I. Alternatives to Comparative Advantage — Economies of Scale

The fact that the largest share of world trade consists of the exchange of similar (manufactured) goods between similar (industrial) countries motivated interest in explanations of trade other than comparative advantage. What alternative candidates are there? There is in fact a very old one: increasing returns to scale. The fundamental idea behind comparative advantage is that countries trade in order to exploit their differences. Another possibility is that they might trade in order to specialize, that is, to become more productive by doing less but doing it better. Of course the exploitation of differences quite likely involves some specialization, so it is not completely clear that the two bases for trade are in fact distinct. Formal models help here.

a Types of scale economies

In practice these occur in great variety, so a classification of the more important attributes is useful.

*Internal vs. external* (to the firm). Internal require consideration of imperfect competition, so start with external and assume perfect competition.

*National vs. international.* Economies of scale may depend on the scale of operations within a nation (e.g., large plant size) or on the scale of operations globally (e.g., division of labor and free trade in intermediate goods). Either type might be either internal or external to the firm.

*Aggregative vs. dis-aggregative.* Increasing returns to scale may be a property of manufacturing generally or of individual manufactured goods. The former can be modeled with a single homogeneous good possessing scale economies, but the latter requires a large number $n$ of distinct goods, each by itself characterized by increasing returns to scale.

These three considerations generate eight types of scale economies, each relevant to the world in which we live. Consider first national, aggregative, increasing returns external to the firm.
b National, aggregative economies of scale external to the firm

Increasing returns to scale can obviously furnish a basis for trade and specialization not related to autarky price differences. To illustrate, consider a simple model in which there are two identical economies with two-good Ricardian technologies, except that the $B$-good sector has national, aggregative IRS that are external to the firm. $A$ is produced with CRS, and we measure units so that one unit of labor is required in each country to produce one unit of $A$. Each country's technology is summarized by:

$$L_A + L_B = L \quad \quad \quad A = L_A \quad \quad \quad B = k(L_B)L_B \quad \text{where: } k(L_B) = k_0 L_B^{\alpha - 1} \text{ for } \alpha > 1.$$

The individual firm takes $k$ as a parameter and so acts as though there were CRS. (NOTE: If $\alpha = 1$ the model reduces to the simple Ricardian model with $a_A = 1$ by choice of units and $a_B = 1/k_0$.) For the economy as a whole, have the PPF:

Since the IRS are external to the firm, in equilibrium price = average cost, so that $PB = L_B$, where $A$ is numeráire (so the wage = 1). Thus $P = [L - L_A]/B$ (= LE/ED in the figure, if D is autarky equilibrium.)

If there are two such identical countries, there is no basis for comparative advantage trade, since relative autarky prices will be the same. But there is obviously a basis for trade: with IRS in the $B$ industry it is not globally efficient for both countries to produce both goods. (Show this geometrically.)

It is clear that the no-trade case, with each country continuing to do what it had done in autarky even after all trade barriers are gone, is indeed a free-trade equilibrium. But with scale economies this cannot be Pareto-optimal.
Though the no-trade situation is clearly an equilibrium, one soon suspects that it may not be a very stable one. For suppose the home economy produces a few more \( B \), and the foreign economy fewer, than in the no-trade equilibrium. Then the home \( k \) is greater than \( k^* \), the foreign analog, so home \( B \) firms can undersell their foreign rivals, while the two countries can still produce \( A \) on equal terms. Thus the new situation is not an equilibrium. Furthermore, one would expect home \( B \) producers to increase their market share at the expense of their foreign competitors, so that foreign resources move into the \( A \) sector. If so, this development causes \( k \) to rise further and \( k^* \) to fall, increasing the home \( B \) advantage still more. This process must continue until a new equilibrium is reached with the home economy producing only \( B \) and/or the foreign economy producing only \( A \). Thus there is a second equilibrium. Indeed there is more than that: since the two countries are identical we can find a third equilibrium by simply reversing the home and foreign roles in the second. Thus we have two mirror-image equilibria with at least one country completely specialized separated by an equilibrium where both countries produce both goods. Our dynamic argument suggests that the latter is unstable and the former stable, though of course such dynamics reflect an arbitrary description of disequilibrium behavior. The basic idea behind comparative advantage is that countries should do what they can do relatively well; this implies some particular role in the world economy. Scale economies on the other hand require countries to concentrate on a small number of tasks; who does what is secondary. Thus scale economies introduce a bias towards a multiplicity of equilibria. This is our second major result.

