

## External Economies and Diseconomies

Howard S. Ellis; William Fellner

*The American Economic Review*, Vol. 33, No. 3 (Sep., 1943), 493-511.

Stable URL:

<http://links.jstor.org/sici?sici=0002-8282%28194309%2933%3A3%3C493%3AEEAD%3E2.0.CO%3B2-8>

*The American Economic Review* is currently published by American Economic Association.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/aea.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [jstor-info@umich.edu](mailto:jstor-info@umich.edu).



# The American Economic Review

VOLUME XXXIII

SEPTEMBER, 1943

NUMBER THREE

## EXTERNAL ECONOMIES AND DISECONOMIES

By HOWARD S. ELLIS and WILLIAM FELLNER\*

### I. *Introduction*

Along with its answer to the principal problem which it set for itself as to how competition allocates resources amongst various uses, neo-classical economics bequeathed to the present generation the much debated proposition that competition causes output under "diminishing returns" to exceed, and under "increasing returns" to fall short of, an output corresponding to the social optimum. Since this issue involves only negligibly any disagreement as to the economic facts, since it is indeed almost a purely quantitative problem not complicated by ethical or other preconceptions, one may find difficulty in understanding how theorists such as Marshall, Pigou, Viner, Graham, Hicks, and Lange could be ranged squarely against Allyn Young, Knight, and Stigler. The answer seems to be that apparently simple technical concepts are often fraught with confusing ambiguities; and the extensive discussion of this subject over a period of years<sup>1</sup> reveals that these ambiguities have already become perennial.

The effort to resolve the issues disputed amongst such distinguished protagonists is in itself an interesting undertaking in the development of theoretical ideas, but there is also no lack of pragmatic justification. The interpretation of certain commonly employed cost functions, of opportunity cost, and of marginal *versus* average cost can not be purely "academic" matters. Furthermore, the Marshall-Pigou proposals of taxes upon "industries of diminishing returns" and bounties upon those of "increasing return," and the Pigou-Graham proposals of a protective tariff in certain instances upon much the same theoretical argument, lead directly to important questions of fiscal and commercial policy. Finally the recent descriptions of price systems under socialism

\* Both authors are members of the department of economics at the University of California, Berkeley.

<sup>1</sup> The footnotes, including n. 22, p. 503, supply a fairly complete bibliography on the subject.

show that the present issues intimately affect not only the formulation of policies in a private enterprise economy but the very concept of the socially optimum employment of resources under (a more or less ideal form of) socialism.

Throughout the period of classical political economy it was a settled but vaguely supported conviction that the price of agricultural produce tends to increase under the influence of diminishing returns, but that "it is the natural effect of improvements to diminish gradually the real price of almost all manufactures."<sup>2</sup> This bad mixture of a dubious "law" of economic history on the one hand, and a truncated part of the static principle of combining proportions on the other may have been engendered originally out of resentment against the Corn Laws and the landed gentry; but it seems to have been perpetuated to the present day through simple confusion. Marshall attempted to give quantitative precision to the notion that "diminishing returns" somehow represent a less favorable application of economic resources than "increasing returns."<sup>3</sup> His demonstration, which ran in terms of the effect of taxes and bounties upon consumers' surpluses in the two cases, is sorely limited by the author's admission that it requires quite special elasticities of demand and supply and that it assumes that marginal utility of money to be constant; and it is entirely destroyed by the failure of the author to include producers' surpluses into the social surplus to be maximized.<sup>4</sup>

The advocates of the tax-bounty thesis have mostly referred to Pigou rather than to the original Marshallian version of the thesis. We turn first to Pigou's analysis so far as it pertains to "diminishing returns" (in his own terminology), which includes also rising transfer costs, leaving until later the case of "increasing returns" (in his terminology), which actually means external economies.

## II. *Diminishing Returns and Rising Transfer Costs ("External Diseconomies"), Without External Economies*

In *Wealth and Welfare*,<sup>5</sup> the first edition of the work later to be more widely known as *The Economics of Welfare*, Pigou draws two positively inclined cost functions (as in Fig. 1), the lower of which,

<sup>2</sup> Adam Smith, *The Wealth of Nations*, Cannan ed., (New York, 1937) Bk. I, chap. 9, p. 242 *et passim*; cf. also J. S. Mill, *Principles of Political Economy*, Gonner ed., (London, 1909), Bk. I, chap. 12, p. 184 *et passim*.

<sup>3</sup> Alfred Marshall, *Principles of Economics*, 3rd ed. (London, 1895), pp. 518-531; and in the corresponding chapter (Bk. V, chap. 12) in later editions.

<sup>4</sup> Cf. J. R. Hicks, "The Rehabilitation of Consumers' Surplus," *Rev. of Econ. Stud.*, Vol. 8 (Feb., 1941), pp. 112-15.

<sup>5</sup> A. C. Pigou, *Wealth and Welfare* (London, 1912), pp. 172-79.

labeled  $S_1$ , he calls "a supply curve of the ordinary type," and the upper, labeled  $S_2$ , "a curve of marginal supply prices." The function  $S_1$  is further described as showing at each point the cost or price at which the corresponding output on the abscissa can be maintained in the long run, and function  $S_2$  as showing at each point "the difference made to *aggregate* expenses"<sup>6</sup> by the production of one more unit. With austere brevity, Pigou concludes directly from the description of the two functions that the intersection of  $S_1$  with the demand schedule at  $C$  corresponds to output and price under competition, whereas the intersection of  $S_2$  with the demand schedule at  $I$  represents the correct output<sup>7</sup> under an ideal allocation of social resources.

