that the world economy comprises a large number of (small) industries, so that any industry can be fully accommodated in any country, no matter what its size.⁵

It is well known that perfect competition and Bertrand competition predict similar outcomes in models with homogeneous goods and constant costs. This is true as well—it turns out—in an autarky equilibrium with external economies of scale. However, the predictions of the alternative behavioral assumptions diverge in a world with national external economies and international trade. With Bertrand competition, a firm may be small when its price is the same as that of its (domestic) rivals—as will be the case in the equilibria that we discuss—yet it may recognize that by announcing a price below that offered by other firms in the industry, it can grow to non-negligible scale. We shall see that the different out-of-equilibrium beliefs under perfect and

⁵In our analysis the scope of the external effects is supposed to be the industry. So it is essential for us to assume that industries are small relative to country sizes. As advocated by Ethier (1982b), one could also think of externalities that transcend both industry and country boundaries. However, external economies at the industry level are plausible in the presence of industry-specific knowledge and other non-rival inputs, and the literature cited in footnote 3 points to their empirical significance.

Bertrand competition yield very different conclusions about the possibility of multiple equilibria and the determinants of the trade pattern.

Moreover, the potential role for country size to influence the trade pattern disappears when industries are small in relation to the size of national economies. In a world with many industries à la Dornbusch, Fischer and Samuelson (1977), a small country can realize its comparative advantage in whichever sectors it is strongest by serving world demands in a relatively small range of industries. We will find that country size plays no role in determining a chain of comparative advantage, although it does affect the dividing line between imports and exports along that chain. The assumption of a large number of small industries also helps to justify our assumption that firms take factor prices as given. A Bertrand competitor may entertain the prospect of growing large in its industry. But, even if it does so, it will remain small in relation to the aggregate economy so long as each industry uses a small fraction of the economy's resources.

In summary, we study a model with a continuum of industries and national industry-level external economies. We assume that firms in an industry produce a homogeneous product and engage in Bertrand (price) competition in an integrated world market. Countries may exhibit Ricardian technological differences, but they share the same prospects for scale economies, which are assumed to be a property of the good being produced. We find in this setting that the trade equilibrium typically is unique and that “natural” comparative advantage determines an ordering of industries.⁶ Most surprising, perhaps, are our findings about welfare. We consider for this purpose a country whose residents hold CES preferences over the set of goods, with possible asymmetries in the weighting of different goods, and with any finite elasticity of substitution. This includes, of course, the Cobb-Douglas case that has been used in the literature to demonstrate the possibility of losses from trade based on external economies of scale. We find, in contrast, that in spite of the generic inefficiency of the autarky and trade equilibria, gains from trade are assured.

The remainder of the paper is organized as follows. The next section develops our alternative model with Bertrand competition and a continuum of industries. There, we derive our predictions about the pattern of trade. Section 3 contains a brief review of the efficiency properties of the model, which sets the stage for the gains-from-trade analysis in the subsequent section. Section 5 concludes.

⁶If the countries share identical technologies for some or many goods, the model does not predict which goods are produced in each country. But factor prices and consumption levels typically are uniquely determined even in this case.

2 A Model of External Economies in a Continuum of Industries

We study an economy with a continuum of industries, as in Dornbusch, Fischer and Samuelson (1977; hereinafter, DFS). Goods are indexed by $i \in [0, 1]$. For the time being, we assume only that preferences in each country are such that a change in the price of a single good has a negligible effect on the demand for all goods other than itself.⁷ Demands need not be identical in the two countries, nor need they be homothetic.

Goods are produced by a single primary factor, labor. Throughout the paper, we take the home wage as numeraire. Technology exhibits constant or increasing returns due to external economies at the local industry level. In the home country, $a_i/A_i(x_i)$ units of labor are required for a unit of output, where x_i is aggregate home production of good i and $A_i(\cdot)$ is non-decreasing, concave, and has an elasticity smaller than one.⁸ In the foreign country, the unit labor requirement is $a_i^*/A_i(x_i^*)$. There are n_i potential producers of good i in the home country and n_i^* potential producers in the foreign country, where $n_i \geq 2$ and $n_i^* \geq 2$. The number of potential producers in each country can be very large, but is assumed to be finite. Firms in each industry engage in price (Bertrand) competition in the integrated world market. Each firm recognizes that, if it sets a price above that quoted by any other firm, it will consummate no sales, if it sets a price that is the lowest among all quoted prices, it will capture the entirety of world demand, and if it sets a price that is tied for lowest among a group of competitors, it will share demand with these rivals.

