

10

Predecessors

Nothing is more usual, among states which have made some advances in commerce, than to look on the progress of their neighbors with a suspicious eye, to consider all trading states as their rivals, and to suppose that it is impossible for them to flourish, but at their expense. In opposition to this narrow and malignant opinion, I will venture to assert, that the increase of riches and commerce in any one nation, instead of hurting, commonly promotes the riches and commerce of all its neighbors; and that a state can scarcely carry its trade and industry very far, where all the surrounding states are buried in ignorance, sloth, and barbarism. (David Hume, "Of the Jealousy of Trade," *Essays* (1741–42))

[Opponents of support of new industry by government offer] the proposition, that industry, if left to itself, will naturally find its way to the most useful and profitable employment. Whence it is inferred that manufactures, without the aid of government, will grow up as soon and as fast as the natural state of things and the interest of the community may require.

Against the solidity of this hypothesis . . . very cogent reasons may be offered. These have relation to the strong influence of habit and the spirit of imitation; the fear of want of success in untried enterprises; the intrinsic difficulties incident to first essays towards a competition with those who have previously attained to perfection in the business to be attempted: the bounties, premiums, and other artificial encouragements with which foreign nations second the exertions of their own citizens, in the branches in which they are to be rivaled. (Alexander Hamilton, *Report on Manufactures*, 1791, p. 203)

Much of what has been described in this book builds upon the fundamental contributions of earlier writers. This chapter is intended as an acknowledgment of our debt to these predecessors.¹ Though a few early examples will be offered, we will attempt no comprehensive survey of the entire literature, but instead we will emphasize five strands in the literature that are most pertinent to our work: (1) the multiplicity of equilibria that arise when there are scale economies or

the costs of entry are substantial, (2) the tendency of the presence of scale economies to result in provision of any product by only a single supplier, (3) the effects of productivity growth in one country in classical models without scale economies, (4) the absence of certainty that the market mechanism will select among the set of potential equilibria one that is (or approximates) the social optimum, and (5) the structure of the formal models.

10.1 Scale Economies, Entry Costs and Multiple Equilibria

More than half a century ago, A. C. Pigou was among the first to explore the policy implications of multiple equilibria. In his groundbreaking *Economics of Welfare* there is a chapter that focuses on the case where there is more than one potential equilibrium. After showing how start-up costs can generate a multiplicity of stable equilibria, he writes:

Benefit might be secured by a *temporary* bounty [i.e., a subsidy] (or temporary protection) so arranged as to jerk the industrial system out of its present poise at a position of relative [i.e., local] maximum, and induce it to settle down again at the position of absolute maximum—the highest hill-top of all. This is the analytical basis of the argument for the *temporary* protection, or other encouragement of infant industries; and if the right infants are selected, the right amounts of protection accorded, and this protection removed again at the right time, the argument is perfectly valid. (1932, p. 141)²

In other words, Pigou emphasizes that where entry cost into an industry is substantial, more than just a single outcome is likely to be possible for the market mechanism. Moreover, some of the possible outcomes may be far inferior to the true optimum, “the highest hill-top of all.” And since an inferior outcome is apt to be stable, government intervention may be helpful or even required to free the economy from such an undesirable equilibrium.

Recognition and systematic examination of the idea that trade models with scale economies are characterized by multiple equilibria probably was first provided by Alfred Marshall. He examined the subject carefully in one of his earliest writings (1879), a piece initially printed for private circulation. Copies were sent to leading economists in the United Kingdom and other countries, and key portions were reproduced in several books before the monograph was published and became generally available in 1930. The international trade portion of the essay is 28 pages long, of which 16 pages deal primarily with the

case of multiple equilibria. Thus it is clear that the notion was not a passing thought nor one whose significance was unrecognized by the author.

His multiple-equilibrium discussion deals explicitly with international trade. Marshall's formal analysis is entirely geometric, and uses the now-familiar offer curve diagram as its tool. The analysis is consequently framed in terms of a two-country, two-good model. He shows that in the presence of scale economies, and only in this case, the offer curves can have several intersections. (Figures 10.1 and 10.2 reproduce Marshall's graphs for the two cases.) Figure 10.2 is the scale economies case. Marshall calls it the exceptional case, "in which an increase in the amount of wares which a country produces for exportation effects a very great diminution in the expenses at which she can produce them" (1879, pp. 5–6). In this graph we see the three equilibrium points, *A*, *B*, and *C*. Marshall argues that intersection *B* is an unstable equilibrium, while *A* and *C* are stable (see Marshall's arrows in his figures 8 and 9).

