

MARGINAL ANALYSIS AND EMPIRICAL RESEARCH

By FRITZ MACHLUP*

Certain critics of "conventional" economic theory from time to time voice surprise at the general acceptance of marginalism and at "the confidence of the textbook writers in the validity of the marginal analysis."¹ They disapprove of allowing the principle of marginalism to play the rôle of a fundamental postulate in the teaching of economics.

Marginalism Implied in the Economic Principle

These critics would probably revolt against all those definitions of economics which contain marginalism as an implicit criterion. Marginalism, as the logical process of "finding a maximum," is clearly implied in the so-called *economic principle*—striving to achieve with given means a maximum of ends.

Economics in a narrow sense is confined to such aspects of conduct as can be explained with reference to the principles of maximizing satisfaction, income, or profit. Under definitions of this sort any deviations from the marginal principle would be extra-economic. Yet, to refuse to deal with any type of business conduct that cannot qualify by the strict standards of marginalism may justly be regarded as a lazy man's excuse. If certain types of business conduct can be found in reality with regularity and consistency, it is undoubtedly desirable to analyze them regardless of their "economic rationale."² And if some of these allegedly "non-economic" aspects of conduct can be explained within the conceptual framework of economics, one may prefer definitions which admit behavior types not strictly subject to marginal analysis among the proper subject matter of economic theory.

Interpretation of Business Behavior

To recognize the study of certain types of merely "traditional" conduct as legitimately within the province of economic theory is one thing; it is another to accept as correct the interpretations of business behavior offered by the critics of marginal analysis. Unable to see how marginal analysis can be applied to their material, these critics have concluded that marginalism should be discarded. It can be shown, however, that

* The author is professor of economics at the University of Buffalo.

¹ Richard A. Lester, "Shortcomings of Marginal Analysis for Wage-Employment Problems," *Am. Econ. Rev.*, Vol. XXXVI, No. 1 (Mar., 1946), p. 63.

² Cf. the admonition that "if an economist finds a procedure widely established in fact, he ought to regard it with more respect than he would be inclined to give in the light of his own analytic method." R. F. Harrod, "Price and Cost in Entrepreneurs' Policy," *Oxford Economic Papers*, No. 2 (1939), p. 7.

the alleged "inapplicability" of marginal analysis is often due to a failure to understand it, to faulty research techniques, or to mistaken interpretations of "findings."

This is not to deny that a goodly portion of all business behavior may be non-rational, thoughtless, blindly repetitive, deliberately traditional, or motivated by extra-economic objectives. But the material thus far presented as the result of empirical research has not proved what the analysts intended to prove. In some instances their findings were the result of careful research, based on a thorough knowledge of economic theory, but their interpretations were still questionable. In other instances the whole approach of the research project was so faulty that the findings as well as the interpretations are all but worthless except as targets for critical discussion.

I. MARGINAL ANALYSIS OF THE SINGLE FIRM

Any attempt to "test" marginalist theory through empirical research presupposes full understanding of the theory. It is necessary to know precisely what the theory says, what it implies, and what it intends to do. Since it has been developed gradually over a period of more than a century,³ it will not suffice to take any particular writer as one's authority or any particular exposition as one's text. Earlier versions lack the necessary refinements and methodological foundations; later formulations often take for granted necessary assumptions or qualifications made in previous expositions. To criticize the theory because of the errors and omissions in any treatise, however representative, is unfair.

The following statement of essential elements in the marginalist analysis of the single business firm attempts merely to give major emphasis to points often overlooked or misunderstood.

The Determination of Output and Employment

The theory of the "equilibrium of the single firm" is not as ambitious as is often believed. It does not attempt to give all the reasons why a given firm makes the type or quality of product which it makes; why it produces the output that it produces; why it employs the workers that it employs; or why it charges the prices that it charges. It is probably an understatement of the importance of the historical situation when Hall and Hitch modestly remark: "There is usually some element in the prices ruling at any time which can only be explained in the light of the history of the industry."⁴ The phrase "usually some element" does

³ Cournot was among the earlier expositors of marginal analysis of the single firm.

⁴ R. L. Hall and C. J. Hitch, "Price Theory and Business Behavior," *Oxford Economic Papers*, No. 2 (1939), p. 33.

not do justice to the part played by historical antecedents in the determination of product, output, employment, and prices. The rôle of the past in shaping the actual conditions under which the firm operates, in developing the routine of its responses to changes in conditions, and in impressing it with experiences which have taught it to size up and anticipate these changes as the basis for its decisions—this rôle is by no means denied by marginal analysis. The rôle of the past in the process of adjusting the present to the anticipated future is essential in all theory of human conduct. It is implied in the very attempt of constructing a pattern of behavior of the single firm.

Instead of giving a complete explanation of the “determination” of output, prices, and employment by the firm, marginal analysis really intends to explain the effects which certain *changes* in conditions may have upon the actions of the firm. What kind of changes may cause the firm to raise prices? to increase output? to reduce employment? What conditions may influence the firm to continue with the same prices, output, employment, in the face of actual or anticipated changes? Economic theory, static as well as dynamic, is essentially a theory of adjustment to change. The concept of equilibrium is a tool in this theory of change; the marginal calculus is its dominating principle.

A. *Marginal Revenue and Cost of Output*

Subjectivity of Cost and Revenue

The proposition that the firm will attempt to equate marginal cost and marginal revenue is logically implied in the assumption that the firm will attempt to maximize its profit (or minimize its losses). It should hardly be necessary to mention that all the relevant magnitudes involved—cost, revenue, profit—are subjective—that is, perceived or fancied by the men whose decisions or actions are to be explained (the business men)—rather than “objective”—that is, calculated by disinterested men who are observing these actions from the outside and are explaining them (statisticians and economists as theorists—not as consultants).

The marginal cost that guides the producer is the addition to his total cost which he expects would be caused by added production. An outside observer, if he had expert knowledge of the production techniques and full insight into the cost situation of the producing firm, might arrive at a different, “objective” figure of the firm’s marginal cost; but what the observer thinks is not necessarily the same as what the producer thinks. The producer’s actual decision is based on what he himself thinks; it is based on “subjective” cost expectations.