2.1 Autarky Equilibrium

Figure 1 shows the inverse demand curve for good i , labelled DD , for a given level of aggregate spending and a given level of any price aggregate that may affect the demand for this good. Firms in the industry take aggregate spending and the relevant price index (if any) as given. The figure also depicts the average cost curve, which is labelled CC . Intersection points are our candidates for industry equilibrium.

Suppose firms in industry i announce a price above \tilde{p}_i in Figure 1. Then a deviant firm could announce a price a bit lower than that offered by the others. The firm would capture all industry sales, and since its price would exceed its average cost at the resulting scale of production, it would earn a profit. Evidently, the price will be bid down at least to \tilde{p}_i . But a further price cut by any firm would spell losses, since the deviant would capture the market but fail to cover its costs. When the DD curve is everywhere steeper than the CC curve, as in Figure 1, then

⁷In this, we follow Wilson (1980). The Cobb-Douglas preferences invoked by DFS, as well as CES, Stone-Geary, and symmetric translog all have this property, as do many others.

⁸As in the literature, an elasticity of productivity with respect to output smaller than one is needed to ensure a non-negative marginal productivity of labor.

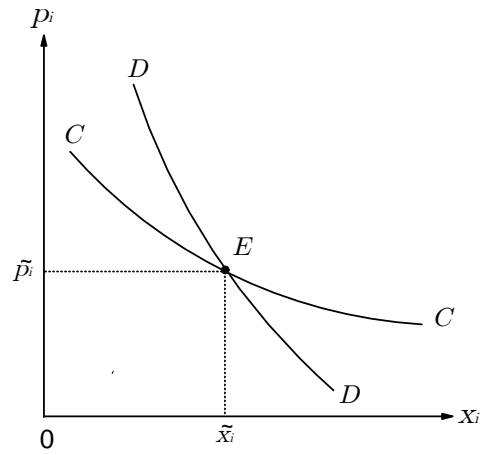


Figure 1: Unique Intersection of DD and CC

the unique intersection point depicts the industry equilibrium. In this equilibrium, price equals industry average cost and an arbitrary (and possibly large) number of firms make sales but earn zero profits.

In Figure 2, the DD and CC curves have multiple intersections at E' , E'' , and E . Neither E' nor E'' represents an equilibrium, because if all firms charge the price associated with either of these points, a deviant firm can announce a lower price such that, at the associated demand level, price exceeds average cost. The equilibrium in industry i is at the lowest intersection of DD and CC , so long as DD cuts CC from above.⁹

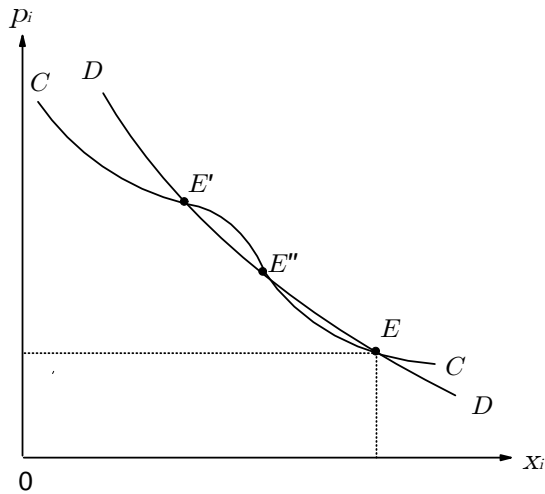


Figure 2: Multiple Intersections of DD and CC

⁹If no such intersection exists, then the price would tend toward zero and output would tend toward infinity. But then the assumption that firms take wages as given becomes untenable.

The general equilibrium requires, of course, that every industry equilibrates and that the labor market clears. There is no need to develop the notation to express this for the general case. Instead, we describe a special case that will prove useful later on. Suppose utility takes the form

$$U = \left(\int_0^1 b_i c_i^\rho di \right)^{1/\rho}, \quad (1)$$

where c_i is consumption of good i , with $\rho < 1$ and $\int_0^1 b_i di = 1$. These preferences imply a constant elasticity of substitution $\sigma = 1/(1 - \rho)$ between every pair of goods. The demand for good i is

$$c_i = \frac{b_i^\sigma p_i^{-\sigma} E}{P^{1-\sigma}},$$

where E is aggregate spending and $P = \left[\int_0^1 b_i^\sigma p_i^{1-\sigma} di \right]^{1/(1-\sigma)}$ is the exact price index associated with the CES utility function. With these demands, the condition that DD is steeper than CC at output x_i is equivalent to

$$\sigma \theta_i(x_i) < 1,$$

where $\theta_i \equiv A'_i(x_i)x_i/A(x_i)$ is the elasticity of the productivity function with respect to output. A sufficient condition for existence of an industry equilibrium with finite x_i is $\sigma \theta_i(x) < 1$ for all x .¹⁰ We will assume that an equilibrium with finite output exists in every industry, either because these sufficient conditions are satisfied, or otherwise.