Marshall offers two interpretations of the scale economies case. One is the standard static concept entailing a negatively sloping average cost curve of the usual sort: "They would then represent a case in which the trade between the two countries could not grow up gradually; but could be carried on with profit to both if it were once started on a large scale by any external cause" (1879, p. 15). Marshall's other type of scale economies (like that described by Krugman and Venables 1992) entails

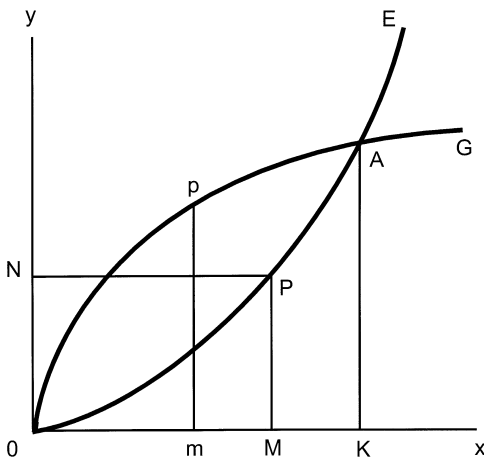


Figure 10.1
Marshall's two-country, two-good model (redrawn from the original)

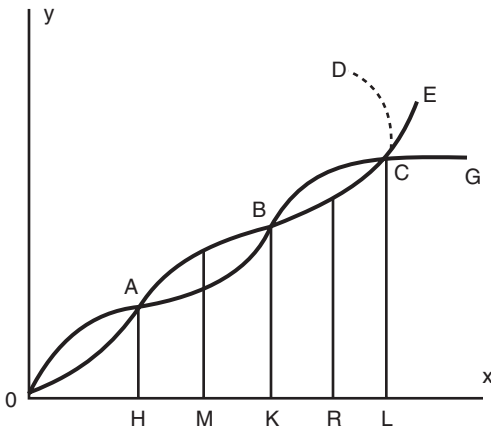


Figure 10.2
Marshall's scale economies model (redrawn from the original)

an introduction process that can require a substantial amount of time: "The introduction of the economies which were requisite in order to render possible such cases as this on a large scale have seldom been effected within a short space of time. The lapse of generations has been required for that development of England's invention and economies in manufacture which was above attributed in part to her export trade. . . . Special knowledge, special skill and special machinery are to a greater or less extent required for the manufacture of these implements" (1879, pp. 13–14).

Moreover economies of this second type may be at least partly irreversible:

[S]uppose that an increase in the amount of cloth produced for exportation leads to the introduction of extensive economies. Such economies when they have once been obtained are not readily lost. Developments of mechanical appliances, of division of labor, and of organization of transport, when they have once been effected are not readily abandoned. Capital and skilled labor which have once been devoted to any particular industry, may indeed become depreciated in value when there is a falling off in the demand for the wares which they produce; but they cannot quickly be converted to other occupations." (1879, p. 27)

It is surely fair to associate this second type of scale economies with the high cost of entry and the resulting retainability property that underlies much of the discussion of this book.

Since Marshall, a number of writers have made valuable contributions and additions to the basic analysis of multiplicity of equilibria in

international trade (see, e.g., Matthews 1949–50; Meade 1952; Kemp 1964, 1969; Chipman 1965; Ethier 1979; Helpman and Krugman 1985, pp. 209–213; Krugman 1991; Krugman and Venables 1992; Romer 1994). The earlier of these discussions generally employed geometric methods of analysis, and therefore found it convenient to confine the study to the two-good, two-country case. Like Marshall, they also generally examined a situation involving a very small number (frequently three) equilibria. Apparently, we are the first to have shown (1992, 1994) that multiple equilibria do not merely exist but that they increase exponentially with the number of traded commodities, and that they constitute a region with robust qualitative characteristics and substantial economic implications.

It has also long been recognized that some of the multiple equilibria that result from scale economies and high entry costs can have undesirable welfare properties. It even led Frank Graham to argue as early as 1923 that permanent protection by a country of an industry that excludes imports of that industry's products altogether can sometimes be advantageous to the country that does so.³ Marshall himself did not discuss the implications of multiple equilibria for the workings of the market mechanism and the reliability of the market's promotion of economic efficiency and the general welfare. This task was left to his successor, A.C. Pigou. We have already seen in the quotation from his *Economics of Welfare* earlier in this section how effectively he carried out that task. Pigou noted that where equilibrium is not unique market forces may well lead the economy to a local optimum that is far inferior (from the point of view of one, or more, or even all of the affected parties) to the true global maximum. He also noted that in such circumstances the public sector can play a useful role in improving the results yielded by the market mechanism, with either temporary or permanent intervention conceivably serving this purpose (*Economics of Welfare*, 4th ed., ch. 3, esp. p. 141). Thus Pigou clearly anticipated the general directions of our policy discussion.