One may perhaps assume that the producer is intensely interested in knowing his cost and that, in general, he has the experience which

enables him to know it. Yet, one must not assume that all producers "really" know their cost in the sense in which an efficiency expert would determine it; several of them may lack the interest or experience; they may not find it worth their while to dig too deeply into the mysteries of their business. (After all, we know that there are good business men and bad, and that the majority is somewhere between good and bad.) But this does not invalidate the proposition that the producer is guided by marginal cost.⁵

The same thing is true with regard to price expectations and sales expectations. It is the "demand as seen by the seller" from which his revenue expectations stem. The increase in demand which is relevant in the analysis of the firm need not be "the real thing"; it may precede an "actual" increase in demand, lag behind it, or be entirely imaginary. The business man does what he does on the basis of what *he* thinks, regardless of whether you agree with him or not.

Marginal analysis of the firm should not be understood to imply anything but subjective estimates, guesses and hunches.

The Range of Price and Output Variations

Beginning students of economics who watch their instructor draw demand and cost curves covering half the blackboard may be misled into believing that the business man is supposed to visualize the possibilities of producing and selling amounts of output ranging from almost zero up to two or three times the amounts that he is currently producing and selling; that the business man is supposed to figure out how much he might be able to sell at prices several times as high as the current price, and how much at prices only one-half or one-third as high. The curve draftsman, indeed, seems to ascribe extraordinary powers of imagination to the business wizards.

Misunderstandings of this sort, and erroneous criticisms of marginal analysis, could be avoided if it were made clear to the students that the length of the curves, *i.e.*, the wide range they cover, was chiefly designed to enable those in the back rows of the class room to make out what goes on on the blackboard; and to permit them to practice curve analysis without using magnifying glasses. The range of possibilities—prices, sales, outputs—which a business man may have in mind is probably quite narrow. Rarely will a business man bother pondering the probable effects of a price increase or cut by 50 per cent; but he may easily think about what a 10 or 15 per cent price change might do to his sales; or what discount it might take to land some additional orders.

The principles of analysis are not altered by the realization that the

⁵ One may wish, of course, to qualify any social implications of the proposition once the subjective character of the relevant cost data is recognized.

alternatives which business men weigh concerning prices or production volumes cover a much more moderate range than the curves which teachers of economics draw to depict the pattern of marginal calculus.

The Time-Range of Anticipations

In view of the known attempts to derive statistical cost curves from accounting data—which of necessity refer to conditions of the past—it is important to mention that the marginal cost and marginal revenue concepts in the analysis of the equilibrium of the firm refer to expectations of future conditions. To be sure, past experience is always in the background of anticipations of the future, and past accounting records may form a firm point of departure for evaluating prospective and hypothetical cost and revenue figures. But anticipations alone are the relevant variables in the marginal calculus of the firm.

What is the time-range of the significant anticipations? How far into the future do they reach, and what period, if any, is given special emphasis? Is tomorrow more important than next year or several years hence? Is it the “short run” or the “long run” which controls current action?

When a firm wishes to increase production, it usually has a choice of expanding the equipment and productive capacity of its plant or of stepping up the output of the existing plant with unchanged equipment. If productive capacity is already well utilized, the marginal cost of producing larger outputs will be higher in the existing establishment with unchanged equipment than in an establishment with adjusted, increased equipment. If several degrees of adjustment in the productive equipment are possible, several marginal cost functions will be “given” and several different outputs will be “the equilibrium output” under given sales expectations.

To cope with these problems economists have made the distinction between the “short period,” assuming no adaptation of equipment, and the “long period,” assuming complete adaptation of equipment. Students often believe that the latter period is called “long” because it takes a long time to expand the plant. This need not be the case. A better understanding of the concepts might be achieved by associating the degree of planned plant adjustment with the length of time for which the changed production volume is expected to be maintained. If an increased demand is expected to prevail for a short period only, it will not pay to invest in plant expansion, and “short-run cost” will determine output. On the other hand, if demand is expected to continue at the higher level for a sufficiently long period, an expansion of the establishment will be considered a profitable investment, and “long-run cost” will determine output. Needless to say, many intermediate periods, that

is, several degrees of plant adjustment with different marginal cost conditions, may exist.

On the basis of this reasoning one will recognize it as a misunderstanding to argue that short-run cost is of controlling influence on the ground that we always live and work in the short period. The duration for which demand conditions are expected to prevail will determine the relevant "period" of cost anticipations. Of course, this relevance is again subjectively determined, not by the "objective" judgment of the economist.

The time-range of the anticipations with regard to the demand and selling outlook is subject to similar considerations. It is a mistake to think that the relevant "period" for demand and marginal revenue expectations is determined by the length of time it takes for today's production to reach the market.⁶ If a price reduction is apt to spoil the market for a long time to come, or a price increase to harm customer loyalty, the effects on future profits will hardly be neglected in considering current actions. If a firm were to regard a certain price change as a desirable step for the time being, but feared that a later reversal might be difficult or costly, it would weigh this anticipated future cost or loss against the short-run benefit.

Anticipations of this sort, complementary or competing with one another, are not exceptions to marginal analysis but are part and parcel of it. To be sure, when an instructor teaches graphical analysis, he will do well to abstract from complicated cost and revenue anticipations and to concentrate on those that can be neatly packed away in geometric curves.

The Numerical Definiteness of the Estimates

The geometric curves and arithmetic schedules by which the instructor presents marginal cost and marginal revenue of the firm seem to leave no room for doubt that these anticipations take the form of estimates of definite numerical values. While this may be necessary for teaching purposes, it should not mislead the student into believing that every action of the business man is in fact the result of a conscious decision, made after careful calculations of differential revenue and cost.

Business men do not always "calculate" before they make decisions, and they do not always "decide" before they act. For they think that they know their business well enough without having to make repeated calculations; and their actions are frequently routine.⁷ But routine is based on principles which were once considered and decided upon and have then been frequently applied with decreasing need for conscious

⁶ Richard A. Lester, *Economics of Labor* (New York, 1941), p. 181.