In the autarky equilibrium (denoted again by tildes), aggregate spending \tilde{E} equals aggregate income, L . Then product market equilibrium requires

$$\tilde{x}_i = \frac{b_i^\sigma \tilde{p}_i^{-\sigma} L}{\tilde{P}^{1-\sigma}} \text{ for all } i, \quad (2)$$

while price equals average cost implies

$$\tilde{p}_i = \frac{a_i}{A_i(\tilde{x}_i)} \text{ for all } i. \quad (3)$$

This gives a system of equations that jointly determine outputs and prices in all industries.

We note in passing that the autarky equilibrium with Bertrand competition is the same as would arise under perfect competition, were firms to take their own productivity as given. Price in every industry is equal to average cost. Firms earn zero profits. Demand equals supply for all goods. And the labor market clears.

¹⁰Note that Ethier (1982, p.1247) invokes essentially the same assumption as a condition for uniqueness and stability of equilibrium in his two-industry model.

2.2 Trade Equilibrium

We now open the economy to international trade. Again, competition between the potential producers in a country drives price for every good i down to average cost. But now it also must be the case that a potential producer in the “other” country does not wish to shave price further. If home producers in industry i quote the price $wa_i/A(x_i)$, a foreign firm could announce a price a bit lower than that, capture sales of x_i , and achieve labor productivity of $a_i^*/A_i(x_i)$. Such a strategy would be profitable if its per unit cost, $w^*a_i^*/A_i(x_i)$, were less than the quoted price. Therefore, concentration of industry i in the home country requires

$$\frac{a_i}{A_i(x_i)} \leq \frac{w^*a_i^*}{A_i(x_i)}.$$

(Note that $w = 1$, because the home wage is numeraire.) Similarly, if production of good i is concentrated in the foreign country, it must be that

$$\frac{a_i}{A_i(x_i^*)} \geq \frac{w^*a_i^*}{A_i(x_i^*)}.$$

Assuming that each good is produced in only one location (which will be true in equilibrium), the equilibrium price of a good equals the lesser of the two average costs evaluated at the equilibrium scale of world production; i.e.,

$$p_i = \min \left[\frac{a_i}{A_i(\bar{x}_i)}, \frac{w^*a_i^*}{A_i(\bar{x}_i)} \right] \text{ for all } i, \quad (4)$$

where \bar{x}_i is world output of good i in the trade equilibrium.

Now order the goods so that $\alpha_i \equiv a_i/a_i^*$ is increasing in i and define I by $\alpha_I = w^*$. Then (4) implies that goods with $i \leq I$ are produced in the home country and goods with $i > I$ are produced in the foreign country.¹¹ Figure 3—familiar from DFS—depicts this relationship between w^* and I as the curve AA . It depicts as well a second relationship between these variables implied by the labor-market-clearing condition. The greater is I , the fewer are the goods produced in the foreign country and the lower must be the foreign wage for the labor market there to clear. This relationship is labelled BB . The equilibrium marginal good and relative wage are found at the intersection of the two curves. This determination of I and w^* is, of course, exactly the same as in DFS, as extended by Wilson (1980) to include economies with more general demands.

Two observations are in order. First, the pattern of specialization is well determined. The possibility of multiple locations for a given industry disappears once firms recognize that, even

¹¹For notational convenience, we adopt the convention that the marginal good is produced in the home country, although this is neither determined by the model nor consequential.

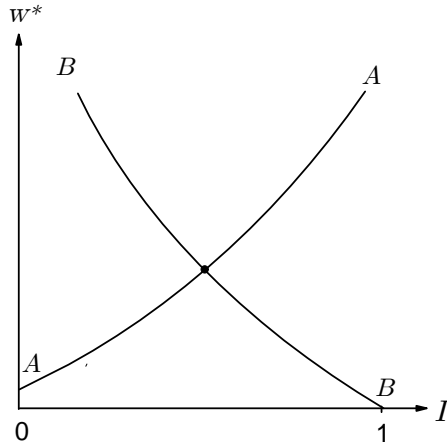


Figure 3: Free Trade Equilibrium

if small in equilibrium, each can overcome any coordination failure by becoming large when the opportunity presents itself. Moreover, there is no danger that a country will lack the resources to accommodate a particular industry at the scale required to meet world demand. In a world of many industries, countries can adjust on the extensive margin by hosting greater or fewer numbers of them, without sacrificing the benefits that come from concentrating an industry in a single location.