The importance of multiple equilibria has also been clear to more recent writers: "The most interesting consequence of external economies is the existence of multiple equilibria," wrote John Chipman in 1965 (p. 749). But where such a multiplicity of candidate equilibria are present the market loses the manifest destiny to which a unique equilibrium would force it. There are then a number of possible end points, and as Graham pointed out three-quarters of a century ago, many of these equilibria can be quite undesirable, at least for one of the trading countries. Moreover, as Paul Romer (1994, p. 6) put it, in such circum-

stances, “We are forced to admit that the world as we know it is the result of a long string of chance outcomes.” Paul Krugman (1991, p. 652) adds, “Once one has multiple equilibria, however, there is an obvious question: which equilibrium actually gets established? . . . On the one side there is the belief that the choice among multiple equilibria is essentially resolved by *history*: that past events set the preconditions that drive the economy to one or another steady state. . . . On the other side, however, is the view that the key determinant of choice of equilibrium is *expectations*.”

Another branch of the literature, that focussed on innovation rather than international trade, has also emphasized that multiple equilibria can give historical accident an important role in the determination of the allocation of productive activities among countries (e.g., see David 1985, and the various recent writings of Brian Arthur, e.g., 1989). This analysis, too, raises the possibility that the market mechanism may conceivably lead to outcomes that do not most effectively promote the public interest, and that some of those outcomes may be at least partially irreversible, all of this implying that such considerations can open a useful economic role for government.

We see, then, that Krugman’s remark—“In the emerging literature on increasing returns and externalities, multiple equilibria are not a nuisance but a central part of the story”—is clearly justified (1991, pp. 651–52).

10.2 Scale Economies and Tendency to Single Suppliers (Perfectly Specialized Equilibrium)

We have made extensive use in this book of the case of perfectly specialized equilibria because, among other reasons, scale economies tend to elicit assignments in which any particular commodity is provided by only a single supplier. Since we are concerned with entire economies rather than single firms, in our discussion “single supplier” means exclusive production of the item in question within a single country, but it does not necessarily imply that the item is supplied by a monopoly firm. Nevertheless, the argument associating scale economies with production within a single economy is the same as the standard “natural monopoly” analysis that associates scale economies with supply of a good or service by a single enterprise.

As Viner often emphasized to one of us, the search for the first writer to have produced any particular idea is a hunt for a chimera. When-

ever a particularly early reference is discovered, it only serves to set off attempts to find earlier references, usually a bit less sophisticated but arguably on the point. Thus we will cite Cournot (1838, p. 76 of the English translation) as the author of a relatively early statement recognizing the incompatibility of scale economies and competition.⁴

The notion of “natural monopoly” seems to have been originated by J. S. Mill, though he only hints at its connection with scale economies:

All the natural monopolies (meaning thereby those which are created by circumstances, and not by law) which produce or aggravate the disparities in the remuneration of different kinds of labor, operate similarly between different employments of capital. If a business can only be advantageously carried on by a large capital, this in most countries limits so narrowly the class of persons who can enter into the employment, that they are enabled to keep their rate of profit above the general level. (*Principles of Economics*, bk. II, ch. 15, p. 394).

There is probably no way to trace the origin of the observation that equilibrium under scale economies tends to entail single-source supply. The idea is that where several suppliers coexist, the arrangement can be expected to be unstable because any supplier in such an initial state that succeeds in expanding its relative output will be able to underprice its rivals. This idea seems so obvious that it probably occurred to relatively early observers. Certainly Marshall and F. Y. Edgeworth treat the observation as a commonplace that hardly merits discussion. Though Marshall clearly devoted considerable space to his analyses of monopoly and economies of scale (both internal and external⁵), he seems only to have referred to our issue in one footnote. And there he treated it as a fairly clear-cut matter that some earlier writers had inexplicably overlooked:

Some, among whom Cournot himself is to be counted, have before them what is in effect the supply schedule of an individual firm; representing that an increase in its output gives it command over so great internal economies as much to diminish its expenses of production; and they follow their mathematics boldly, but apparently without noticing that their premises lead inevitably to the conclusion that, whatever firm first gets a good start will obtain a monopoly of the whole business of its trade in its district.” (*Principles*, p. 459, fn)

Similarly, while no mention of the subject occurs in Edgeworth’s noted article, “The Pure Theory of Monopoly” (1896) or in “The Laws of Increasing and Diminishing Returns” (1911, both reprinted in *Papers*

D), he covers the issue in one sentence in a book review of 1905: “The liability of an industry to be monopolized when it obeys the law of increasing returns creates peculiar difficulty in the application of the geometrical method to supply” (*Papers III*, p. 141).