I have assumed the two countries completely identical because that is the most dramatic way to demonstrate that scale economies provide a basis for trade independently of comparative advantage. But even if tastes and sizes differ, it will be possible to find a separating equilibrium with both countries producing both goods, provided it is possible to divide evenly an equilibrium quantity of world \( B \) production between the two countries. Only if they differ sufficiently in size and if equilibrium world demand for \( B \) is sufficiently great will this not be possible. (The assumption of identical countries made the point that IRS was a basis for trade independent of comparative advantage, so now we no longer need it. Will still take identical countries as our point of departure, but we will also allow the countries to differ in \( k_{x} \), in \( a \), in \( L \) and in preferences.)
The other equilibria might involve one country completely specialized in $B$ and/or one completely specialized in $A$. It makes a great deal of difference which occurs. Consider an equilibrium in which the home country specializes in $B$ production, with the (identical) foreign economy still producing both goods. Since $k > k^*$, foreign wages must be lower than home wages for this to be an equilibrium. As the foreign economy adjusted to the new equilibrium shown in the figure, its budget line swung inward, so its real income must have declined. The international price line must have become steeper, in order to support an equilibrium with production of both goods above the autarky point: good $B$ must become relatively more expensive. As the home economy adjusted to this equilibrium, it jumped down and to the right along its production possibility frontier to the $B$ axis, so that its now steeper budget line, pivoting on the PPF's endpoint, swung out. Thus the home economy was made better off. Thus trade has benefitted the home economy, relative to autarky, but has made the foreign worse off. There is quite likely a "mirror image" equilibrium in which the roles of the two countries reverse (this is obviously true if they are identical). Thus potential international conflict is inherent. This is our third major result. This comes about in any equilibrium in which the diversified economy produces some $B$ but less than in autarky. With identical economies it is associated with a large world demand in equilibrium for $B$, so that one country alone cannot satisfy it. But, especially with dissimilar countries, it is easy to construct examples in which either the larger or the smaller could be a loser from trade, and in which a diversified smaller country gains.

Even with identical countries, both can gain in a Graham equilibrium if preferences are such that the movement from autarky to free trade produces a large enough shift to $B$ production that the country with the smaller $B$ sector still has a larger one than in autarky. But it is still true that the countries fare differently: the country with the larger $B$ sector is better off.
Suppose now that, instead, the dynamic adjustment ends with the foreign economy specialized in $A$ and the home economy producing both goods. Since both countries produce $A$, wages must be the same internationally, in sharp contrast to the previous case. Thus residents of both countries fare the same. Home residents must be better off than in autarky, for the same reason as in the previous case, and foreign residents will also be better off if and only if the home $B$ industry is larger with free trade than the foreign $B$ industry was in autarky, since in that case the foreign budget line must pivot out. There will again be another “mirror image” equilibrium, as long as the foreign economy is larger than is necessary to produce the world output of $B$, but unlike the previous case this is of no consequence, because everyone fares the same regardless of country of residence. With identical economies, a factor-price equalization equilibrium is associated with a large world demand in equilibrium for $A$, so that one country alone cannot satisfy it. But it is easy to construct examples in which either the larger or the smaller country specializes in $A$, and in which both countries lose relative to autarky.

Even with identical countries, both can lose in a factor-price equalization equilibrium if preferences are such that the movement from autarky to free trade produces a large enough shift to $A$ production that the country with the $B$ sector still has a smaller one than in autarky. But the two countries are still equally well off: everyone loses from the move from autarky to free trade, which exacerbates the externality.