Second, the pattern of specialization conforms to the pattern of comparative advantage, when the latter is evaluated at a common scale of production. We have assumed in our formulation (as in the standard approach) that countries share the same scale economies, which are a property of the industry and not the place of production. Countries differ only with respect to a set of Hicks-neutral (and scale neutral) technology parameters. Then, with no risk of coordination failure, the country that enjoys the relative technological advantage in an industry will capture that industry—and its associated scale economies—in the general equilibrium. All that remains to be determined, as in DFS, is the margin between industries that locate at home and those that locate abroad. Indeed, the DFS model is a special case of ours with $A_i(x_i) \equiv 1$ for all i . Here, as there, the dividing line in the chain of comparative advantage is determined by the relative sizes of the countries, the pattern of world demand, and the pattern of technologies in the various industries.

Note that country size plays no role in determining relative productivity. Since every individual industry is small in relation to the size of either economy, any country can produce enough of any good to meet the pressures of world demand. The relative size of the two countries affects the margin I between the set of each country’s import goods and its export goods, but it cannot affect the ordering of goods in those sets.

In short, by incorporating many industries and amending firms’ perceptions about their

own productivity, we have overturned several conclusions from the canonical analysis. Multiple equilibria are not pervasive and “natural” comparative advantage rules.

A pedagogically interesting special case arises when the countries share identical technologies and differ only in labor supplies. This setting has been used by some authors to argue that trade induced by external economies will not equalize factor prices in a pair of countries that differ only in size.¹² However, this conclusion also rests on the assumption that individual industries are large in relation to the sizes of the national economies. In our model, when $a_i = a_i^*$ for all i , if $w^* \neq 1$, then all industries would locate in the low-wage country. This is not consistent, of course, with labor-market clearing in the other country, so factor prices must be equalized. With equal factor prices and identical technologies, the pattern of specialization is not determined. What is determined is only the aggregate employment levels in the two countries which, in equilibrium, matches the exogenous labor supplies. A small country can employ its labor force in equilibrium by producing and exporting a moderate number of goods produced in relatively small quantities or a smaller number of goods that require greater numbers of workers to meet world demand. In any case, differences in country size are accommodated by the number of goods produced in each country and not by the identities of which goods are produced where.

3 Efficiency and the First-Best

Both in autarky and in a trade equilibrium, efficiency requires equality between the marginal rate of substitution and the marginal rate of transformation for every pair of goods. In our setting, utility-maximizing consumers equate as usual the marginal rate of substitution between goods to the relative price. But, in general, the marginal rate of transformation is not equal to the relative price due to the presence of production externalities. Therefore, neither the autarky equilibrium nor the free-trade equilibrium achieves an efficient allocation of labor.

To see this, note that the marginal product of labor in industry i in the autarky equilibrium is $A_i(\tilde{x}_i)/a_i(1 - \theta_i)$.¹³ It follows that the autarky marginal rate of transformation between good i and good j is

$$\widetilde{MRT}_{ij} = \left[\frac{1 - \theta_i(\tilde{x}_i)}{1 - \theta_j(\tilde{x}_j)} \right] \frac{a_i/A_i(\tilde{x}_i)}{a_j/A_j(\tilde{x}_j)} = \left[\frac{1 - \theta_i(\tilde{x}_i)}{1 - \theta_j(\tilde{x}_j)} \right] \frac{\tilde{p}_i}{\tilde{p}_j},$$

where the second equality follows from (3); i.e. from price equals average cost. Therefore, the autarky relative outputs of goods i and j can be efficient only if it happens that $\theta_i(\tilde{x}_i) = \theta_j(\tilde{x}_j)$. Similarly, the free-trade relative outputs can be efficient only if $\theta_i(x_i) = \theta_j(x_j)$. With relative

¹²See, for example, Markusen and Melvin (1981) and Ethier (1982a).

¹³Let L_i be employment in industry i . Then $x_i = A_i(x_i)L_i/a_i$, so $dx_i = (A_i/a_i)dL_i + (A'_i/a_i)dx_i$ or

$$\frac{dx_i}{dL_i} = \frac{A_i}{a_i} \left(\frac{1}{1 - \theta_i} \right).$$

prices equal to relative average costs, they will not generally be equal to relative *marginal* costs. A set of Pigouvian subsidies could be used to achieve the autarky first best. A similar set of subsidies, together with a set of optimum tariffs, could do likewise for a country that trades.