Wicksell treats the subject with equal brevity (*Lectures*, pp. 131 and 233). He tells us: “If . . . the law of increasing returns applies without qualification. . . . The whole industry will be dominated by a more or less completely monopolistic association and all smaller concerns will disappear. . . . any attempt on the part of the smaller enterprise at effective competition . . . would be fruitless. . . . (p. 131).

Frank Knight (1924), too, emphasized the conflict between competition and scale economies, making it a central point in his criticism of Frank Graham’s 1923 contention (see above) that under scale economies a country may well obtain losses rather than gains from trade.

Viner, in his justly noted article, “Cost Curves and Supply Curves,” approaches the matter somewhat more extensively and in a different way:

The familiar proposition that net internal economies of large-scale production and long-run stable equilibrium are inconsistent under competitive conditions is clearly illustrated [by the cost-curve diagrams]. . . . [The firm] in short-run equilibrium when . . . its short-run marginal cost . . . is equal to price . . . will not be in long-run equilibrium, however, for its long-run marginal cost will then be . . . less than price.⁶ Provided that no change in its output will affect market price, it will pay this concern to enlarge its plant whatever the price may be, and whatever its existing scale of plant may be. If thereby it grows so large that its operations exert a significant influence on price, we pass out of the realm of atomistic competition and approach that of partial monopoly. . . . (p. 40)

The subsequent evolution of the concept of “natural monopoly” is explored extensively by William Sharkey (1982). In any event, it is clear that the concept has since become a commonplace and is nowadays discussed at length in any volume on industrial organization or on the economics of regulation (e.g., see Kahn 1971, ch. 4, esp. pp. 116–123) and Kasserman and Mayo (1995, pp. 413–21). In more recent times Krugman has written extensively on the effects of economies of scale on international trade. His work has attracted considerable modern interest to the subject. We cite here one observation particularly pertinent to this book. Krugman and Venables tell us:

Then a country with a somewhat stronger initial position in some industry than its competitors may find itself with an advantage that cumulates over time.

Producers of final goods will find that the country with the larger industry supports a larger base of intermediate producers, which gives them low enough costs to export to other markets. . . . Thus, each industry will tend to concentrate in one of the countries . . . a dynamic process of regional specialization and differentiation. . . . (pp. 4–5)

Thus the notion that scale economies tend to lead to perfectly specialized equilibria⁷ is widely recognized. Nevertheless, we should remind the reader that our analysis produces the same economic results when nonspecialized equilibria are included in the model.

10.3 Faster Productivity Growth in One Country in Models without Scale Economies

At least since 1953, after the noted Inaugural Lecture by J. R. Hicks, specialists in international trade have recognized that increases in the productivity of a lagging economy can sometimes decrease the overall welfare of a trading-partner country that is more advanced technologically. Nevertheless, debate on the subject has continued among economic historians (see, e.g., McCloskey 1981, pp. 133–83), and a substantial number of economists apparently still believe that such a productivity increase is either certain or very likely to be beneficial to both countries.

Our discussion of productivity follows the lead of the specialists, employing the same premises as they did to show, once again, that enhancement of productivity in one country can sometimes (but will not always) be harmful to another country. It also follows them in basing the analysis on Ricardian models whose trade equilibrium is shifted by improvements in productivity in one country or the other. However, we carry our analysis several important steps further. First, we focus on the effects of the productivity change upon total world output and the national income it provides to the two countries. We show that the previously unemphasized effects on world output, and the way in which they are shared among the trading countries, are at the heart of the issue. Our approach permits us to describe the full set of equilibria that correspond to the different productivity levels and to make a direct comparison of any pair of these equilibria in terms of their benefits to the world and to each of the trading countries. Second, our approach makes it possible to add to J. R. Hicks' generalization about the characteristics that determine whether in any particular case growth of productivity in country 2 is beneficial or detrimental to country 1.

In his Inaugural Lecture (1953) Professor Hicks sketched out an intuitive Ricardian model of the effect of increased productivity in country 2 on its own welfare and on that of its trading partner, country 1. He concluded first that uniform increases in productivity in a trading partner benefit both countries, and then went on to distinguish two other cases. In the first case the improvements in country 2 are concentrated in its export industries, and his conclusion was that this improvement is beneficial to both countries. The intuitive reason is clear. Since country 2, which experiences technical progress in production of good X , is an exporter of X to country 1, the latter has no foreign sales of X to lose as a result, and it can now purchase X more cheaply than before.⁸ In the second case discussed by Hicks, the improvements are concentrated in country 2's import industries, and he concludes that although this is good for country 2, country 1 is worse off. Intuitively this is so because some or all of its exports of X may then be taken over by country 2, and/or the prices country 1 receives for those exports will tend to fall.