⁷ See George Katona, "Psychological Analysis of Business Decisions and Expectations," *Am. Econ. Rev.*, Vol. XXXVI, No. 1 (Mar., 1946), p. 53.

choices.⁸ The feeling that calculations are not always necessary is usually based upon an ability to size up a situation without reducing its dimensions to definite numerical values.⁹

The business man who is persuaded to accept a large order with a price discount or some other concession usually weighs the probability that he will have to make the same concession to his other customers. This is one of the business man's considerations included in the "calculation" of marginal revenue. In order to explain this to the student, or to reduce it to curves and schedules, the economics teacher makes "exact" calculations; in order to make up his mind whether to take or reject the order, the business man ordinarily needs no arithmetic, mental or written, and indeed needs no concrete figures. Yet his reasoning or his routine behavior is most conveniently analyzed in terms of marginal revenue.

When the marginal revenue is negative, that is to say, where gross receipts after accepting the additional order (with the price concession) would be smaller than without it, no further consideration is necessary. But if the dollar volume of sales can be increased by accepting the order (taking full account of all repercussions on future marketing possibilities), the business man must take another step in his reasoning: will it pay to make more sales in view of the additional cost of producing the larger output? If conditions have not changed, he will not have to make new calculations; if changes have occurred or are expected, some figuring may be required. But it is a type of figuring for which usually no accounting records are consulted, no memoranda prepared and of which no records are made. Often the business man can do this "figuring" in his head; if not, he may take a piece of scrap paper, jot down a few round numbers, reach his conclusion, and throw the paper in the waste basket.

The theorist's contention that such reasoning is typically based either on additional cost or on total cost—and hence most conveniently described in terms of marginal cost—is contradicted by certain empirical researchers who claim that most business men calculate on the basis of average cost even if they lose money by doing so. With this contradiction we shall deal later.

Non-Pecuniary Considerations

Marginal analysis of the equilibrium of the single firm rests on the assumption that the business firm attempts to maximize its profits. To

⁸ Discussing the difference between "routine behavior" and "genuine decisions," Dr. Katona explains with regard to routine actions that "principles, well understood in their original context, tend to be carried over from one situation to another." *Ibid.*, p. 49. Genuine decisions are made when expectations "change radically." *Ibid.*, p. 53.

⁹ Although I do not know either the width or length of my automobile, I am quite capable of making adequate comparisons between these magnitudes and the space between two parked cars, which I estimate again without thinking of feet, inches, or any numbers.

make this assumption is not to deny that the men who run a business may be motivated also by other considerations.

That a business man is motivated by considerations other than the maximization of money profits does not necessarily make his conduct "uneconomic." The economic theorist finds no difficulty in fitting into the pattern of "economic" conduct (that is, into the conceptual scheme of consistent maximization of satisfaction within a given preference system) the householder and consumer who makes donations to friends or the church; or the seller of labor services who chooses a badly-paying but less strenuous job in preference to one that pays more but calls for more exertion. Likewise, there is nothing essentially "uneconomic" in the conduct of a business man who chooses to pay higher prices for raw material purchased from a fraternity brother, or to sell at a special discount to members of his church, or who refrains from embarking on a promising expansion of his business because he prefers an easier life.

There are economic theorists who would include considerations of this sort among the data for the marginal calculus of the firm. The satisfaction from favoring his friends through higher purchase prices or lower selling prices is a special reward or "revenue" to the business man; he may ask himself how much it is worth to him, and we may conceivably add it to his revenue curve. To give up an easier life, expend greater efforts and increase his worries are among the business man's "costs" when he considers an expansion of his business; we may conceivably add it to his "cost" curve. Any number and type of non-pecuniary sacrifices and rewards could thus be included, at some sort of "money equivalent," among the costs and revenues that make up the profits of the firm: the marginal calculus of the firm would become all-inclusive.

It seems to be methodologically sounder if we do not reduce the non-pecuniary satisfactions and dissatisfactions (utilities and disutilities) of the business man to money terms and do not try to make them part of the profit maximization scheme of the firm. If *whatever* a business man does is explained by the principle of profit maximization—because he does what he likes to do, and he likes to do what maximizes the sum of his pecuniary and non-pecuniary profits—the analysis acquires the character of a system of definitions and tautologies, and loses much of its value as an explanation of reality. It is preferable to separate the non-pecuniary factors of business conduct from those which are regular items in the formation of money profits.

This methodological controversy is not too important. Not much depends on whether non-pecuniary considerations of the business man are translated into money terms or, instead, treated as exceptions and qualifications in the explanation of typical business conduct. The pur-

pose of the analysis of the firm is not to explain all actions of each and every firm in existence; we are satisfied if we can explain certain strong tendencies in a representative sector of business. The chief aim of the analysis, moreover, is to show the probable effects of certain changes; if the direction in which output or price is likely to move as a result of a certain change in "data" is not affected by the existence and strength of non-pecuniary factors in business conduct, their inclusion in or exclusion from the marginal analysis of the firm is not a crucial matter.

As a matter of fact, the nature, strength and effects of non-pecuniary considerations in business behavior are problems that need to be investigated. One may presume that producing larger production volumes, paying higher wage rates, or charging lower product prices than would be compatible with a maximum of money profits may involve for the business man a gain in social prestige or a certain measure of inner satisfaction.¹⁰ It is not impossible that considerations of this sort substantially weaken the forces believed to be at work on the basis of a strictly pecuniary marginal calculus.

During the war we were able to observe that patriotism was a strong force in the production policy of American business. There can be no doubt that many firms produced far beyond the point of highest money profits. To be sure, they made large profits, but in many instances they could have made still more money without the last, particularly expensive, portions of output. Their conduct was not defined by the principle of maximization of money profits.¹¹

Another of the possibly important qualifications in the analysis of the firm refers to the conflict of interests between the hired managers and the owners of the business. The interest of the former in inordinately large outlays or investments may be capable of description in terms of a

¹⁰ A gain in social prestige may sometimes increase the good will of a firm on which it expects to cash in later. If such a gain is an aim of the firm's policy, it should be treated as a part of its pecuniary considerations. For example, a firm may grant extraordinarily high wage rates as a part of its selling and advertising expense; that is to say, it may hope that its "generous labor policy" will make its products more popular. A portion of current labor cost of the firm would then properly be allocated to future rather than current output.