It should be clear from the discussion that the autarky equilibrium achieves the first best in the special case in which all industries bear a constant and common degree of scale economies.¹⁴ If $\theta_i(x) \equiv \theta$ for all x and i , the equality between price and average cost in every industry ensures equality between relative prices and marginal rates of transformation. With trade, the first best still requires a set of optimal tariffs. But gains from (free) trade can be established for this special case in the usual way.¹⁵

4 Gains from Trade

To study the prospects for gains from trade with more general production technologies, we posit CES preferences for the home country. Domestic preferences are described by the utility function in (1). We make no assumptions about the weighting of the various goods in utility, about foreign preferences, or about the productivity functions $A_i(x_i)$. In particular, we allow for varying degrees of scale economies across industries and non-constant output elasticities in any or all of them.

4.1 Elastic Demands

We begin with $\sigma \geq 1$, which generates demands that are price elastic. Let us suppose that there are losses from trade; i.e., $U < \tilde{U}$. Since the home wage is numeraire in both the autarky and trade equilibria, $U < \tilde{U}$ requires $P > \tilde{P}$; i.e., a lower *real* income with trade than without implies a higher price index.

In Figure 4, we show the autarky equilibrium in some industry i . It is found at \tilde{E} , the intersection between $\tilde{D}\tilde{D}$ and $\tilde{C}\tilde{C}$, and has

$$\tilde{x}_i = \frac{b_i^\sigma \tilde{P}_i^{-\sigma} L}{\tilde{P}^{1-\sigma}}$$

as we have noted before (see Equation 2).

¹⁴See Chipman (1970) for an, early, formal demonstration of this point.

¹⁵For the case of constant and common elasticities θ , the free-trade output vector maximizes the value of domestic output given prices. We can prove gains from trade using the standard, revealed-preference argument: The value of free-trade consumption vector at the prices of the trade equilibrium equals the value of the free-trade output vector at those prices, which exceeds the value of autarky output vector at those prices, which in turn equals the value of the autarky consumption vector at those prices. So, consumers could buy the autarky consumption bundle in the trade equilibrium with the income available to them. If they chose not to do so, it must be that they prefer the trade consumption bundle. Notice that this argument makes no assumptions about demand.

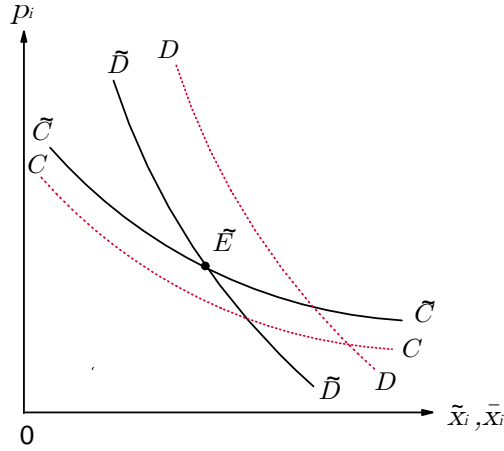


Figure 4: Hypothesized Industry Equilibrium Assuming Losses from Trade

In a world with trade, producers of good i face the world demand

$$\frac{b_i^\sigma p_i^{-\sigma} L}{P^{1-\sigma}} + c_i^*(\mathbf{p}, w^* L^*),$$

where $c_i^*(\mathbf{p}, w^* L^*) > 0$ is the foreign demand for good i in the trade equilibrium (an arbitrary function of the vector of prices \mathbf{p} and foreign income, $w^* L^*$). The hypothesis that $P > \tilde{P}$ implies that $P^{1-\sigma} < \tilde{P}^{1-\sigma}$ under the assumption that demand is elastic ($\sigma \geq 1$). Then, the substitution effect of the higher price level outweighs the income effect on the demand for good i , so the inverse demand curve with trade— DD in the figure—lies to the right of $\tilde{D}\tilde{D}$ under the maintained hypothesis. The average cost curve with trade is depicted by CC and is given by

$$\min \left[\frac{a_i}{A(\bar{x}_i)}, \frac{w^* a_i^*}{A(\bar{x}_i)} \right]$$

where \bar{x}_i again is the total world output of good i . This curve coincides with $\tilde{C}\tilde{C}$ for $i \leq I$; i.e., for goods that are produced by the home country in the trade equilibrium. It lies below $\tilde{C}\tilde{C}$ (as depicted in Figure 4) for goods that are produced abroad ($i > I$); these goods are imported by the home country only because their average cost is less than it would be with domestic production. In either case, the intersection of DD and CC lies below and to the right of \tilde{E} .