Hicks's contribution was quickly followed by a number of expansions, formalizations and commentaries, including E. J. Mishan (1955), H. G. Johnson (1955), W. M. Corden (1956), and Findlay and Grubert (1959).

Hicks's fruitful line of thought was taken up again by Dornbush, Fischer, and Samuelson (1977) in an explicit Ricardian model. This groundbreaking paper is, perhaps, most noteworthy because of its definitive break with the two-commodity models, presenting an ingenious graphic device that permits analysis of the extreme cases in which the number of goods is infinite. On the subject we are now discussing, they concluded, like Hicks, that technological change spread uniformly among the products of the improving country is good for both countries. They also pointed out, however, that the international transfer of technology from a high-wage country to a less advanced, low-wage country can be harmful to the welfare of the transferring country. In an illuminating paper that builds on the ideas of both papers, Krugman (1985) took up the subject of trade between a technologically advanced country and its less advanced trading partner, adding several pertinent and illuminating assumptions. He assumed that the technologically advanced country was likely to make progress more rapidly in its more technologically advanced sectors and traced out the effect of this progress on both countries using a method of analysis similar to that of Dornbush, Fischer, and Samuelson. He found

an interesting asymmetry. Progress in the advanced country is always beneficial to both countries, while progress in the less advanced country, while always beneficial to it, can, depending on circumstances, either be harmful or beneficial to the more advanced country. He pointed out that these results can be interpreted in terms of the tendency of the advanced country to make export-biased improvements and of the less advanced country to make improvements that are more import biased.

More recently Stafford et al. (Johnson and Stafford 1993, 1998; Hymans and Stafford 1995) have analyzed the effect of the improvement in a single industry in one of the countries. They found, consistent with the earlier work, that if the industry starts from a very low level of productivity, and the product is therefore entirely imported, the effects of an initial productivity increase in the low-productivity country, *if it affects trade*, are good for the improving country but harmful to the trading partner. However, at a later stage of productivity growth, when the industry is shifted entirely to the improving country, further improvements are beneficial to both. Whether the overall change yields a net gain or net loss to the other country depends on the balance of the two phases. To summarize, all this work offers the conclusion that productivity improvements in one country are always good (or, rather, never harmful) for it⁹ but that the effect on its trading partner depends on the balance between the damaging effect on importing industries and the beneficial effect on exporting industries.

In this book we introduce different but complementary techniques that make it possible to ask and answer a different but complementary set of questions. Instead of looking at the effect of changes in productivity near an existing equilibrium,¹⁰ we examine all possible productivity parameters and determine which sets of productivity parameters yield the best results for one country or for the other. It emerges, as we have seen, that the attributes that make another country the very best trading partner from the self-interested viewpoint of a developed country require that ideal trading partner to be the exporter of only a few products but to have very high productivity levels in the provision of those goods. This best outcome for the developed country is consequently usually a very poor one for its partly developed partner.

All these results require the presence of a reasonably substantial number (usually six or more) of industries, but then they are valid for all parameter choices, that is, for all models with a large number of

products. Much of the previous literature has tended to deal either with models containing only two products or, at the other extreme (following the pathbreaking work of Dornbush, Fischer, and Samuelson), with an infinite number of products. It is easy to demonstrate that for our multiparameter analysis at least the model with two or three products behaves very differently from models with, say, six or more products. For example, in the case of a small number of products the interests of the trading partners are not characterized by the sources of conflict that always arise when the number of products is larger.

Our approach in this book also brings out a natural connection between this theory and one with economies of scale. The vast number of multiple equilibria that one generally finds in Ricardian models modified to entail economies of scale (see Gomory 1994) may be totally different from the linear-model equilibria that result as several industries are shifted from one country to the other by productivity changes. Nevertheless, all of these scale economies equilibria are included in the set of equilibria that are produced by our linear model if the range of parameter values in the latter is suitably wide. This introduces the tight connection, which we have called the *correspondence principle*, between the family of linear models with changing productivity levels and the results that are obtained from models with economies of scale. The implication is that there are behavioral attributes and policy consequences that carry over from the one case to the other, and that do not seem to have been observed before.

10.4 Governmental Promotion of Entry of an Economy into New Industries: Benefits, Costs, and Market Failure

One of the basic conclusions of our analysis is that entry into a new industry and acquisition of a substantial share of its world market is, in an apparently wide range of circumstances, beneficial to the country that succeeds in such an undertaking. Indeed, we suggest that it may sometimes be appropriate for government to encourage such a development. Obviously this analysis is not unrelated to the old infant industry argument, though we trust that the reader will find it carried out here on a more sophisticated level.