¹¹ Observance of laws and regulations presents a special problem for the analysis of business conduct. It will depend on business morals whether prohibited, unlawful alternatives may be regarded as definitely excluded and therefore non-existent; or whether they may be considered as possibilities subject only to certain peculiar risks. Assume, for example, that a price ceiling is fixed for the sale of a product, and fines are provided for violations. To the business man who is unconditionally law-abiding the ceiling price is the only possible price, regardless of how insistently some of his customers may tempt him with higher bids. To the business man, however, who abides by the law only because of the risk of being found out and fined, "demand prices" above the ceiling are real possibilities and the risks of penalties are additions to cost or deductions from revenue. If the sanctions for violations include jail sentences, the risk becomes largely non-pecuniary and it is up to the potential violator, or to the theorizing economist, whether or not that risk will be "converted" into money terms. Black-market prices are in part the result of such risk conversions.

pecuniary calculus, but it is not maximization of the firm's profits which serves here as the standard of conduct. Maximization of salaries and bonuses of professional managers may constitute a standard of business conduct different from that implied in the customary marginal analysis of the firm. The extent to which the two standards would result in sharply different action under otherwise similar conditions is another open question in need of investigation. At this juncture we know only that a qualification must be made. How much it may modify the results of marginal analysis of the single firm we do not know.

B. *Marginal Productivity and Cost of Input*

The Firm, the Industry, the Economy

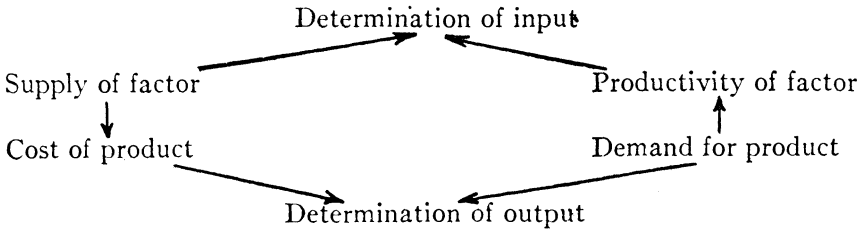
Marginal productivity has different meanings in the equilibrium theories of the single firm, the industry, and the whole economy. In the theories of demand for particular "factors of production" (productive services) by the industry or economy as a whole marginal productivity analysis is of another methodological character than in the theory of factor employment by the individual firm: the level of abstraction and the frame of reference are different.

In this article we are concerned only with the analysis of the single firm. Like marginal product cost and marginal revenue in the theory of the firm's output, marginal factor cost and marginal productivity are the variables in the theory of the firm's input.

Determination of Input and Output

In a sense, the determination of input on the basis of factor cost and factor product is merely the reverse side of the determination of output on the basis of product cost and revenue. In the former, the cost of and revenue from employing additional factors are balanced; in the latter, the cost of and revenue from producing additional product are balanced. Before we draw curves for the cost of production of a good, we must assume that the supply curves of the factors of production are known, because the buying prices of factors are among the things that make up production cost. Before we draw curves for the revenue productivity of a factor we must assume that the demand curves for the products made with the help of this factor are known, because the selling prices of products are among the things that make up factor productivity. Hence, in each pair of curves one of the curves comprises the data shown in one curve of the other pair.

The interrelationship between the four curves (or functions) can be shown schematically as follows:



A fifth set of data, the production function, showing the technological transformation of factors into products, is implied in both pairs of curves: in the analysis of output it is among the data determining the cost of production; in the analysis of input it is among the data determining the productivity of the factor.¹²

These remarks should make it clear that neither of the two analyses is prior to the other. They are of strictly equal rank, merely two ways of looking at the same thing, namely, the conduct of a single firm maximizing its profits. The only difference is that the significant magnitudes of the analysis are, on the one side, units of factors (such as labor hours) and, on the other side, units of product.

Marginal Net Revenue Productivity

When we speak in the analysis of the firm of "marginal productivity" of a factor, this is an abbreviation for longer but synonymous expressions such as "marginal value productivity" or "marginal net revenue productivity."

The following steps are pedagogically expedient in explaining the concept of marginal net revenue productivity:

- (1) Determine by how much a given physical volume of production, X , is increased if the employment of a particular factor is increased slightly (*e.g.*, by one unit), and call the output increase the factor's "marginal physical product," MPP .
- (2) Determine the selling price, P , at which MPP can be sold.
- (3) Multiply MPP by P in order to obtain the "value of the marginal physical product," $VMPP$.
- (4) Determine whether X , because of the sale of MPP , has to be sold at a price lower than it would sell if MPP were not sold; if so, multiply this price reduction, ΔP , by X , and obtain the "revenue loss on sales because of price cut," $X \Delta P$.

¹² This shows that the customary analysis lacks elegance. Production cost and factor productivity are "derived" rather than "original" data. One could do more elegantly with only three sets of data: (a) the possibilities of buying productive services (the factor supply function), (b) the possibilities of transforming them into products (the production function), and (c) the possibilities of selling the products (the product demand function).

- (5) Deduct $X \Delta P$ from $VMPP$ in order to obtain the "marginal gross revenue product," $MGRP$.
- (6) Determine whether the production of MPP was connected with increased or decreased outlays for any other complementary or substitutable means of production (materials, fuel, lubricants, labor of any sort, capital funds, wear and tear of equipment, etc.), exclusive of the factor in question, and call them (positive or negative) "incidental expenses," ΔC .
- (7) Deduct ΔC from $MGRP$ in order to obtain the "marginal net revenue product," $MNRP$.

The use of the word "revenue" as an adjectival modifier is preferred by many writers in order to stress (a) the distinction between physical product and money product, and (b) the fact that marginal revenue is less than selling price if it takes a price cut to dispose of additional output. The use of the word "net" is preferred in order to stress the fact that additional output will rarely be produced efficiently by increasing the employment of one particular factor while leaving all other outlays unchanged; as a rule, some other adjustments will be appropriate. That "marginal productivity" refers regularly to a net revenue product has been clear to economic theorists for over fifty years.¹³

Technology, Market and Supply Conditions

The marginal net revenue product of a factor, at some level of employment, becomes zero or negative. This may be due to technological difficulties—shown in step (1) of the above scheme—or to difficulties in marketing—shown in step (4)—or to difficulties with other supplies and expenses—shown in step (6).

