The Xistence of X-Efficiency

By GEORGE J. STIGLER*

Harvey Leibenstein called attention in an influential article (1966) to a source of economic inefficiency which was given the awful name of X-[in]efficiency. He cited studies in which misallocations of resources due to monopoly or tariffs had trifling social costs, whereas simple failure to attain the production frontier apparently led to social losses of a vastly greater magnitude. I propose to argue that this type of inefficiency can usefully be assimilated into the traditional theory of allocative inefficiency.

It is a question (to be discussed below) whether one ascribes failures to reach the ultimate limits of output from given inputs in any state of technology to inadequacy of knowledge alone, or adds also inadequate "motivation." Leibenstein (1966) separates the two:

> It is obvious that not every change in technique implies a change in knowledge. The knowledge may have been there already, and a change in circumstances induced the change in technique. In addition, knowledge may not be used to capacity just as capital or labor may be underutilized. More important, a good deal of our knowledge is vague. [pp. 404–05]

He ascribes increases in X-efficiency to 1) increases in motivational efficiency—workers are stimulated by incentive pay, or management by competition or other adversities; and 2) improvements in the inefficient markets for knowledge. I shall first deny the propriety of treating changes in motivation as a source of changes in output, and then proceed to the question of knowledge.

I. Motivational Losses

Leibenstein (1973) has emphasized that X-inefficiency arises largely from losses of output due to motivational deficiencies of resource owners:

* Graduate School of Business, University of Chicago.

> [F]or the same set of human inputs purchased and the same knowledge of production techniques available to the firm, a variety of output results are possible. If individuals can choose, to some degree, the APQT bundles [choice of Activity, Pace, Quality of work, Time spent] they like, they are unlikely to choose a set of bundles that will maximize the value of output. [p. 768]

If management seeks to impose output-maximizing $APQT$ bundles on the workers, indeed, these assignments of tasks would likely be " . . . less efficient than those that individuals would choose themselves under an acceptable set of [managerial] restraints" (p. 769).

In this case, and in every motivational case, the question is: what is output? Surely no person ever seeks to maximize the output of any one thing: even the single proprietor, unassisted by hired labor, does not seek to maximize the output of corn: he seeks to maximize utility, and surely other products including leisure and health as well as corn enter into his utility function. When more of one goal is achieved at the cost of less of another goal, the increase in output due to (say) increased effort is not an increase in "efficiency"; it is a *change* in output.

The concept of motivational efficiency seems to extend also to the task of getting a "predetermined output" from hired factors (see Leibenstein, 1966, pp. 408, 412). There are important and pervasive problems in all contracts between people, in seeking the fulfillment of the reciprocal contractual promises, and substantial resources are necessary to enforce the agreements (see Armen Alchian and Harold Demsetz). Both the avoidance of unpleasant tasks and the enforcement activity designed to curtail this avoidance can be carried on to the utility-maximizing degree and generate no inefficiency in producing utility. Output and utility would be larger if resources were not

necessary to the enforcement of contracts, but output and utility would also be larger if water boiled at 180°F or a day had 25 hours. New techniques of contract enforcement may be as productive as other improvements of technology.

Thus X-inefficiency attributed to motivational factors characterizes as inefficiency either the existence and pursuit of other desired outputs or the expenditure of resources required for the optimal enforcement of contracts. This tunnel vision of output seems entirely unrewarding: it imposes one person's goal upon other persons who have never accepted that goal. There is no waste in this sort of X-inefficiency: waste is a foregone product that could be acquired for less than its cost.

Leibenstein achieves much of the importance of motivation in X-efficiency by that ancient and powerful scientific technique, definition. When he copiously illustrates that "changes in incentives will change productivity per man" (1966, p. 401), he is assigning motivation an independent role whereas ordinary economic language would classify the methods of remuneration of employees as a part of the state of technology. Again, when an Egyptian petroleum refinery becomes more "efficient" with a change of management, we are told that "It is quite possible that had the motivation existed in sufficient strength, this change could have taken place earlier" (1966, p. 398). Potential motivation could indeed rewrite all history: if only the Romans had tried hard enough, surely they could have discovered America. (Thus motivation can be invoked to explain every unperformed task that is physically possible, no matter how unrewarding.) We may sympathize with Leibenstein's desire to associate his X-efficiency with economic behavior, but this shotgun marriage is not fertile.

II. The State of Technology

The near-universal tradition in modern economic theory is to postulate a maximum possible output from given quantities of productive inputs—this is *the* production function—and to assert that each firm operates on this production frontier as a simple corollary of profit or utility maximization. The merit of this conventional tradition is also its demerit: it eliminates the problem of the choice of technology.

Alfred Marshall followed an entirely different approach, and it is remarkable that he had virtually no followers. He proposed to characterize production possibilities by the *average* outputs obtained from given inputs, and in particular labelled the user of this average relationship the Representative Firm:

> We shall have to analyse carefully the normal cost of producing a commodity, relatively to a given aggregate volume of production; and for this purpose we shall have to study *the expenses of a representative producer* for that aggregate volume. On the one hand we shall not want to select some new producer just struggling into business, who works under many disadvantages, and has to be content for a time with little or no profits, but who is satisfied with the fact that he is establishing a connection and taking the first steps towards building up a successful business; nor on the other hand shall we want to take a firm which by exceptionally long-sustained ability and good fortune has got together a vast business, and huge well-ordered workshops that give it a superiority over almost all its rivals. But our representative firm must be one which has had a fairly long life, and fair success, which is managed with normal ability, and which has normal access to the economies, external and internal, which belong to that aggregate volume of production; account being taken of the class of goods produced, the conditions of marketing them and the economic environment generally. [p. 317, and Bk. IV, ch. 13]

Marshall suggested two causes of variation among firms in costs of a given output: the age of the firm (which he emphasized), and variations in entrepreneurial capacity. Strictly speaking, the latter element (the departure from "normal ability") is inappropriate: differences in quality of an input do not lead to differences in outputs from given inputs.