Viner tells us that: "Modern writers usually credit Alexander Hamilton or Friedrich List, or even John Stuart Mill, with the first presentation of the 'infant industry' argument for protection to young industries. It is of much earlier origin, however, and is closely

related both in principle and in its history to the monopoly privileges granted to trading companies opening up new and hazardous trades and to inventions (the ‘patents of monopoly’)” (1937, p. 71). Yet, the passages cited by Viner, while clearly expressing the general idea, do not explore it nearly as thoroughly as Hamilton does, and thereby omit several features of the argument that are relevant to our analysis.¹¹

Hamilton takes the position that protection of infant industries is called for primarily, but not exclusively, by the absence of complete freedom of international trade:

If the system of perfect liberty to industry and commerce were the prevailing system of nations, the arguments which dissuade a country, in the predicament of the United States, from the zealous pursuit of manufactures, would doubtless have great force. It will not be affirmed that they might not be permitted, with few exceptions, to serve as a rule of national conduct. In such a state of things, each country would have the full benefit of its peculiar advantages to compensate for its deficiencies or disadvantages. If one nation were in a condition to supply manufactured articles on better terms than another, that other might find an abundant indemnification in a superior capacity to furnish the produce of the soil. And a free exchange, mutually beneficial, of the commodities which each was able to supply, on the best terms, might be carried on between them, supporting, in full vigor, the industry of each. (*Report on Manufactures*, pp. 200–201)

Having gone on to point out that trade was hardly free (and having offered several supplementary arguments), Hamilton continues:

The spontaneous transition to new pursuits, in a community long habituated to different ones, may be expected to be attended with proportionably greater difficulty. . . . The apprehension of failing in new attempts is, perhaps, a more serious impediment. . . . To this it is of importance that the confidence of cautious, sagacious capitalists, both citizens and foreigners, should be excited. And to inspire this description of persons with confidence, it is essential that they should be made to see in any project which is new—and for that reason alone, if for no other, precarious—the prospect of such a degree of countenance and support from governments as may be capable of overcoming the obstacles inseparable from first experiments.

The superiority antecedently enjoyed by nations who have preoccupied and perfected a branch of industry, constitutes a more formidable obstacle than either of those which have been mentioned, to the introduction of the same branch into a country in which it did not before exist. To maintain, between the recent establishments of one country, and the long-matured establishments of another country, a competition upon equal terms, both as to quality and price, is, in most cases, impracticable. The disparity, in the one, or in the other,

or in both, must necessarily be so considerable, as to forbid a successful rivalry, without the extraordinary aid and protection of government. (pp. 204–205)

According to Pigou (1906, p. 15), Colbert described protective duties as the “crutches to teach the new manufactures to walk” (no source given by Pigou). But Hamilton went well beyond this, citing fear of the unknown, response to unknown risks by the capital market, and the advantages of having gotten there first as primary reasons why market forces cannot ensure the discontinuous leaps necessary to bring an economy from an inferior equilibrium to one that serves its interests better.

There is no need to reproduce in any detail the similar arguments of Friedrich List¹² (1841; see also 1885, p. 300) or those of J. S. Mill (bk. V, ch. X, p. 423). It is only worth noting their common emphasis, in Mill’s words, upon the possibility that: “The superiority of one country over another in a branch of production, often arises only from having begun it sooner. There may be no inherent advantage on one part, or disadvantage on the other, but only a present superiority of acquired skill and experience.”

Moreover all these discussions focus on high start-up costs rather than on static economies of scale, and consider those costs to result from lack of experience in the prospective new industries, fear of uncertainty, capital market limitations, and other sources of high start-up costs as reasons why automatic market forces may perform imperfectly in the presence of multiple equilibria. Of course, the modern term “multiple equilibria” is never used, but the idea of the entry of an economy into an industry new to it is patently the same as the move from one such equilibrium to another.

The modern trade literature, with its repeated discussion of nonuniqueness of trade equilibria in the presence of scale economies, clearly implies that the move from one such equilibrium to another may be able to benefit a country. However, the issue seems only to have been reraised explicitly by Graham (1923). As already noted, Graham argued that if a country has a comparative disadvantage in an industry subject to scale economies, it may nevertheless be advantageous to it to prevent importation of that commodity and produce the item entirely domestically. Wilfred Ethier (1982) provides a careful formal analysis of the issue that seems to be accepted as the resolution of the discussion.