On the other hand, it is possible that both the marginal physical product and the marginal gross revenue product are zero and, nevertheless, the marginal net revenue product is positive. This will be the case if additional units of factor are used only to secure "incidental reductions in expenses" for other means of production (*i.e.*, substitution) rather than an increased production volume. For example, an additional unskilled laborer may be employed as another watchman to reduce the "use" of certain materials which are in heavy demand outside of the plant. Or he may be employed to dust or cleanse certain valuable equipment and thus reduce outlays for repairs or replacements. Substitution

¹³ *E.g.*, "the net product . . . is the net increase in the money value of . . . total output after allowing for incidental expenses." (Alfred Marshall, *Principles of Economics*, 8th Edition, p. 521.) For a more detailed discussion of the concept, see my essay, "On the Meaning of the Marginal Product," *Explorations in Economics*, Contributed in Honor of F. W. Taussig (New York and London, 1936), pp. 250-63.

of this sort is nearly always possible¹⁴ and will usually make for positive marginal net revenue productivities even where marginal gross revenue productivities are negative because of limitations in the demand for the product.

Marginal productivity reflects all sorts of technological possibilities. An increased amount of the factor may be used (a) for reducing other expenses without increasing total output (substitution in the narrow sense), (b) for increasing total output with no or few adjustments in the use of other factors (substitution in a wider sense), and (c) for increasing total output with corresponding increases in the use of other factors (inclusive of long-run adjustments, possibly without any substitution). In the last case the incidental expenses will certainly absorb the major portion of the marginal gross revenue product.

Marginal productivity reflects also all possible situations in the demand for the product. If demand is completely inelastic beyond a certain volume, that is, if additional output is not saleable at all, the effect upon marginal productivity is not any worse than if larger outputs can be marketed at severely reduced prices. For whenever the elasticity of demand is less than unity, gross revenue from larger outputs would be lower than from smaller outputs. Hence, the marginal gross revenue product of the factor would become negative. Possibilities of landing additional orders at a price discount but without affecting the rest of the business (that is, possibilities of price discrimination) would show in the fact that no deduction for revenue loss would have to be made from the value of the marginal physical product. Whatever views the firm may have concerning the market for its product are fully reflected in the marginal productivity of the factors employed.

Marginal productivity, finally, reflects all possible conditions of supply of complementary and substitutable factors. Extreme scarcity of a complementary factor may cause a most rapid decline in marginal productivity. Increased supply of a substitutable factor may drastically reduce the whole marginal productivity schedule.

While the conditions of supply of complementary and substitutable factors are among the data determining the marginal productivity of a particular factor of production, the conditions of its own supply are regarded as a separate matter. The "incidental expenses" of increased

¹⁴ The assumption of fixed coefficients of production sometimes affords convenient and permissible simplifications of analysis. But in actual fact, substitution is a practical possibility in almost any production. Beginners sometimes think that substitution of labor for capital must mean the scrapping of machines and shifting of their functions to hand labor. Better care or maintenance work for equipment, postponing the need for replacement, constitutes a clear case of substitution of labor for capital. Increased utilization of plant capacity with increased employment and output also raises the ratio of labor to capital and is another form of substitution.

employment of the factor do not include any of the cost of that factor. The cost of the factor itself is not a part of its marginal net productivity but, instead, is the counterpart with which a balance is sought.

Marginal Factor Cost

Where the supply of the factor is perfectly elastic at a given point, that is, where the firm may be able to employ an additional amount without having to pay for it a higher price per unit, the "marginal factor cost" is equal to the price of the factor (wage rate). If, however, by purchasing or employing more of a factor the firm bids up the price not only of the additional units of the factor but also of the units previously employed, this increase in outlay is a part of the cost of the additional employment. The additionally employed factors would cost the firm not only what they themselves are paid but also the incidental increase in the pay of their fellow factors.

Marginal factor cost, in other words, is the total increase in payment for the particular type of productive service: it consists of (1) the price (wage) paid to the additionally employed, and (2) the price increase (wage increase) paid for the amount of services employed before the addition. In the case of labor, these increases may be due to union action anticipated because of the increased demand for labor, or to the impossibility of discriminating against older employees when new ones can be attracted only at higher rates of pay.

In considering any increase in employment the employer will ask himself whether the additional services will "pay for themselves," that is, what they will cost him and what they will be worth to him. This is all that the economist means when he says that the employer, maximizing his profits, equates marginal factor cost with marginal productivity.

Monopoly, Monopsony, Discontinuities

Neither the existence of monopoly nor of monopsony need invalidate the proposition that the firm will equate marginal productivity and marginal cost of input. For any degree of monopoly is fully reflected in marginal net revenue productivity, and any degree of monopsony is fully reflected in marginal factor cost.¹⁵

¹⁵ To be sure, there may be a large difference between the price of the factor and the value of its marginal physical product. This difference is due to (a) the reduction in product price that the firm must grant to its customers in order to dispose of an increased output and (b) the increase in factor price that the firm must grant to its suppliers or employees in order to acquire an increased input. These two parts of the spread between the price of the factor and the value of its marginal physical product have been called (a) "monopolistic exploitation" and (b) "monopsonistic exploitation" of the factor. These terms, misleading in several respects, are merely to remind the student of the fact that the spread would not exist if the firm were (a) selling its products under pure competition and (b) buying its factors under pure competition.

Discontinuity of the marginal productivity and marginal factor cost curves, however, may make it impossible for the two magnitudes to be equal. If marginal factor cost at a certain level of employment is below marginal productivity but would be above it at the next higher possible level of employment, the firm will stop short of the latter. Moderate jerks from "marginal cost below revenue" to "marginal cost above revenue" are nothing unusual in arithmetic illustrations; in geometric curves they occur only under special assumptions.