As the figure makes clear, our hypothesis that $P > \tilde{P}$ implies $p_i < \tilde{p}_i$ for an arbitrary good i , be it one that is exported in equilibrium or one that is imported. Therefore, all prices are lower in the trade equilibrium than in the autarky equilibrium (relative to the domestic wage). But this is a contradiction, because the price index cannot rise if each element in the index falls. We conclude that $P < \tilde{P}$ and thus $U > \tilde{U}$; i.e., the home country gains from trade!

Since $P < \tilde{P}$ and $\sigma \geq 1$, the inverse demand curve DD may, in fact, be to the left or to

the right of $\tilde{D}\tilde{D}$. Apparently, aggregate output in some industries might fall as the result of the opening of trade. Nonetheless, the impact of these cases cannot be so great as to negate the overall gains that come from international specialization according to comparative advantage and the realization of greater scale economies due to the expansion of the market.

4.2 Inelastic Demands

Next consider the case of $\sigma < 1$; i.e., inelastic demands. Since there are gains from trade as σ approaches one from below¹⁶, and the model is continuous in the parameter σ , trade losses for some σ would imply the existence of a $\hat{\sigma}$ such that $P(\hat{\sigma}) = \tilde{P}(\hat{\sigma})$, in the obvious notation. Let us suppose this to be true, and consider further the autarky and trade equilibria that would arise when $\sigma = \hat{\sigma}$.

The demand function with trade and the price-equals-average cost relationship (4) imply

$$\begin{aligned} \frac{\bar{x}_i}{A_i(\bar{x}_i)} &= \frac{b_i^{\hat{\sigma}} p_i^{1-\hat{\sigma}} L}{P^{1-\hat{\sigma}} \min[a_i, a_i^* w^*]} + \frac{c_i^*(\mathbf{p}, w^* L^*)}{A_i(\bar{x}_i)} \\ &> \frac{b_i^{\hat{\sigma}} p_i^{1-\hat{\sigma}} L}{P^{1-\hat{\sigma}} a_i}, \end{aligned}$$

where the inequality in the second line follows from the fact that $a_i \geq \min[a_i, a_i^* w^*]$ and $c_i^*(\mathbf{p}, w^* L^*) > 0$. Now suppose that $p_i > \tilde{p}_i$. Then, with $\hat{\sigma} < 1$ and $P = \tilde{P}$,

$$\frac{b_i^{\hat{\sigma}} p_i^{1-\hat{\sigma}} L}{P^{1-\hat{\sigma}} a_i} > \frac{b_i^{\hat{\sigma}} \tilde{p}_i^{1-\hat{\sigma}} L}{\tilde{P}^{1-\hat{\sigma}} a_i} = \frac{\tilde{x}_i}{A_i(\tilde{x}_i)}.$$

Therefore $\bar{x}_i/A_i(\bar{x}_i) > \tilde{x}_i/A_i(\tilde{x}_i)$. But the function $x_i/A_i(x_i)$ is increasing in x_i by the assumption that $\theta_i(x_i) < 1$. So the string of inequalities implies $\bar{x}_i > \tilde{x}_i$, which in turn implies $p_i < \tilde{p}_i$. This contradicts our supposition that $p_i > \tilde{p}_i$. We conclude that $p_i < \tilde{p}_i$ for all i , which in turn implies $P < \tilde{P}$. It follows that there can exist no $\hat{\sigma}$ for which the price index under autarky is the same as the price index with trade. Again, the home country gains from trade!

Figure 5 illustrates the autarky and free-trade equilibrium in a typical industry i when $\sigma < 1$. With $P < \tilde{P}$ and $\sigma < 1$, the DD curve necessarily lies to the right of the $\tilde{D}\tilde{D}$ curve. The CC curve is identical to $\tilde{C}\tilde{C}$ for $i \leq I$ and strictly below it for $i > I$. It follows that $\bar{x}_i > \tilde{x}_i$ for all i . In this case, trade expands the world output of every good. All prices fall relative to the home

¹⁶The limiting case of the CES as $\sigma \rightarrow 1$ is equivalent to Cobb-Douglas preferences. Gains from trade for these preferences can be established directly. With a constant spending share b_i on good i in the home country,

$$\frac{\bar{x}_i}{A(\bar{x}_i)} = \frac{b_i L}{\min[a_i, w^* a_i^*]} + \frac{c_i^*(\mathbf{p}, w^* L^*)}{A(\bar{x}_i)} > \frac{b_i L}{a_i} = \frac{\tilde{x}_i}{A(\tilde{x}_i)}.$$

But this implies $\bar{x}_i > \tilde{x}_i$ for all i and therefore $p_i < \tilde{p}_i$ for all i . The home worker's real income increases in terms of every good.