In the universe of discourse of Pigou's problem, economic theory now operates with a number of cost functions presently to be de-

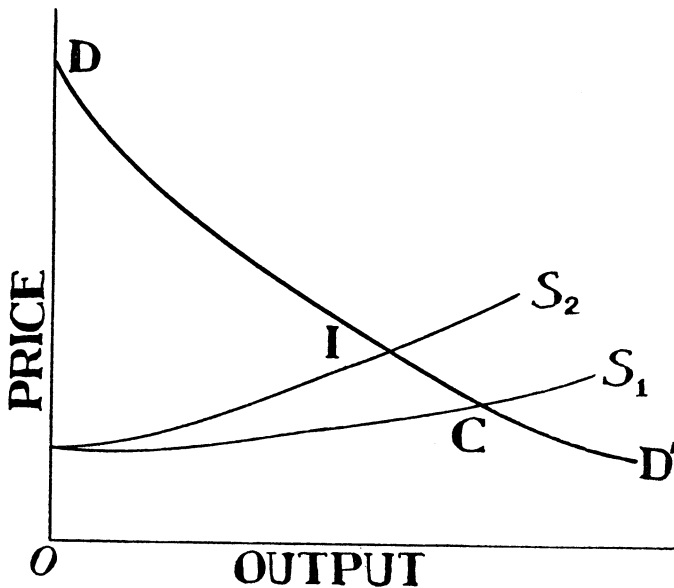


FIG. 1—Pigou's Cost Functions.

scribed; but from the exposition of *Wealth and Welfare* it is impossible to discover which of three mathematical functions Pigou intended to employ. Retrospectively, in answer to criticisms, Pigou acknowledged one possible interpretation but denied that he had intended to draw his conclusions on this basis, proposed another interpretation, and finally, even upon this interpretation, limited his thesis to a special argument concerning international trade.

<sup>6</sup> Our italics.

<sup>7</sup> By implication also the ideal price, but Pigou does not stress this.

In a review of *Wealth and Welfare* appearing about a year after the publication of the book, Allyn Young hailed Pigou's  $S_2$  curve as a "new and powerful instrument of economic analysis" especially as applied to monopoly, but denied that it proved a divergence of competitive from the ideal output.<sup>8</sup> Young's criticism also did not distinguish the three interpretations we shall point out. In fact it did not need to do so, as it is equally valid upon any of them; but Pigou did not believe this to be the case, and so we must examine the merits of each separately.

The three possible interpretations of Pigou's functions arise from the fact that increasing costs in an industry may come from (1) diminishing returns due to the presence of a factor which is fixed in supply for the industry; (2) rising transfer costs due to the presence of a factor which can be drawn in greater amounts from other industries only by a rise in its price; or (3) a combination of (1) and (2). In 1920 Pigou focused attention upon the *first* interpretation by admitting it as a legitimate reading of his cost curves but not what he had intended.<sup>9</sup> This interpretation permits us to give unambiguous definitions to the  $S_1$  and  $S_2$  curves of Figure 1 and to trace out the reasoning on this basis. The "supply curve of the ordinary type,"  $S_1$ , in Pigou's language is the usual "marginal cost curve," that is, the curve indicating the cost of production of the marginal unit of output; and  $S_2$  a function which adds to  $S_1$  at each point the *aggregate* increment of costs on all intramarginal units of output.<sup>10</sup> Thus, if an expansion of output from 50,000 to 51,000 units involves a rise of cost at the margin from \$1.00 to \$1.01, the ordinate of  $S_1$  at 51,000 units is \$1.01, and the ordinate of  $S_2$  is the difference between 50,000 times \$1.00 and 51,000 times \$1.01, or \$1,510, divided by 1000, or \$1.51.

On the present interpretation of Pigou (which he acknowledged as a possibility) increasing costs arise solely from rising transfer costs; and to make this interpretation explicit, in Figure 2 we give to  $S_1$  and  $S_2$ , respectively, the distinctive labels of  $\epsilon$  and  $\gamma$ . Pigou's  $S_1$  (our  $\epsilon$ ) shows the cost of the marginal unit in isolation, or more explicitly: marginal cost excluding all increments of transfer cost; alternatively  $\epsilon$  shows also *average cost* per unit of output, including transfer rent. Pigou's  $S_2$

<sup>8</sup> Allyn A. Young, "Pigou's Wealth and Welfare," *Quart. Jour. of Econ.*, Vol. 27 (Aug., 1913), pp. 672-86. Edgeworth also called the function "the marginal increment of cost from the viewpoint of a monopolist"; cf. F. Y. Edgeworth, "Contributions to the Theory of Railroad Rates, IV: A Digression on Professor Pigou's Thesis," *Econ. Jour.*, Vol. 23 (June, 1913), p. 211.

<sup>9</sup> A. C. Pigou, *The Economics of Welfare*, 1st ed. (London, 1920), pp. 934-36.

<sup>10</sup> If the  $n$ -th unit of the output is produced at a cost of  $f(n)$ , then  $S_1 = f(n)$ , and

$$S_2 = \frac{d[nf(n)]}{dn} = f(n) + nf'(n)$$

(our  $\gamma$ ) shows "marginal costs" also but in a different sense—marginal cost including the *total* increment to transfer costs *on all units*.

Pigou originally maintained that the intersection at *C* represents competitive equilibrium, but that the ideal allocation of resources would be given by reducing output through appropriate taxation to a

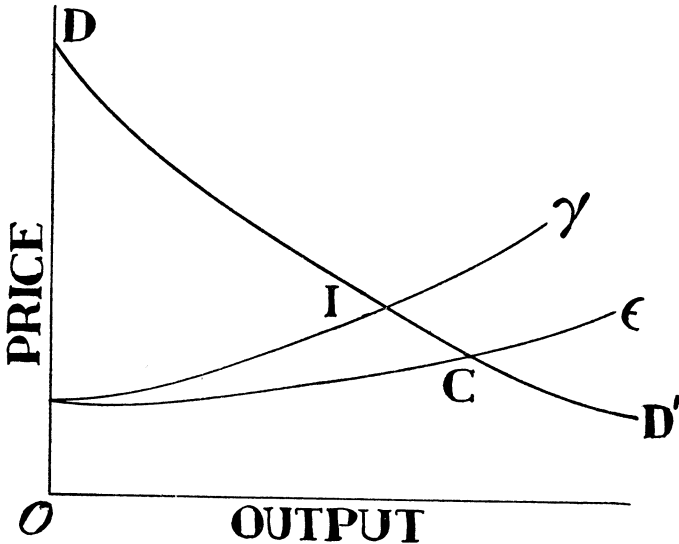


FIG. 2—Rising Transfer Costs.