For example, marginal net revenue productivity may precipitously drop at a given employment if the product is sold under certain oligopoly conditions (involving high elasticity of demand in the case of a price increase and low elasticity in the case of a price reduction¹⁶) and if the factor is not easily substitutable for other factors. The marginal factor cost curve might intersect this marginal productivity curve in its vertical portion. Likewise, marginal factor cost may precipitously rise at a given employment if the factor is bought or hired under certain oligopsony conditions (involving high elasticity of supply in the case of a reduction in the factor price and low elasticity in the case of a raise.¹⁷) The marginal productivity curve might intersect this marginal factor cost curve in its vertical portion. Under such circumstances the firm would be in equilibrium, with its profits maximized, at a volume of input (employment) at which marginal factor cost is below marginal productivity.

Subjectivity, Range, Concreteness

Almost everything that has been said in earlier sections concerning the meaning of marginal revenue and marginal cost of output holds true, *mutatis mutandis*, in regard of the meaning of marginal productivity and marginal cost of input. More specifically, we should emphasize that

- (1) the concepts are to be understood as referring to subjective estimates and conjectures;
- (2) the range of imagined variations of the magnitudes in question may be rather narrow;
- (3) the time-range of the relevant anticipations will depend on the circumstances of each case and will rarely be confined to the short run;
- (4) the estimates need not be reduced to definite numerical values;
- (5) non-pecuniary considerations may effectively compete with those pertaining to the maximization of money profits.

It is probably unnecessary to expatiate again on these points in connection with marginal productivity analysis. Only on the subject of

¹⁶ Under such oligopoly conditions the firm will maximize profits at a volume of output at which marginal revenue is above marginal cost.

¹⁷ Oligopsony in the labor market is probably not as frequent as oligopoly in the product market.

numerical definiteness does further discussion seem advisable, especially in view of what was said above about the concept of marginal net revenue productivity. The process by which this magnitude may be derived, involving seven separate "steps" and at least as many variables, is rather formidable. If this analytical pattern were taken as a realistic description in photographic likeness of the actual reasoning of the typical employer, the employer would have to be endowed with talents which only few possess in reality.

An analogy may explain the apparent contradiction.

The "Extreme Difficulty of Calculating"

What sort of considerations are behind the routine decision of the driver of an automobile to overtake a truck proceeding ahead of him at slower speed? What factors influence his decision? Assume that he is faced with the alternative of either slowing down and staying behind the truck or of passing it before a car which is approaching from the opposite direction will have reached the spot. As an experienced driver he somehow takes into account (a) the speed at which the truck is going, (b) the remaining distance between himself and the truck, (c) the speed at which he is proceeding, (d) the possible acceleration of his speed, (e) the distance between him and the car approaching from the opposite direction, (f) the speed at which that car is approaching; and probably also the condition of the road (concrete or dirt, wet or dry, straight or winding, level or uphill), the degree of visibility (light or dark, clear or foggy), the condition of the tires and brakes of his car, and—let us hope—his own condition (fresh or tired, sober or alcoholized) permitting him to judge the enumerated factors.

Clearly, the driver of the automobile will not "measure" the variables; he will not "calculate" the time needed for the vehicles to cover the estimated distances at the estimated rates of speed; and, of course, none of the "estimates" will be expressed in numerical values. Even so, without measurements, numerical estimates or calculations, he will in a routine way do the indicated "sizing-up" of the total situation. He will not break it down into its elements. Yet a "theory of overtaking" would have to include all these elements (and perhaps others besides) and would have to state how changes in any of the factors were likely to affect the decisions or actions of the driver.¹⁸ The "extreme difficulty of calculating,"¹⁹ the fact that "it would be utterly impractical"²⁰ to attempt to work out and ascertain the exact magnitudes of the

¹⁸ Very cautious drivers are apt to work with so wide safety margins that small changes in the "variables" may not affect the actions. Timid souls may refuse to pass at all when another car is in sight.

¹⁹ Lester, *Am. Econ. Rev.*, Vol. XXXVI, No. 1, p. 72.

²⁰ Lester, *ibid.*, p. 75.

variables which the theorist alleges to be significant, show merely that the *explanation* of an action must often include steps of reasoning which the acting individual himself does not *consciously* perform (because the action has become routine) and which perhaps he would never be *able* to perform in scientific exactness (because such exactness is not necessary in everyday life). To call, on these grounds, the theory "invalid," "unrealistic" or "inapplicable" is to reveal failure to understand the basic methodological constitution of most social sciences.

Imagine an empirical researcher attempting to test by a naïve questionnaire method the "theory of overtaking," questioning hundreds of drivers about their ability to estimate distances and speed, and to calculate the relevant time intervals and the degrees in which a small change in any one of the variables affected the result. Would he not obtain the most hopeless assortment of answers? Would not these answers support the conclusion that the assumptions of the theorists had been wrong and that one must look for other explanations? Yet I can hardly believe that any sensible person would deny the relevance of the enumerated variables and would contend, for example, that speed and distance of the approaching automobile could not have been taken into account by the driver passing the truck, because he was not good in mathematics.²¹

The Analysis of Change Needs No Exactness

The business man who equates marginal net revenue productivity and marginal factor cost when he decides how many to employ need not engage in higher mathematics, geometry, or clairvoyance. Ordinarily he would not even consult with his accountant or efficiency expert in order to arrive at his decision; he would not make any tests or formal calculations; he would simply rely on his sense or his "feel" of the situation. There is nothing very exact about this sort of estimate. On the basis of hundreds of previous experiences of a similar nature the business man would "just know," in a vague and rough way, whether or not it would pay him to hire more men.

The subjectivity of his judgments is obvious. Just as different drivers may reach different conclusions about the advisability of passing another car under given "objective" conditions, different business men will have different "hunches" in a given situation. The subordinates or partners of the man who makes a decision may sharply disagree with him; they may see the situation quite differently. They may be more optimistic about the possibilities of obtaining more orders with only

²¹ Driving at night, when he has nothing to go by except the size and brilliance of the headlights of the approaching cars, the experienced driver becomes conscious of the fact that in daytime he has better ways of sizing up their speed and distance. With reduced visibility he will "calculate" with greater safety margins.

slight price concessions or through increased sales efforts (which would raise both the marginal revenue and marginal productivity curves drawn by the theorist to characterize their considerations). Or they may be more certain about the technical possibility of achieving a larger output by certain production methods (which would lower the marginal cost curve, and could raise or lower the marginal productivity curves). Some decision, usually a routine decision without debate, is made, or at least some action is taken; and the decision or action is necessarily affected by the business man's conjectures concerning sales possibilities and production possibilities.