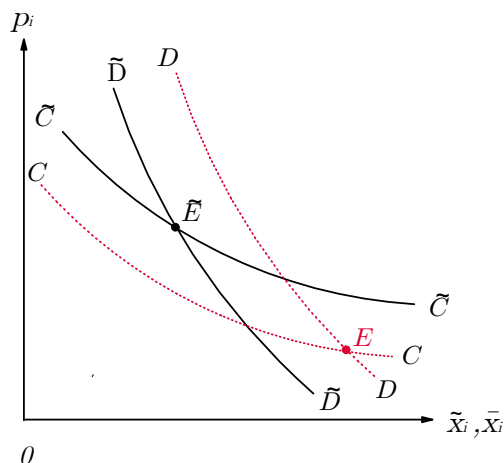


Figure 5: Autarky and Trade Equilibrium for $\sigma < 1$

wage, and domestic workers benefit as a result.

We have established gains from trade for the home country when its residents have CES preferences, independent of the elasticity of substitution, the pattern of comparative advantage, the pattern of scale economies, and the details of foreign preferences. This goes some way toward overturning the “paradoxical” welfare implications that derive from the standard analysis, and the discomfort generated therefrom.

5 Concluding Remarks

The extensive literature on external economies typically posits a competitive environment with many small firms. Researchers have used the small size of producers in a competitive equilibrium to justify an assumption that firms neglect the impact of their output decisions on aggregate scale and productivity. We have argued that it is not the equilibrium size of a firm that matters for its assessment of productivity, but rather the nature of industry competition and, in particular, whether a firm can foresee growing large as a result of its actions. When industry participants believe that they can grow large, the potential for internalization is enough to coordinate locational decisions even when externalities remain pervasive in equilibrium.

The literature also typically focuses on an environment with only two industries. Often the assumption is made for convenience, although it might be appropriate if external scale economies spread throughout broad sectors. If, instead, external economies are generated by specific activities, a model with many activities may be more appropriate. Then, country size no longer plays the role that it does in the standard analysis, because a small country can accommodate all of some activities so long as it hosts relatively few of them.

These observations may have more general applicability than our simple Ricardian model. For example, Grossman and Rossi-Hansberg (2008) study task trade between similar countries. They assume that firms perform a continuum of tasks to generate output and that productivity in each task depends on how frequently it is performed by all firms in that same location. In that setting, a firm can grow large by serving as a potential supplier of a task to others. Although no outsourcing takes place in equilibrium, the potential for same is enough to coordinate locational decisions for each task. Moreover, the fact that each task is small means that any task might be concentrated in a small country.

Another relevant application can be found in urban economics. The existence of cities seems evidence of agglomeration economies that are, at least partially, external to the firm. As in our setup, a site can have small or zero productivity if no firm locates there, but could be as productive as New York if ten million people reside nearby. Multiple equilibria are pervasive in models of systems of cities, as sites may or may not host cities depending on whether agents coordinate to live there. To address this problem, the urban economics literature has invoked the actions of large ‘city corporations’ (Henderson, 1974) that can subsidize factors to encourage the internalization of externalities and thereby coordinate location decisions.¹⁷ Our theory provides an alternative mechanism for coordination in naturally productive sites. If firms engage in Bertrand competition, each will recognize that by locating alone in a “good” site it can potentially achieve efficient scale as a “one-company town.” Such a producer might still be small in relation to the aggregate economy, if the number of towns or cities is large. In equilibrium, all towns may house numerous firms, yet the equilibrium allocation of activities across space will be unique. In short, our conclusions about trade with many industries may apply as well to agglomeration with many cities.

Our paper demonstrates that the size of industries (or cities) and the nature of the competition ought to be carefully considered in future modeling of external scale economies.

¹⁷Durable housing investments can also serve as a coordinating mechanism in dynamic models, as shown by Anas (1992) and Henderson and Venables (2008).