magnitude given by the intersection at *I*. Allyn Young accepted this description of competitive equilibrium (as do all parties to the dispute) but denied that the *total* increment of cost to the industry could be regarded as a cost to society. In Pigou's excellent paraphrase, "In other words, according to Professor Young's view, the excess of marginal supply prices [our  $\gamma$ ] in industries of diminishing returns [read: 'increasing costs'] over the corresponding supply prices [our  $\epsilon$ ] is merely a nominal excess of money paid, and not a real excess representing resources employed."<sup>11</sup>

If Young had spoken the magical word "rent," it seems probable that Pigou would have capitulated completely. For Young could have put his proposition in these words: "If the expansion of an industry gives a factor a higher per unit remuneration, whether or not that higher price induces a greater aggregate [social] supply of the factor, the units already being supplied earn producers' rents [or increase the previous rent]; and rent is not a cost in social resources." Consequently if the output of a commodity expands, the rise in transfer

<sup>11</sup> Pigou, *op. cit.*, p. 935.

costs (*i.e.*, in the value) of the intramarginal units of the transferred resource is not part of the marginal social cost of producing the commodity under consideration. The marginal social (opportunity) cost of transferring resources yielding  $n$  units is merely the cost of transferring the resources required for the production of the  $n$ -th unit. This cost is expressed by  $\epsilon$  not by  $\gamma$ . The  $\gamma$  function is not a social cost curve because it includes increments to rent.

Pigou accepted this criticism without reservation so far as concerns transfer costs; and, we may add, he accepted it gracefully and without seeing, at that time, its full import. For he proceeded to argue in the sentences immediately following that his conclusions are valid for diminishing returns, although not for rising transfer costs:

The reason why diminishing returns in terms of money [read: "increasing costs"] appear when they do appear is, in general, not that the money price of factors employed is increased, but that that proportionate combination of different factors, which it is most economical to employ when  $(x + \Delta x)$  units of commodities are being produced is a less efficient proportionate combination than that which it is most economical to employ when  $x$  units are being produced; and the extra cost involved in this fact is real, not merely nominal. For these reasons Professor Young's objection, as a general objection, fails.<sup>12</sup>

Employing a useful notation introduced by Mrs. Joan Robinson,<sup>13</sup> we show in Figure 3 the functions germane to diminishing returns, transfer costs assumed constant.

Mathematically the definitions of  $\alpha$  and  $\beta$ , respectively, are precisely the same as for  $\gamma$  and  $\epsilon$ , since both  $\alpha$  and  $\gamma$  satisfy the requirements of Pigou's  $S_2$ , and  $\beta$  and  $\epsilon$  the requirements of  $S_1$ . Only the economic implication is changed:  $\alpha$  refers to marginal cost including the total increment of Ricardian rent, and  $\beta$  to marginal cost excluding Ricardian rent or average cost including average Ricardian rent. The function  $\delta$ , to which there is no counterpart in the case of rising transfer costs, is the familiar curve of average cost excluding rent. According to the first edition of *The Economics of Welfare*, the intersection at  $C$  shows the competitive solution, as before; and the intersection at  $I$  again is held to represent the socially ideal output, though this time on the grounds that less efficient combinations of factors signify a social "extra cost."<sup>14</sup>

<sup>12</sup> Pigou, *op. cit.*, p. 936.

<sup>13</sup> Joan Robinson, *The Economics of Imperfect Competition* (London, 1933), chap. 10, "A Digression on the Four Cost Curves." While her chapter fails to distinguish  $\epsilon$  and restricts the concept of rent to "Ricardian" surplus on the fixed factor, it is in general accurate and illuminating, and has contributed indirectly very much toward the present analysis.

<sup>14</sup> Pigou's contention (*cf.* 1st ed., p. 194) that of two roads connecting the same two

However real the "extra cost" from these grounds, the application of Young's reasoning proves it to be adequately included in  $\beta$ , the cost at the margin, *i.e.*, the incremental cost in the variable factor. Unless rent is a social cost, it is erroneous to envisage the social marginal cost as including the increment to rent as in the function  $\alpha$ .

To maximize the aggregate of producers' and consumers' surpluses, the relevant magnitude is  $DCA$ , not  $DIA$ , if the marginal utility of

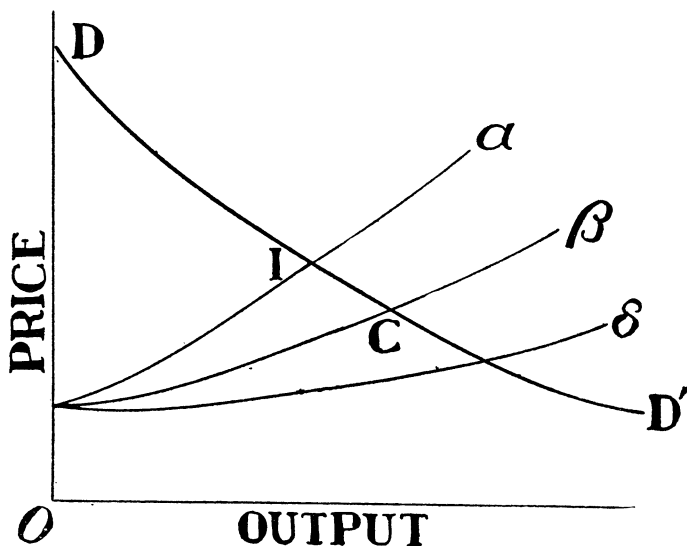


FIG. 3—Diminishing Returns.