The way in which changes in the essential variables will affect the probable decisions and actions of the business man is not much different if the curves which the theorist draws to depict their conjectures are a little higher or lower, steeper or flatter. These curves are helpful to the student of economics in figuring out the probable effects of change — in learning in what direction output, prices and employment are likely to be altered, and under what circumstances increases or decreases are likely to be drastic or negligible. Better markets or higher costs are likely to affect business men of different vision or daring in rather similar ways; and any differences can be conveniently "typed" in terms of shapes, positions and shifts of the curves into which the theorist condenses the business men's conjectures.

Equipped with this understanding of the meaning and purposes of marginal analysis, we may proceed to a discussion of the findings of empirical research which purportedly failed to verify it — or by which it was deemed to be contradicted and disproved.

II. EMPIRICAL RESEARCH ON THE SINGLE FIRM

There is not as yet available any large amount of material derived from systematic empirical research on the business conduct of the single firm. But almost everybody interested in these questions has had occasional conversations with business men, and the impressions gained from such inquiries into the business men's experiences often form an empirical basis for the doubts which so-called "realistic" critics entertain of "theoretical" analysis.

I submit that the few systematic and the many casual researchers have often been misled by pitfalls of semantics and terminology and by a naïve acceptance of rationalizations in lieu of genuine explanations of actions.

Economists' Vocabulary and Business Language

The vast majority of business men have never heard of expressions such as elasticity of demand or supply, sloping demand curves, mar-

ginal revenue, marginal cost. If they do not know the words or the concepts, how can they be supposed to think in these terms? A scattered few of the men may have been exposed to such words and ideas in half-forgotten college courses, but they have found in practice they had no use for a vocabulary unknown to their associates, superiors, subordinates, and fellow business men. Thus the most essential terms in which economists explain business conduct do not exist in the business man's vocabulary. Does this not prove that the explanations are unrealistic or definitely false?

Only an inexperienced researcher could draw such a conclusion. The technical terms used in the explanation of an action need not have any part in the thinking of the acting individual. A mental process in everyday life may often be most conveniently described for scientific purposes in a language which is quite foreign to the process itself.

To ask a business man about the "elasticity of demand" for his product is just as helpful as inquiring into the customs of an indigenous Fiji Islander by interviewing him in the King's English. But with a little ingenuity it is possible to translate ideas from the business man's language into that of the economist, and *vice versa*. Questions such as "Do you think you might sell more of this product if you cut the price by 10 per cent?" or "How much business do you think you would lose if you raised your price by 10 per cent?" will evoke intelligent answers in most cases provided the questions are readily reformulated and adapted to the peculiarities of the particular man and his business. Often it will be necessary to know a good deal of the technology, customs and jargon of the trade, and even of the personal idiosyncrasies of the men, before one can ask the right questions. A set formulation of questions will hardly fit any large number of business men in different fields and, hence, questionnaires to be filled out by them will rarely yield useful results.

Rationalizations of Decisions or Actions

Psychologists will readily confirm that statements by interviewed individuals about the motives and reasons for their actions are unreliable or at least incomplete. Even if a person tries to reconstruct for himself in his memory the motives and reasons for one of his past actions, he will usually end up with a rationalization full of afterthoughts that may make his actions appear more plausible to himself. Explanations given to an interviewer or investigator are still more likely to be rationalizations in terms that may make the particular actions appear plausible and justified to the inquirer. In order to be understood (and respected) the interviewed person will often choose for his "explanations" patterns of reasoning which he believes to be recognized as "sound" and "fair" by others. Most of these rationalizations may be subjectively honest and

truthful. It takes an experienced analyst to disentangle actual from imaginary reasons, and to separate relevant from irrelevant data, and essential from decorative bits of the information furnished. Written replies to questionnaires are hopelessly inadequate for such purposes.²²

Questions of business policy are particularly difficult objects of inquiry because the business man usually is anxious to show by his answers that he is intelligent, well informed, and fair. The standards of fairness and business ethics to which he wishes to conform are often those which he believes are accepted by his lawyers, accountants, customers, competitors, fellow citizens, economists and whatnot. Only through detailed discussions of different situations and decisions, actual as well as hypothetical, will an investigator succeed in bringing out true patterns of conduct of the individual business man.²³

A. *Average Cost and Price*

One of the conclusions of casual or systematic empirical research on the business firm is that business men do not pursue a policy of maximizing profits, and of pricing according to the marginal cost and marginal revenue principle, but instead follow rules of pricing on the basis of average cost calculations even where this is inconsistent with profit maximization.

We shall attempt to reinterpret the findings of systematic research along these lines. For this purpose we must first clear up some misunderstandings which appear to have contributed to the support for the average-cost theory of pricing.²⁴

Averaging Fluctuating Costs and Prices

In discussions with business men I have found that two different types

²² Cf. George Katona, *Price Control and Business* (Bloomington, Ind., 1945), p. 210. He states that "only detailed interviews can probe into the motives behind business decisions."

²³ For further comments on the difficulties of good empirical research on business conduct, see my paper "Evaluation of the Practical Significance of the Theory of Monopolistic Competition," *Am. Econ. Rev.*, Vol. XXIX (1939), p. 233. After discussing the policies of my former business partners I concluded (p. 234): "An investigator who would have based his findings on their answers to questionnaires or even on personal interviews, would have come to erroneous results. An investigator who could have seen all the actually or potentially available statistics would have come to no results at all. The only possibility for a fruitful empirical inquiry into these problems lies, I think, in the more subtle technique of analyzing a series of single business decisions through close personal contact with those responsible for the decisions."

²⁴ According to modern theory price equals average cost (inclusive of normal profit) chiefly under the pressure of competition. The individual firm will charge a price above or below average cost depending on the situation and in line with the marginal calculus. However, when price has risen above average cost, other firms will expand production and new firms will enter the industry and their competition tends to reduce price to the average cost level. Thus it is not the price policy of the individual firm but the pressure of actual or potential competition which makes prices equal to average cost. In contrast with this, the theory advanced by the critics of marginal analysis asserts that firms set their prices according to average cost regardless of the state of competition and regardless of the market situation.

of averages must be distinguished: averages over time and averages as a function of the volume of output.