The final possibility is that both countries specialize. If the specialization outputs are consistent with an international equilibrium, it matters not that scale economies are involved in $B$ production. (For example, the model will be identical to a Ricardian model with $k$ fixed at the equilibrium level.) In this sense scale economies don't matter. Note that in this case the international equilibrium is Pareto-optimal, unlike both of the other cases, where too little $B$ is produced.

To summarize, with scale economies there are four potential types of equilibria. (i) Equilibria in which both countries produce both goods. We might call this the knife-edge case since there is some reason to regard such equilibria as unstable. In this case world production is inefficient, residents of both countries have identical free-trade consumption possibilities, and either country may be better or worse off than in
autarky. (ii) Equilibria in which one country specializes in $B$ production and the other produces both goods. In this case the specialized country gains from trade and the other might well lose, and there can easily be “mirror image” equilibria in which the countries reverse roles. We might call this the Graham case since it corresponds to Frank Graham's (1923) argument for protection. (iii) Equilibria in which one country specializes in $A$ production and the other produces both goods. We might call this the factor-price equalization case since wages must be equal in the two countries. In this case the possibility of multiple equilibria may be of little real interest because the two countries fare the same, even if both lose relative to autarky. (iv) Equilibria in which both countries specialize. We might call this the Ricardian case because scale economies really make no difference.

These types of equilibria are not mutually exclusive. That is, if tastes, technology and size imply multiple equilibria, the equilibria could be (and usually will be) of different types.

Knife-edge equilibria are of perhaps little practical relevance, whereas Ricardian equilibria are simple to analyze but add little of real importance to the understanding of the world economy beyond what can be derived from models without scale economies at all. It is the possibility of factor-price equalization and (especially) Graham equilibria that produce the real value added that can come from consideration of increasing returns to scale. These equilibria can produce positive and normative implications in sharp contrast to those of comparative advantage and can therefore be used in support of quite different policy recommendations. They are of direct relevance to the "grand old debates" over the wisdom of participating in the international trading system.

But this analysis of national external economies of scale is less than fully satisfying, for several reasons, and has accordingly had to play a role very much subservient to that of comparative advantage. For one thing, the indeterminacy of results due to the likelihood of multiple equilibria renders the theory cumbersome to use. Also, the “grand old debates” may be important, but they are not all-important, and our investigation of scale economies was motivated in large part by a desire to address more directly a world in which the lion's share of trade consists of the exchange of similar commodities between similar economies. But the influence of scale economies, enhancing the possibility of specialization, is to move us in just the opposite direction! Finally, it is unsatisfying to confine our analysis of scale economies to those that are external to the firm but internal to the national industry. Additional methods of modeling scale economies are needed.

c Dis-aggregative economies of scale

The above theory had a tendency for multiple equilibria, complete specialization and uncertain welfare properties. These features, and the general inconvenience of Graham equilibria, caused scale economies
to be incorporated into trade models much less frequently than most economists probably thought that
their inherent importance warranted. But this could also motivate looking at other types of scale
economies.

Suppose now that we maintain the same model as above (so that the scale economies are external to the
firm and we can assume perfect competition), except that the $B$ sector now consists of $n$ distinct varieties,
$B_i$, each with the technology described above. Note the following concerning international trade in such a
model.

(i) Knife-edge equilibria (those to which our intuitive instability argument apply) arise whenever both
countries produce any two goods (either two varieties of $B$, or $A$ and one variety of $B$) in common.

(ii) Graham equilibria emerge whenever the two countries produce some variety of $B$ in common but in
different amounts. The country with the larger production must have the higher wage and therefore
cannot be producing any $A$.

(iii) Factor-price equalization equilibria will again feature both countries producing some $A$, but no
variety of $B$ in common, so that both are equally well off with free trade. If one country specializes in $A$
all trade will consist of the inter-industry exchange of $A$ for $B$, but if both countries produce some
varieties of $B$ there will also be an intra-industry exchange of $B$ varieties. We would expect the latter to
be relatively more important the more similar the two countries are; If the two are exactly alike, there
will be an equilibrium with only intra-industry trade.

(iv) Ricardian equilibria, with each good and variety produced in exactly one location, may likewise
involve inter-industry and/or intra-industry trade.