The reason Marshall's approach was not adopted by the science is lucidly presented in the leading attack that was made on the

representative firm by Lionel Robbins. In a once-famous essay, Robbins argued persuasively that when costs of firms differed because of quality of entrepreneurs (or other inputs), the differences in productivity would be reflected in differences in profits (or other input prices). Just as differences in efficiency of workers are reflected in their wages, so differences in entrepreneurial skills (including the choice of technology) will be reflected in their "profits." He states:

> There is no more need for us to assume a representative firm or representative producer, than there is for us to assume a representative piece of land, a representative machine, or a representative worker. [p. 393]

Robbins was of course correct: the Representative Firm is not needed to reconcile the existence of differences among entrepreneurs with the existence of stable competitive equilibrium.

What one may lament, however, is the failure of Robbins and Leibenstein, and all of us in between, to recognize the problem of determining which technologies will be used by each firm (and, for that matter, each person). The choice is fundamentally a matter of investment in knowledge: the costs and returns of acquiring various kinds and amounts of technological information vary systematically with various characteristics of a firm: its size, the age of its present capital assets, the experience of its managers, the prospects of the trade. No attention has been paid by economists to the analysis of the optimal amount of technological knowledge that a firm should possess. Leibenstein deserves credit for reviving this Marshallian question, but his attention to X-inefficiency as the explanation is an act of concealment: it simply postulates the differences in technology among firms which should be explained.

III. The Interpretation of Output Differences

We observe two farmers with reasonably homogeneous land and equipment, who nevertheless obtain substantially different amounts of corn. We measure this corn output over some period of time to reduce the effects of stochastic variation (i.e., unenumerated inputs such as weather). The observed variation is due, perhaps, to differences in knowledge, including the knowledge of technology or the knowledge of how far to carry the application of each productive factor. The farmers will differ in the cost of learning new things or the expected returns from new knowledge—one may be planning to leave agriculture shortly—so they "rationally" devote different amounts of resources to acquiring knowledge. Or one is simply more intelligent than the other, and learns more quickly or thinks more precisely (for example, makes fewer mistakes in arithmetic).

The effects of these variations in output are all attributed to specific inputs, and in the present case chiefly to the differences in entrepreneurial capacity. In neoclassical economics, the producer is always at a production frontier, but his frontier may be above or below that of other producers. The procedure allocates the foregone product to some factor, so in turn the owner of that factor will be incited to allocate it correctly.

Leibenstein does not attempt to understand the allocation of "inefficient" resources, and hence does not see the necessity for attributing his X-inefficiency to specific inputs. Just as automobile accidents are palpable inefficiencies to many people so X-inefficiency is a palpable inefficiency to Leibenstein. But accidents and "inefficiencies" are associated with returns as well as costs, and a useful theory must take both sides of their roles into account.

Indeed, Leibenstein's apparatus does not allow him to analyze effectively concrete economic problems. Consider his argument that monopoly is less efficient than competition. To reach this result, he must assume that 1) monopolists do not maximize profits, and 2) competitors are driven closer to "the" minimum costs by the entry of new rivals, some of whom are efficient, by a Darwinian process. The first assumption is an abandonment of formal theory, and one which we shall naturally refuse to accept until we are given a better theory. It "solves" the question of the effect of monopoly on efficiency without argument or evidence. The latter assumption of competitive selection coolly

ignores the problem of general equilibrium (where do the driven-out entrepreneurs go?, and where do the efficient entrepreneurs come from?), and fails to demonstrate (or even to argue) that inflows and outflows of entrepreneurs of various qualities will converge on a high-efficiency equilibrium in each competitive industry.

Earlier I defined waste as the situation in which foregone products could be obtained for less than they cost. Waste can arise *ex post* because *ex ante* plans rested upon erroneous predictions. This type of waste is unavoidable, although its magnitude is subject to control. Waste can also arise in the absence of uncertainty if the economic agent is not engaged in maximizing behavior. Unless one is prepared to take the mighty methodological leap into the unknown that a nonmaximizing theory requires, waste is not a useful economic concept. Waste is error within the framework of modern economic analysis, and it will not become a useful concept until we have a theory of error.

REFERENCES

A. A. Alchian and H. Demsetz, "Production, Information Costs, and Economic Organization," *Amer. Econ. Rev.,* Dec. 1972, *62,* 777–95.

H. Leibenstein, "Allocative Efficiency vs. 'X-Efficiency'," *Amer. Econ. Rev.,* June 1966, *56,* 392–415.

———, "Competition and X-Efficiency: Reply," *J. Polit. Econ.,* May/June 1973, *81,* 765–77.

A. Marshall, *Principles of Economics,* 8th ed., London 1948, 317.

L. Robbins, "The Representative Firm," *Econ. J.,* Sept. 1928, *38,* 387–404.