Recently a number of economists have begun to express more policy-oriented reservations about unrestricted free trade. The Spencer-Brander (1983) article has led to an extended and illuminating discussion (see also Tyson 1992).

A more general investigation of the advantages and disadvantages to the countries involved of the move to a new equilibrium through changes in the productivity performance of one of these trading partners has appeared recently in three pathbreaking papers (e.g., Johnson and Stafford 1993). These have already been summarized here and in chapter 4, and we will not repeat the discussion here. We need only conclude that in the area under consideration—the possibility of welfare gains through moves from one equilibrium to another and the obstacles to achievement of such gains by the unaided market mechanism—our discussion also owes much to earlier writers.

10.5 Relation to Other Scale Economies Models

Finally, we examine where our model fits in with the models employed elsewhere in the literature. The previous trade literature has emphasized three different variants of industry scale economies. While each corresponds to real and probably significant phenomena, they require markedly different analytic methods and yield very different conclusions (e.g., see Krugman 1984, pp. 109–110).

One set of widely used models of scale economies assumes them to be internal to the firm. As we have noted, this leads us to expect markets to be monopolistic or subject to monopolistic competition, and unless the markets are perfectly contestable, it is likely to entail nonzero profits.¹³ Helpman, Grossman, and Krugman have been the pioneers in the use of this approach, and have produced entirely new, fundamental and extremely illuminating results with its aid (e.g., see Krugman 1979; Helpman 1984; Helpman and Krugman 1985; Grossman and Helpman 1991).

The second of the previously studied scale economies models entails worldwide externalities in which every producer firm in an industry benefits from the expansion of other firms in the industry, no matter in what countries they may be located. Though this situation was investigated by eminent scholars including Viner (1937), Ethier (1979), and Helpman and Krugman (1985), we will not discuss it further here because it goes in a direction so very different from ours.

The third group of models, with which the analysis in this book can be associated, is perhaps the most traditional in the literature. It assumes that firms are perfectly competitive, that they operate under constant or diminishing returns to scale, but that industry scale economies are produced by externalities that depend on the geographic proximity of the firms in question. Scale economies therefore benefit the firms within an industry only in a given country. Competition then will, of course, drive profits to zero. Examples of the many writings using this approach include Kemp (1969) and Ethier (1982). The concept goes back to Marshall's *Principles*. (For a good review of the history, see Chipman 1965, p. 740 ff). Our model is associated with this third group because we too assume that profits are zero despite the presence of scale economies.¹⁴

Though this is probably the most widely used of the scale economies constructs, it has always aroused controversy. It is sometimes suggested that this case rarely arises except where specialized labor is most effectively trained by experience on the job and the labor force is immobile internationally. There are many more cases, however, in which proximity generates economies external to the firm because the activities of one firm lend support to those of others. The modern semiconductor industry or, indeed, any complex manufacturing industry, is dependent on a host of specialized and experienced suppliers, especially of services, whose absence greatly complicates the startup of an industry and whose presence contributes greatly to efficiency. In such cases high start-up costs can indeed yield a range of ex ante scale economies for the industry, and yet entail constant returns for the firm, effective competition and, hence, a tendency to zero economic profits.

10.6 On Analytical Tools

Obviously some of the methods we have employed here are new. So far as we have been able to determine, the concept of the region of equilibrium points and its depiction in our basic graph have never been used before. Consequently the shape of that region and its economic implications have not been studied earlier. Nor has there previously been a calculation of the number of specialized equilibrium points and the rapidity of the growth of that number as a function of the number of traded commodities.

There have, however, clearly been predecessors in dealing, first, with the number of traded commodities and, second, with the effects on

equilibrium of changes in parameter values, that is, in terms of comparative statics. Here, as will be shown in this section, while the literature plainly provides very powerful tools for comparative statics analysis, most of these tools permit only the study of local variations in parameters, that is, the effects of minuscule changes in their values. An exception to this is the work of Frank Stafford and his collaborators described in chapter 8. Our comparative statics analysis, in contrast, can be considered global, since it can deal with changes in parameter values of any feasible magnitude. This can be important because, as we have seen here, the qualitative character of the effects of small changes can be extremely different from those of more substantial variations.

Turning, first, to the number of traded commodities in a model, it has already been noted that almost all studies of trade in the scale economies case had, until the second half of the 1970s, considered worlds with only two traded goods. These studies were able to show that even with two goods, scale economies introduce multiple equilibria of which more than one can be stable, and some of which may not maximize welfare (e.g., see Matthews 1949–50; Meade 1952; Kemp 1969; Ethier 1979.) Then Dornbush, Fischer, and Samuelson achieved a major breakthrough in providing a procedure that deals with an infinite number of commodities. Since then, Helpman and Krugman (1985) and Grossman and Helpman (1991) have effectively incorporated into their analyses trade in n commodities, with the number n unspecified. Thus our use of n traded goods has a number of eminent predecessors.