Although we have the same classes of equilibria as before, they no longer have the same plausibility. If
the two countries have the same technology, then, as long as the demand for each variety of $B$ is not too
large, any pattern of world demand can be accommodated by some factor-price equalization equilibrium.

**d International economies of scale**

External economies have generally been identified in the literature with an increased division of labor
made possible by a larger market: Adam Smith's pin factory and the Swiss watch industry are the
prominent hoary examples. Less common are examples having to do with a larger volume of public
information generated by a larger industry. In principle none of these requires an industry to be
physically located in one place. A dispersed industry can realize a great division of labor if intermediate components can be shipped from place to place; public information can be dispersed within the industry if communication is efficient. What matters, under these conditions, is the global size of the industry, not its geographical concentration.

This suggests that the returns to scale depend upon the size of the world industry, not the national industry. This is what is meant by international returns to scale. Suppose that resources are used to produce $A$ and $m$. $A$ production is characterized, as in the above one factor model, by CRS. $m$ is an index of the scale of operations of the national B industry, subject to increasing returns to scale. With national returns to scale, national B production $B$ is related to $m$ by

$$B = km$$

where $k = k(m), k' > 0$.

With international returns to scale, on the other hand, we have instead

$$B + B^* = k(m + m^*)$$

where $k = k(m + m^*), k' > 0$.

Here an asterisk refers to the foreign country. Of course one might argue that, even in the modern world economy, some determinants of external economies require geographical concentration, so the appropriate general formulation would have $k = k(m, m^*)$ and $k^* = k^*(m, m^*)$, with perhaps the additional requirement that $k$ be relatively more sensitive to $m$ and $k^*$ to $m^*$. But the exposition will be more effective with pure international scale effects.

At first glance it might seem that we have complicated matters enormously. National production possibility frontiers between final goods are not even defined, because productivity in each country's B industry depends upon the size of the other country's B industry. But the situation becomes almost transparent as soon as we focus on patterns of resource allocation rather than on goods.

To see why this is so, consider the world production possibility frontier between $A$ and $B$. This is certainly well defined. A point on it can be found by maximizing world $B$ production for a given feasible volume of world $A$ production, that is, by choosing $m, m^*$ to

maximize: $B + B^* = k(m + m^*)[m + m^*]$  
subject to: $T(m) + T^*(m^*) = \text{some specified value}$. 
Here $T$ and $T^*$ denote the home and foreign production possibility frontiers between $A$ and $B$-resources. It is immediately clear that $B + B^*$ will be maximized by maximizing $m + m^*$: this problem has exactly the same solution as that of choosing $m, m^*$ to

$$\begin{align*}
\text{maximize: } & m + m^* \\
\text{subject to: } & T(m) + T^*(m^*) = \text{some specified value}
\end{align*}$$

and the solutions to problems of the latter sort are just the comparative advantage predictions. That is, efficient patterns of world activity in $A$ and $B$ correspond to efficient patterns in $A$ and $B$-resources, ignoring the scale economies.

Productive efficiency is as in the constant returns model, and firms behave competitively because the economies are external to them. The result is that the complex tendencies associated with Graham equilibria when scale economies are national disappear when they become international.

The second major implication of international economies of scale is that they imply a theory of the intra-industry exchange of intermediate goods between relatively similar economies. The essential idea behind international returns is that a dispersed industry can realize the benefits of a large division of labor if intermediate goods can be shipped within the industry. Thus the more nearly equal in size $m$ and $m^*$ are, the greater the volume of intra-industry trade in $B$ components.

With identical homothetic demands across countries, the pattern of inter-industry trade and specialization is determined in the familiar comparative advantage fashion. This inter-industry trade will comprise all trade if the disparity between countries is great enough for the $A$ exporter to specialize completely in $A$. Small international differences reduce the incentive for inter-industry trade but cause the integrated $B$ industry to be divided relatively evenly between countries, thereby inducing intra-industry trade. In the limiting case where the countries are identical, they will both be self sufficient in $A$. But they can gain from trade by establishing a single, rationalized $B$ industry; all trade will be intra-industry.