References

- [1] Anas, Alex (1992), "On the birth and growth of cities: Laissez-faire and planning compared," *Regional Science and Urban Economics* 22, 243-258.
- [2] Antweiler, Werner and Trefler, Daniel (2002), "Increasing Returns and All That: A View from Trade," *American Economic Review* 92, 93-119.
- [3] Bhagwati, Jagdish N. (1971), "The Generalized Theory of Distortion and Welfare," in J.N. Bhagwati et al., eds., *Trade, Balance of Payments and Growth: Papers in International Economics in Honor of Charles P. Kindleberger* (Amsterdam: North-Holland).
- [4] Caballero, Ricardo J. and Lyons, Richard K. (1989), "The Role of External Economies in U.S. Manufacturing," NBER Working Paper No. 3033.
- [5] Caballero, Ricardo J. and Lyons, Richard K. (1990), "Internal versus External Economies in European Industry," *European Economic Review* 34, 805-826.
- [6] Caballero, Ricardo J. and Lyons, Richard K. (1992), "The Case for External Economies," in A. Cukierman, Z. Hercowitz and L. Leiderman, eds., *Political Economy, Growth, and Business Cycles* (Cambridge MA: The MIT Press).
- [7] Chan, Vei-Lin , Chen, Been-Lon and Kee-Nam Cheung (1995), "External Economies in Taiwan's Manufacturing Industries," *Contemporary Economic Policy* 13, 118-130.
- [8] Chipman, John S. (1965), "A Survey of the Theory of International Trade: Part 2, The Neo-Classical Theory," *Econometrica* 33, 685-760.
- [9] Chipman, John S. (1970), "External Economies of Scale and Competitive Equilibrium," *Quarterly Journal of Economics* 84, 347-363.
- [10] Choi, Jai-Young and Yu, Eden S.H. (2002), "External Economies in the International Trade Theory: A Survey," *Review of International Economics* 10, 708-728.
- [11] Dornbusch, Rudiger, Fischer, Stanley, and Paul A. Samuelson (1977), "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods," *American Economic Review* 67, 823-839.
- [12] Ethier, Wilfred J. (1982a), "Decreasing Costs in International Trade and Frank Graham's Argument for Protection," *Econometrica* 50, 1243-1268.
- [13] Ethier, Wilfred J. (1982b), "National and International Returns to Scale in the Modern Theory of International Trade," *American Economic Review* 73, 389-405.

- [14] Graham, Frank D. (1923), "Some Aspects of Protection Further Considered," *Quarterly Journal of Economics* 37, 199-227.
- [15] Grossman, Gene M. and Rossi-Hansberg, Esteban (2008), "Task Trade Between Similar Countries," Princeton University.
- [16] Helpman, Elhanan (1984), "Increasing Returns, Imperfect Markets, and Trade Theory," in P.B. Kenen and R.W. Jones, eds., *Handbook of International Economics, Vol. 1* (Amsterdam: North Holland).
- [17] Helpman, Elhanan and Krugman, Paul R. (1985), *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy* (Cambridge MA: The MIT Press).
- [18] Henderson, J Vernon (1974), "The Sizes and Types of Cities," *American Economic Review* 64, 640-656.
- [19] Henderson, J Vernon and Venables, Anthony J. (2008), "The Dynamics of City Formation: Finance and Governance," mimeo Brown University.
- [20] Henriksen, Espen R., Midlefarth Knarvik, Karen Helene, and Steen, Frode (2001), "Economies of Scale in European Manufacturing Revisited," CEPR Discussion Paper No. 2896.
- [21] Kemp, Murray C. (1964), *The Pure Theory of International Trade and Investment* (Englewood Cliffs: Prentice Hall).
- [22] Krugman, Paul R. (1995), "Increasing Returns, Imperfect Competition, and The Positive Theory of International Trade," in G.M. Grossman and K. Rogoff, eds., *Handbook of International Economics, Vol. 3* (Amsterdam: North Holland).
- [23] Markusen, James R. and Melvin, James R. (1981), "Trade, Factor Prices, and the Gains from Trade with Increasing Returns to Scale," *Canadian Journal of Economics* 14, 450-469.
- [24] Marshall, Alfred (1879), *The Pure Theory of Foreign Trade*. Reprinted by the London School of Economics and Political Science, 1930.
- [25] Marshall, Alfred (1890), *Principles of Economics* (London: Macmillan and Co.).
- [26] Matthews, R.C.O. (1949-1950), "Reciprocal Demand and Increasing Returns," *Review of Economic Studies* 37, 149-158.
- [27] Melvin, James R. (1969), "Increasing Returns to scale as a Determinant of Trade," *Canadian Journal of Economics* 3, 389-402.

- [28] Ohlin, Bertil (1933), *Interregional and International Trade* (Cambridge MA: Harvard University Press).
- [29] Segoura, Io (1998), "Return to Scale and External Economies: Empirical Evidence from Greek Two-Digit Manufacturing Industries," *Applied Economic Letters* 5, pp.485-490.
- [30] Wilson, Charles A. (1980), "On the General Structure of Ricardian Models with a Continuum of Goods," *Econometrica* 48, 1675-1702.