money is assumed to be constant. Dropping the assumption of constant marginal utility for money renders it impossible to express the aggregate surplus of consumers and producers by areas lying between demand and supply curves. It still remains true, however, that the optimum output will be reached at the intersection of the demand curve with  $\beta$ , not  $\alpha$ . The price will still express for each consumer the value

points the one, assumed to be superior but narrow and therefore subject to diminishing returns, is overexploited in competition unless taxed differentially, seems to have rested on the notion that competitive output is determined by the  $\delta$  function. The contention was proven to be fallacious by Professor Knight, who has shown that the owner of the good road will charge a toll that will raise costs to users to the  $\beta$  level. *Cf.* F. H. Knight, "Fallacies in the Interpretation of Social Cost," *Quart. Jour. of Econ.*, Vol. 38 (Aug., 1924), pp. 582-606, reprinted as chap. VIII in *idem*, *The Ethics of Competition* (New York, 1935), pp. 217-36. Pigou omitted the "two roads" example from the second edition of his book which happened to appear almost simultaneously with Knight's criticism, and does not refer to the latter. No special significance should be attributed to Pigou's recantation of the "two roads" proposition because it was inconsistent with his own position and should have been dropped even if his position had not been modified.



of the change in total utility occasioned by the marginal unit of the commodity, although it will express this change in terms of the variable marginal utility of money rather in a constant unit of measurement.

If marginal social cost equals price, it also equals the value of the marginal addition to the utility of each consumer as expressed in terms of the marginal utility of the money stock actually owned by him. This, however, implies that, given the distribution of wealth and income, the  $\beta$  intersection is optimal. If resources were to be shifted out of the industry in question, each consumer would lose more utility, as expressed in terms of money, than the saving in social cost; and if further resources were to be shifted into the industry, the addition to social cost would outweigh the gain in utility.

That his reply to Young was inadequate, Pigou admitted four years later in the *Economic Journal*,<sup>15</sup> adumbrating a revised statement in the second edition of *The Economics of Welfare*. The revision consisted in the abandonment of the general thesis that, under increasing cost, output under competition exceeds the ideal, and the adoption of the very limited proposition that a divergence occurs only from the viewpoint of one nation against another when it pays agricultural rents to foreign owners in the price of imports.<sup>16</sup>

This limited version of the proposition, retained in the later editions of *The Economics of Welfare*, requires the use of functions which express the effects of rising transfer costs as well as those of Ricardian diminishing returns. In Figure 4, average costs as affected by diminishing returns but not by rising transfer costs<sup>17</sup> are shown by  $\beta$ ; average costs as affected by rising transfer costs but not by diminishing returns<sup>18</sup> are shown by  $\epsilon$ ; and average costs as affected by both forces in conjunction are shown by  $\phi$ . The function  $\phi$  may be described further as marginal cost excluding all Ricardian rent and all increments to transfer rent, or as average cost including average Ricardian and average transfer rent. To this magnitude, at each point,  $\theta$  adds the *total* increment of transfer rent on the intramarginal transferred units, and the *total* increment of Ricardian rent.<sup>19</sup> Pigou's

<sup>15</sup> A. C. Pigou, "Comment," *Econ. Jour.*, Vol. 34 (Mar., 1924), p. 31.

<sup>16</sup> A. C. Pigou, *The Economics of Welfare*, 2nd ed. (London, 1924), pp. 194-95. The third edition (London, 1929), p. 225, made no change. F. Y. Edgeworth, "The Revised Doctrine of Marginal Social Product," *Econ. Jour.*, Vol. 35 (Mar., 1925), p. 35, agreed with this version of Pigou.

<sup>17</sup> *I.e.*, calculated at constant transfer costs.

<sup>18</sup> *I.e.*, calculated as though constant returns prevailed.

<sup>19</sup> Mrs. Robinson makes  $\alpha$  do double duty as a designation for marginal cost including only the total increment of Ricardian rent and as a designation for our  $\theta$ . This is an undesirable ambiguity. Furthermore, since she omitted to distinguish  $\epsilon$ , she has no

argument with respect to foreign trade would then mean that rents transferred abroad, whether those rents originated in diminishing returns or rising transfer costs, are costs to the domestic economy; the competitive purchases of such imports would run to the foreign output as determined by  $\phi$ , whereas ideally the importing country should buy only the foreign output as determined by  $\theta$ .

But even the foreign trade argument, which had been adumbrated by Graham a year<sup>20</sup> previous to Pigou's decision to restrict the argument to rents disbursed to foreign owners, is subject to rather severe

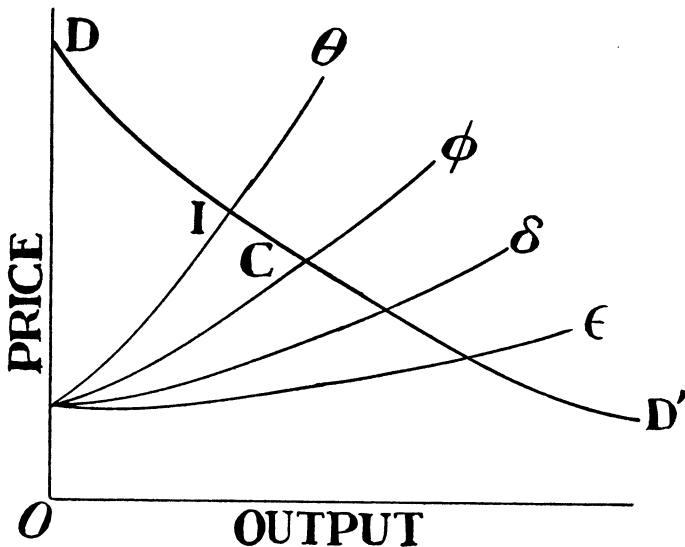


FIG. 4—Diminishing Returns and Rising Transfer Costs.

limitation if not complete rejection. After all, the doctrine of optimum allocation in all of the versions here considered is a "cosmopolitan" doctrine in that it does not distinguish between surpluses accruing to domestic owners *and* consumers on the one hand and surpluses accruing to foreigners on the other. The qualifications required to allow for policies distinguishing between the interests of domestic economic subjects and foreigners are much more extensive than would be suggested

function corresponding to our  $\phi$ , which adds  $\epsilon$  and  $\beta$ .