However, so far as we know, no previous study has investigated the effect of the number of traded goods on the number and properties of the equilibria. We have demonstrated that the number of specialized equilibria in the two-country case grows as $2^n - 2$. More important, we have shown that things that can be true in the two-good case are generally very different when the number of traded goods is larger than, perhaps, a half-dozen. For example, in most of our discussion we have seen that the peak of country 1's region of equilibria is generally to the right of that of country 2, leading to the zone of conflict that we have repeatedly discussed. It can be shown, however, that in the two-good case these two peaks tend to coincide, so a two-good analysis is likely to overlook the zone of conflict and its significance for theory and policy.

There is no need to repeat here the earlier history of formal comparative statics analysis, going back at least to Cournot (1838), carried out

more fully by Hicks and Allen (1934) and others, and then facilitated by Samuelson (1953) and others through use of revealed preference analysis and duality theory (introduced by Roy 1942 and others). The rather evident point here is that these analyses are all carried out with the aid of the differential calculus, and generally end up finding the partial derivative of some endogenous variable with respect to the value of one of the model's parameters (for an abundance of clear examples in trade theory, e.g., see Dixit and Norman 1980). Clearly, such a derivative can, by its very nature, describe only very local changes and responses. Our model, in contrast, lays out the entire region of equilibrium points and thereby brings out the full set of possible changes and responses, both local and global. This is patently critical in a model of scale economies where the uniqueness ensured by the appropriate convexity-concavity conditions is typically violated. So, as Pigou pointed out (see above), a move in the uphill direction can easily bring the economy further away from the true global maximum.¹⁵ Moreover we have seen that while in some regions of our graph a local move will certainly be advantageous to both trading countries, this is emphatically not generally true of larger displacements. Thus here, as elsewhere in this book, we clearly owe a great debt to our predecessors. But we can also claim that building on their work, we have gone some steps further.

10.7 Concluding Remarks: Application to Discussions of Policy

The implications of scale economies, high costs of entry into an industry and multiplicity of equilibria for the role of government, and unrestricted freedom of trade are all part of a subject very much under current discussion. Once more, Krugman is a main contributor to this analysis (e.g., see Krugman 1983, 1984, 1987; Krugman and Lawrence 1993). Other noteworthy examples appear in Chichilnisky and Heal (1986). There is good reason to expect the debate to continue and expand, as international negotiations, influenced by nationalistic political pressures, keep the subject to the fore.

In closing, we would like to point out that both of our analyses in this book are entirely within the realm of free trade. In our analysis of models with economies of scale, we only discussed free trade outcomes. But we pointed out that there were a vast range of possible free trade outcomes, all of which tend to be sustained by market forces. So that free trade, in the presence of retainable industries, tends to

preserve what is, rather than move the world toward any one predetermined outcome.

In our analysis of the effect of different productivity parameters, we considered the wide range of productivity parameters that are possible for a country, and how the changes in these parameters affected the country's national incomes. However, once again, we were discussing alternatives that might be available to a country, each of which was a free trade outcome. For each choice of parameter values it was the free trade equilibrium that we found and that we discussed. The effects we described, the ups and downs of the different national incomes that resulted from those parameter choices, these were all comparisons between outcomes, each of which was the result of free trade. Free trade, here too, allows a vast range of possible outcomes.

This is a section of [doi:10.7551/mitpress/3306.001.0001](https://doi.org/10.7551/mitpress/3306.001.0001)

Global Trade and Conflicting National Interests

By: Ralph E. Gomory, William J. Baumol

Citation:

Global Trade and Conflicting National Interests

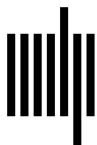
By: Ralph E. Gomory, William J. Baumol

DOI: 10.7551/mitpress/3306.001.0001

ISBN (electronic): 9780262274050

Publisher: The MIT Press

Published: 2022



The MIT Press

©2000 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

This book was set in Palatino by Best-set Typesetter Ltd., Hong Kong

Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Gomory, Ralph E.

Global trade and conflicting national interests / Ralph E. Gomory and William J. Baumol ; with a contribution by Edward N. Wolff.

p. cm.—(Lionel Robbins lectures)

Includes bibliographical references and index.

ISBN 0-262-07209-2 (hc. : alk. paper)

1. Free trade. 2. Protectionism. 3. International trade. I. Baumol, William J. II. Title. III. Series.

HF1713 .G5665 2001

382'.7—dc21

00-056069