The functions  $\alpha$ , lying between  $\theta$  and  $\beta$ , and  $\gamma$ , lying between  $\theta$  and  $\epsilon$  have been omitted in Fig. 4;  $\delta$  has likewise not been drawn.

<sup>20</sup> F. D. Graham, "Some Aspects of Protection Further Considered," *Quart. Jour. of Econ.*, Vol. 37 (Feb., 1923), pp. 199-227; "The Theory of International Values Re-examined," *ibid.*, Vol. 38 (Nov., 1923), pp. 54-86; "Some Fallacies in the Interpretation of Social Costs; a Reply" (to F. H. Knight), *ibid.*, Vol. 39 (Feb., 1925), pp. 324-30.

by the limitation of Pigou's argument to such increases in rents as are included in the value of commodity imports. One would have to exclude from the aggregate to be maximized all consumers' surpluses accruing to foreigners; and also all producers' surpluses accruing to foreign stockholders or other foreign owners of domestic enterprise.

The Pigou doctrine in the latest and narrowest version would lead one to believe that in a competitive world the "national"—as opposed to the "cosmopolitan"—aggregate of surpluses is increased by a tax reducing the output of industries which, by importing raw materials, increase the value of foreign resources. The cosmopolitan net surplus is surely reduced by such a tax; whether the "national" net surplus is increased or reduced depends on how the aggregate of consumers' and producers' rents accruing to domestic economic subjects changes when other industrial activities are partly substituted for those connected with raw material imports.

To replace the "cosmopolitan" approach with the "national" would require a reinterpretation of the entire doctrine no less fundamental than that which would be required if the distribution of wealth within the economy were not to be accepted as "given." In both cases the "votes" of the consumers and of the producers whose behavior determines the allocation of resources is held to be weighted incorrectly. The important qualifications arising from this consideration must either be disregarded, which means accepting the cosmopolitan point of view and taking the distribution of wealth and income as given; or they must be dealt with in the framework of a broad sociological approach extending to questions such as reprisals in international relations, the potential stability of different patterns of distribution within social communities, etc. In no event does it seem satisfactory to confine the reinterpretation of the "cosmopolitan" doctrine to industries which, by importing raw materials, give rise to foreign producers' rents.

The preceding analysis is not concerned with the genuine diseconomies arising from phenomena such as the smoke nuisance, the wasteful exploitation of natural resources, etc.

So far as concerns the present heading—diminishing returns and rising transfer costs—we have found: (1) There is no divergence between the ideal and competitive outputs. (2) Pigou originally believed that output under competition is excessive because the total increment to rent is not included as cost. (3) Under the force of Allyn Young's criticism that rents are not social costs, Pigou gradually attenuated his thesis to the case of imports produced under rising supply price. (4) Writers who invoked the authority of Pigou after 1924 on what

had once been his *general* thesis did so unjustifiably.<sup>21</sup> (5) Even the restricted foreign trade thesis has little or no validity as a single qualification because the entire problem would have to be reformulated if this qualification, in conjunction with more important ones, were to become valid. (6) Since rents are not social costs, the relevant cost function—one which maximizes the total of consumers' and producers' surpluses—is marginal cost in the sense of costs of the marginal unit of output alone (= average cost including average rent), that is,  $\epsilon$  or  $\beta$ , and *not* marginal cost including the total increment of rent, that is,  $\gamma$  or  $\alpha$ . (7) The atomistic single seller notices and acts correctly upon the costs of  $\epsilon$  or  $\beta$ ; he does not notice and should ignore the additional costs incorporated in  $\gamma$  or  $\alpha$ .

### III. *External Economies*

Economists upholding the special tax in the diminishing returns case also maintain the necessity of a bounty for the realization of external economies. One of the most debated issues in the entire discussion of the tax-bounty thesis has always been the reality of external economies.<sup>22</sup> Many supposed examples have proved to be spurious or far-fetched; but we do not propose to begin upon the painstaking inquiries into techniques and economic history which would be necessary to appraise the possibilities. Among the many difficulties and complexities, however, there are a few certainties.

One is that if an "external economy" is an internal economy to another *industry*, the outcome is either monopoly in the second industry, or else the complete exploitation of the internal and hence the disappearance of the external economies.<sup>23</sup> If the outcome is monopoly in the second industry, costs are very unlikely to decline in the first, since a monopolist will respond to a rise in demand with a reduction of price only (1) in case he is operating in the downward range of his *marginal* cost curve (and the elasticity of the new demand curve is not

<sup>21</sup> Oscar Lange (and F. M. Taylor), *On the Economic Theory of Socialism* (Minneapolis, 1938), pp. 98-99; Jacob Viner, "Cost Curves and Supply Curves," *Zeitschrift für Nationalökonomie*, Vol. 3 (Sept., 1931), p. 42.

<sup>22</sup> In addition to the literature elsewhere cited in these pages, see the following: J. H. Clapham, "Of Empty Economic Boxes," *Econ. Jour.*, Vol. 32 (Sept., 1922), pp. 305-14; A. C. Pigou, "Empty Economic Boxes: A Reply," *ibid.*, Vol. 34 (Dec., 1922), pp. 458-65; *idem* and D. H. Robertson, "Those Empty Boxes," *ibid.*, Vol. 34 (Mar., 1924) pp. 16-31; G. J. Stigler, *The Theory of Competitive Price* (New York, 1942) pp. 106, 142-44; K. E. Boulding, *Economic Analysis* (New York, 1941), p. 194; Joan Robinson, *op. cit.*, Appendix, pp. 337-43.

<sup>23</sup> Thus Knight, *op. cit.*, p. 229, and "On Decreasing Cost and Comparative Costs," *Quart. Jour. of Econ.*, Vol. 39 (Feb., 1925), pp. 331-33; and Joan Robinson, *op. cit.*, p. 340.

sufficiently smaller than was that of the old one to offset the downward slope of the marginal cost curve); or (2) if the new demand curve is more elastic than was the old one (and the upward slope of the marginal cost curve in the relevant range is insufficient to offset this circumstance).

Usually the monopolist will raise his price if demand increases, in which case such economies as are internal economies in the "second"

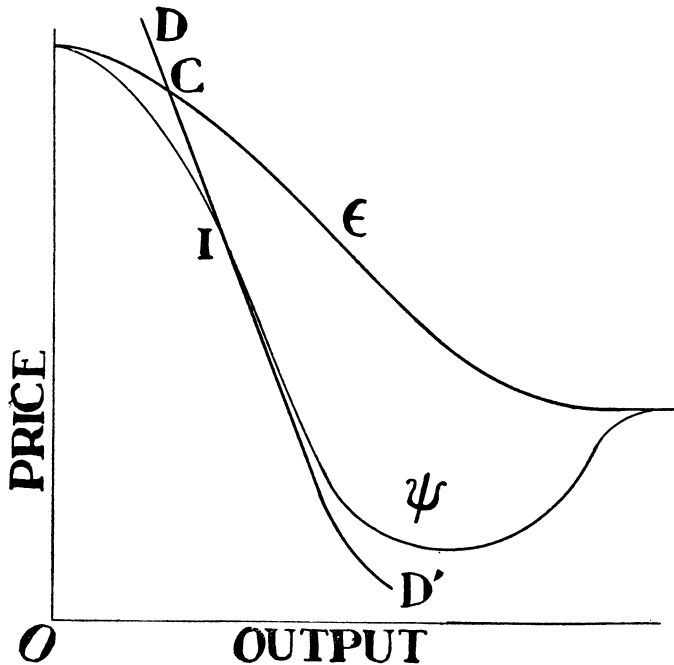


FIG. 5.—External Economies in the Absence of Diminishing Returns and Rising Transfer Costs.

industry will not lead to the realization of economies in the "first" industry. As Mr. Sraffa has suggested, the concept of external economies may, however, be rescued by illustrations not depending upon lowering the price of a commodity supplied by another industry, but upon the better rendering of services.<sup>24</sup> The development of a skilled labor force, the migration of suitable labor, the appearance and progress of professional and trade associations and journals, and the like, are not to be dismissed as unrelated to output in every case.

When and if external economies exist, they must be incorporated

<sup>24</sup> Piero Sraffa, "Laws of Returns under Competitive Conditions," *Econ. Jour.*, Vol. 37 (Dec., 1926), pp. 535-50.

into the structure of economic theory; but it must be confessed that the theoretical treatment of this subject still leaves much to be desired. The first step in reconstruction is a clear description of the cost functions and of competitive equilibrium.

In Figure 5 we assume that there is no "scarce" factor for this particular industry, and that its demand for factors is atomistic so that an extension of output does not cause rising transfer costs. The function  $\epsilon$  signifies, as in earlier contexts, average cost, or marginal cost in the sense of costs for the marginal output in contrast to smaller outputs.<sup>25</sup> The function  $\psi$  is defined somewhat analogously to  $\gamma$  and  $\alpha$  as the total cost difference on *all* units (marginal and intramarginal). If there are no external economies (or none unexploited),  $\epsilon$  and  $\psi$  coincide in a horizontal line; if there are,  $\psi$  lies below  $\epsilon$ . Thus without economies the per unit cost might be \$4.00 for all outputs. With economies the first "unit" costs \$4.00 if only one unit is produced; if "two units" are produced the second unit costs \$3.00 and the costs of the first unit also decline to \$3.00. At an output of two units  $\epsilon = \$3.00$ ,  $\psi = \$2.00$ .

We employ the symbol  $\epsilon$  because external economies have exactly the opposite significance *to the industry* as rising transfer costs. One must be on guard against an extension of the meaning of  $\epsilon$  from the behavior of *output* costs to the price of the variable factor; if  $\epsilon$  declines because of external economies, factor price almost necessarily rises. This probable rise, however, is disregarded in the graphic representation and in the foregoing example, where the industry in question is assumed to be atomistic and consequently constant factor prices are assumed.

The equilibrium to be expected from competition is determined by the intersection of  $DD'$  and  $\epsilon$ . The costs of production at the margin will be equated to price. The circumstance that further expansion of the industry would reduce intramarginal costs will not produce such an expansion, since to the right of  $C$  costs at the margin fall short of price. Now it has been argued that optimum allocation in these circumstances requires subsidizing the industry under consideration in such a manner as to make its output expand to  $I$ , corresponding to the intersection of  $DD'$  with  $\psi$  (instead of  $\epsilon$ ).

This argument was presented by Pigou at the time when he first attempted to establish the proposition that increasing cost industries tend to overexpand in free competition. But while the proposition relating to increasing cost industries was gradually limited to the foreign

<sup>25</sup> On the assumption of free competition, costs to all firms are the same for their equilibrium outputs.

trade proposition previously discussed, the thesis that decreasing cost industries do not expand sufficiently was and is maintained by him and other authors.<sup>26</sup> It is worth pointing out that shifting resources into decreasing cost industries is not the same thing as shifting resources out of competitive increasing cost industries, since competitive industries may also operate under constant costs, not to speak of industries operating under monopoly and monopolistic competition. Besides, the proposition that the social optimum output of increasing cost industries is determined by  $S_2$  functions (rather than  $S_1$  functions)<sup>27</sup> would be basically different from the proposition that the social optimum output of decreasing cost industries is determined by  $S_2$  functions (rather than  $S_1$  functions), even if subsidizing decreasing-cost industries could be carried out only by taxing increasing-cost industries. We shall now be concerned with the second of these two propositions.

The reasoning by which it can be shown that the argument relating to increasing cost industries is faulty does not affect the validity of the argument relating to decreasing cost industries. The crucial point here is that, while in conditions of increasing supply price, the *rise* in intramarginal costs is rent rather than social cost, the *decline* in intramarginal costs attending the expansion of decreasing cost industries is a social economy, *i.e.*, social cost with a negative sign. One might therefore conclude that in conditions of decreasing supply price the  $\psi$  function expresses marginal social cost; and that the social optimum output is  $I$  instead of  $C$ .<sup>28</sup> The competitive output, one might conclude, is determined by the socially "incorrect" cost function  $\varepsilon$ , which fails to express the marginal saving in intramarginal social cost, that is, the saving in social cost on intramarginal units of the resource attending the increase in total output by a marginal unit. The  $\psi$  function contains the necessary correction; and a permanent bounty inducing the production of  $I$  is required to achieve optimum allocation. It would have to be added that a two-dimensional presentation, like that in Figure 5, overstates the deviation from the optimum in case the industry is not atomistic. In this event the  $\psi$  curve shifts upward as output expands, since opportunity costs rise as resources are shifted away from other employments.

This conclusion is correct if the external economies are "reversible": the  $\psi$  function actually expresses marginal social cost in case the

<sup>26</sup> A recent exposition of the thesis is found in Hicks, *op. cit.*

<sup>27</sup> Cf. Figure 1.

<sup>28</sup> We mean, of course, the corresponding distances along the abscissa. In the event of more than one intersection the optimum is reached at the intersection maximizing the expression  $\int [D(x) - \Psi(x)] dx$ , where  $D(x)$  is the demand curve.

economies appear with an expansion of output, but *disappear* if and when output subsequently contracts.<sup>29</sup> In the event of irreversibility the problem acquires different characteristics, however. It may be suggested that irreversible external economies are much more significant than are reversible ones. Certain industries must usually reach some stage of growth before a geographical region starts to develop significantly and also before human and material resources become more specialized. But it is rarely true in these cases that a contraction of the output of any one industry would lead to a loss of the economies in question.

If irreversible external economies are potentially present, competi-

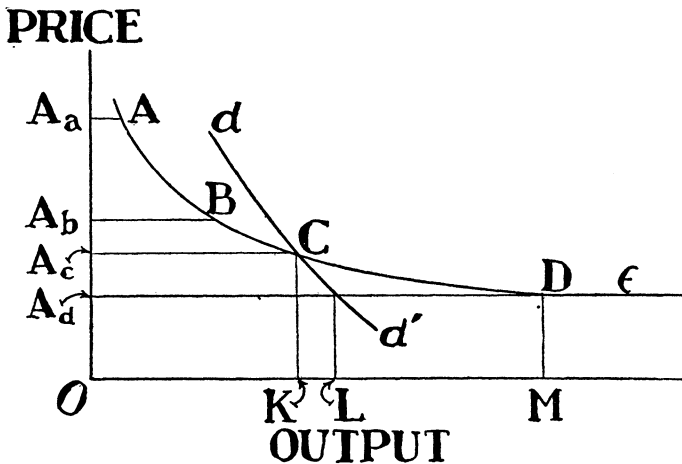


FIG. 6—Irreversible External Economies in the Absence of Diminishing Returns and Rising Transfer Costs.

tive equilibrium fails to achieve optimum allocation; at the same time  $I$  ceases to be the optimum output and the bounty required to achieve optimum allocation is temporary.

In the circumstances now considered the  $\psi$  function loses its significance. What happens in conditions like these can best be expressed by the statement that the  $\epsilon$  function shifts permanently as we move downward along the curve.

Assuming that transfer costs remain unaffected by the output of the industry in question and that the industry does not give rise to Ricardian rents, there will be a point ( $D$ ) on the  $\epsilon$  curve to which an output ( $OM$ ) corresponds that exhausts all potential external econo-

<sup>29</sup> As was pointed out in the preceding paragraph, the function ceases to be a curve if the industry is not atomistic. In this event  $\Psi$  is a function of more than one variable. The "curve" shifts upward when output expands, and downward when it contracts.



mies. Once that output is reached, the  $\varepsilon$  curve will intersect with the ordinates at  $Ad$  and it will be horizontal throughout its course. Before any output—whatsoever is produced, the path of the function is marked by the points  $A, B, C, D$ ; the curve is horizontal only to the right of  $D$ . If a point lying along the curve (such as  $A$  or  $B$ ) has already been realized, a movement backward, toward the ordinate, occurs along a horizontal line (such as  $AA_a$  or  $BA_b$ ); whereas a movement forward, toward  $D$  occurs along the still unused portion of the original function (such as  $ACBD$  or  $BCD$ ).

To say that external economies are mostly of this kind, instead of being “reversible,” may be interpreted to mean that they are typically *dynamic* phenomena. Whether shifts like these should be excluded from equilibrium analysis as Knight suggests and then relegated to “dynamic theory” is clearly a matter of convenience. The opinion may be expressed, however, that, so long as certain shifts of functions are on the same level of predictability as are movements along the functions, it is not very fruitful to distinguish between two types of theory dealing with these two kinds of changes respectively. The shifts reflecting irreversible external economies surely are not always on the same level of predictability as movements along given supply curves. But the difference may in some cases be insufficient to warrant the exclusion of the phenomenon from the type of theory we are concerned with at present.

Returning now to Figure 6, we observe that the conditions sketched there would, in the first approximation, justify a temporary bounty raising the output of the industry from  $OK$  to  $OM$  and thereby producing the lowest obtainable social cost curve for the commodity. After the discontinuation of the bounty the output of the industry drops to  $OL$  and social net output is higher than it was prior to the interference and higher than it would be if the bounty were continued. We said this conclusion was justified only “in the first approximation” because it disregards the waste involved in moving temporarily an excessive stock of resources into the industry under consideration. The perpetual income stream yielded by the capital value of the temporary excess of resources is a loss to be counted against the decline in social cost if the excess of resources (corresponding to the temporary excess output  $OM$  minus  $OL$ ) is completely immobile. The more mobile the excess is, the smaller does the loss become.

Moreover, by directly subsidizing certain processes, such as migration or the specialization of resources, it may be possible to induce, at the “permanent” output level  $OL$ , the economies that would develop automatically at the output level  $OM$ , and thus to avoid the detour over the excessive output  $OM$ . Such temporary subsidies, raising

the output of the industry from *OK* directly to *OL*, might in certain circumstances achieve the maximum obtainable saving in social cost, with no sinking of an excessive stock of specialized resources into the subsidized industry. The detour over *OM* does involve costs of this character. If the costs of a detour are unavoidable the true optimum output which takes account of these costs may be different from *OL* since the temporary excess of output is not the same for the different potential final outputs.<sup>30</sup>

In the preceding pages we considered external economies in isolation, disregarding diminishing returns and rising transfer costs. If, as seems plausible to assume, external economies occur in conjunction with the two cost-raising forces functionally related to industry output, the fundamentals of the preceding analysis are not altered. For a comparison of competitive output with social optimum output it is not necessary to divorce the two *cost-raising* forces from one another, since neither of them upsets the identity of the two outputs in question. But it is necessary to divorce the *cost-diminishing* forces from those *raising* cost, because the divergence of social optimum output from competitive output is determined by the *cost-diminishing* forces, regardless of whether the interaction of all forces makes for increasing, constant or declining supply price.

It should be repeated in this connection that statements like the foregoing one imply that it is fruitful to include in equilibrium theory the analysis of certain phenomena that in a sense are "dynamic," *i.e.*, that reflect themselves in shifts of the curves used in traditional equilibrium theory. Reversible external economies, not implying shifts of the curves, are in all probability unimportant. Whether it is fruitful to take account of "dynamic"—that is, irreversible—external economies in this type of approach depends on the degree of predictability, or, as it may be expressed alternatively, on the "regularity" of the phenomenon called external economies. Inclusion into the body of equilibrium theory may be warranted in some cases, but certainly not in all cases in which costs are declining "historically."

#### IV. Summary

We have found that diminishing returns and external diseconomies (rising transfer costs) do not result in a divergence between social and competitive costs. Social cost equals resource (opportunity) cost, *i.e.*, the cost of production of additional or marginal units. If this cost

<sup>30</sup> If the distance between the demand curve and the original curve increases monotonously to the right of *OK*, then the temporary excess will be smaller for final outputs smaller than *OL*.

advances because of diminishing returns or rising transfer costs, the higher cost to society is completely exhausted by a price covering the increment to costs on marginal units. The increment to rents on intramarginal units are not social costs; they do not enter into competitive supply prices and they should not. The statement made by certain writers<sup>31</sup> that an ideal allocation of resources requires that production be guided by "marginal cost," and not by "average cost" as it is under competition is either meaningless or erroneous. It is meaningless because cost *including* the total increment to rent on intramarginal units ( $\gamma$ ,  $\alpha$ , and  $\theta$ ) is "marginal cost" in one sense; and cost on the marginal units excluding these rents ( $\epsilon$ ,  $\beta$ , and  $\phi$ ) are also "marginal costs" in another sense; and because the second set of functions, forming the supply curves under competition, are both "marginal" and "average," according as the cost is computed incrementally without rent, or averaged including *average* rent. The statement is wrong if it means that functions such as  $\gamma$ ,  $\alpha$ , and  $\theta$  represent social costs.

When unexploited external economies exist, competitive output falls short of the optimum. But again this can not be expressed for all cases as a divergence of "average" and "marginal" cost, the former being represented as the correct guide. In the case of "reversible" external economies—those which disappear when the subsidy is terminated—a marginal function such as  $\psi$ , which subtracts from the increment to costs attributable to the marginal unit in isolation the decrement to costs on intramarginal units occasioned by the external economies, should be regarded as the true social costs function.<sup>32</sup> But in the case of "irreversible" economies, the contrast of average and marginal costs can not be used to indicate the divergence of competitive from social costs, inasmuch as the economies simply cause a downward shift of average costs as a horizontal function up to each realized output; "marginal," as anything distinct from this successive lowering of average costs, has no meaning.

Retrospectively, it is also worth remarking that the "atomistic" character of one producer's output under competition, frequently thought to be crucial in the external economies—diseconomies context, is not decisive of itself. In the "diseconomies" case, as we have seen, the private calculus of the single competitive producer results in no divergence of competitive and ideal output, for rents are not social costs. Where there are *genuine* diseconomies ignored by the competitive producer—smoke nuisance, wasteful exploitation of re-

<sup>31</sup> Lange, *op. cit.*, pp. 98-99, n. 2; A. P. Lerner, "Statics and Dynamics in Socialist Economies," *Econ. Jour.*, Vol. 47 (June, 1937), pp. 253-70.

<sup>32</sup>  $\Psi$ , as was shown, is a function of more than one variable, and hence not a curve, if the industry in question is not atomistic.

sources, etc.—these results follow not from the atomistic character of production, but from technical or institutional circumstances as a consequence of which scarce goods are treated as though they were free; and the divorce of scarcity from effective ownership may be equally complete for atomistic, oligopolistic, and monopolistic private enterprise.

The divergence of competitive from ideal output under external economies is more closely related to the atomistic position of the single competitor because, to the monopolist,  $\psi$  functions are the marginal cost functions if the economies are reversible; and he may take account of the downward shift of the  $\varepsilon$  functions if the economies are irreversible but predictable. It must not be overlooked, however, that, on the demand side, precisely the same circumstance—the non-atomistic position of the single producer—leads to the determination of output by marginal revenue which falls short of demand price. This in and of itself always tends to reduce output below the competitive level.

We have found theoretical possibilities of achieving by state interference more external economies than are given by competition; but it is, to say the least, doubtful whether frequent or extensive interference would be justified in practice upon this basis. Where economies are such as to be permanent (irreversible) once the requisite output is developed, we have an almost perfect analogy with the infant industry argument for tariffs, and little more need be said to emphasize the pitfalls.

The departure of the economist's *free* competition from the ideal of social costs is in fact negligible for external economies and non-existent for the cost-increasing forces. The departures of *actual* competition and the manifold other market forces from free competition are striking, just as the departures of actual forms of the corporate state, socialism, and planning from more or less ideal prototypes might also be striking.