Selling prices frequently fluctuate over time, not only cyclically and seasonally but during the week or the day. In calculations for investment, cyclical price fluctuations will be taken into account and average prices will be estimated. In planning the production of seasonally demanded goods—summer dresses, swimming suits, winter sport clothes, Christmas toys—price discounts for off-season sales will be counted into the average selling price. Hotels in resorts may charge preferential rates for guests arriving on Tuesdays and leaving on Thursdays; wholesale grocers will dispose of over-ripe fruit and vegetables at reduced prices; public utilities may charge lower rates to industrial off-peak customers; in all these cases the firms will have to figure out their average revenue or average price.

Costs may show similar fluctuations over time. Raw materials and fuel prices may vary cyclically and seasonally, electric power rates even over different hours of the day. Seasonal changes of the weather may cause cost differences in several technical processes—natural instead of artificial heat for drying when wind, temperature and humidity are favorable; hydroelectric instead of steam-generated power when rivers carry sufficient water; and so on. These and hundreds of other reasons call for calculations of average costs by the affected business firms.

The average revenues and average costs which must be calculated to take care of such variations over time are not in the least inconsistent with the marginal revenue and marginal cost principles. Indeed, if increases in output are under consideration, the marginal changes of revenue and cost as functions of output will have to comprise any changes over time that will affect revenue or cost. That the firm figures with these averages over time does not mean that it makes its decisions concerning price policies on the basis of an average-cost rule rather than the maximum-profit rule.

Actual versus Potential Average Costs

The absence of the expressions “marginal cost” and “marginal revenue” from the business man’s vocabulary and the fact that he usually explains his price policy in terms of “average cost” account for a good part of the skepticism of the empiricists. Yet, the words used are not indicative of the lines of thinking; the marginal calculus may be followed without pronouncing or knowing any of the terms in question.

In the economist’s jargon, the business man who considers taking more business is supposed to say to himself: “At the increased volume of output, marginal cost will be this much and marginal revenue that much.” (Statement I.) In a literal translation into everyday language, he would

say, "The increase in production will cost me this much and will bring in that much." (Statement II.) He could say it also in a different version: "The increase in business will raise total costs from this to this much, and total receipts from that to that much." (Statement III.) These statements are absolutely equivalent, all expressing the marginal calculus of variations.

The same thing can also be expressed in a fourth, much more complicated way: "The increase in business will change average cost from this to this much, and average price from that to that much; it will, therefore, change profits by changing the margin of so and so much, times an output of this much, to a margin of so and so much, times an output of that much." (Statement IV.) With all its complications the statement is still equivalent to the former ones. It is a bit foolish to divide total costs and receipts by the output figures just in order to multiply afterwards the differences again by the output figures; but it is not incorrect. The average cost figures as such are, of course, irrelevant in the calculation.²⁵

The average cost figures, in spite of their prominent place in our business man's complicated statement, had no place in his actual decision. The decision was based on the profitability of the added business. When not only the current but also the potential average cost—that is, the average cost at a different production volume—and also the change in total receipts are considered, then the reasoning is true marginal calculus, not average-cost reasoning as some mistakenly believe.

Average-Cost Pricing as the Lawyer's Ideal

Generations of lawyers have accepted and proclaimed the fairness of the average-cost standard of pricing. Decades of regulatory experiments and arguments, and a long history of court decisions, have emphasized the average-cost principle as the just basis of pricing. Is it then surprising that business men try to explain their pricing methods by average-cost considerations?

²⁵ This can be easily illustrated by assuming any set of figures. Assume that the firm considers taking new orders for 1,000 tons of product, reducing its average price. Statement IV might read: "The increase in business from 10,000 tons to 11,000 tons will raise total cost from \$80,000 to \$86,900 and, hence, will reduce average cost from \$8.00 to \$7.90; it will raise total receipts from \$99,500 to \$107,800 and, hence, will reduce average price from \$9.95 to \$9.80; it will, therefore, raise profits by changing a margin of \$1.95, times an amount of 10,000, *i.e.*, \$19,500, to a margin of \$1.90, times an amount of 11,000, *i.e.*, \$20,900. Let's take the business."

Statement III would read under the same circumstances: "The increase in business will raise total costs from \$80,000 to \$86,900, that is by \$6,900, and will raise total receipts from \$99,500 to \$107,800, that is by \$8,300. Let's take the business."

Statement II on the same situation would read: "The increase in production will cost me \$6,900 and will bring in \$8,300. Let's take the business."

Statement I, finally, would read: "At the increased volume of output, marginal cost will be \$6.90 and marginal revenue \$8.30. Go ahead."

Corporations in regulated industries are sometimes caught in their official price justifications: a change in the market situation may make it wise and profitable to change the selling price, but that price has been anchored to an average-cost calculation which it is now difficult to disavow. The companies cannot very well submit to their regulatory commissions revised average-cost calculations every time market conditions change. They have to put up with relatively inflexible prices which, were it not for the regulatory authorities, might be as much against their own interests as against those of the consumers.

More often, however, the business man is not conscious of the fact that he uses average-cost considerations merely as rationalizations or justifications. Selling with high profit margins might indicate monopoly and "squeezing of the consumer"; selling below cost might indicate unfair competition and "cutting the throat of the competitor." As a good citizen the business man wishes to avoid both these wicked practices. As long as he can justify his prices as covering "average cost plus a fair profit margin" he can say, to others as well as to himself, that he is living up to the accepted standards of law and decency. If this "fair profit margin" is at times a bit generous and at other times rather thin, he can still justify his price. (That such variations betray his "explanation" of this pricing method as incomplete or untenable may escape his attention as well as that of his inquirers.)

Average-Cost Pricing as the Accountant's Ideal

Selling price must cover average cost inclusive of overhead and fair profit margin if the business enterprise is to live and to prosper. A good accountant regards it as his duty to watch over the soundness of the firm's pricing methods and to warn against prices below full cost.

Practical and academic accountants have sometimes attacked the marginal-cost principle as a fallacy conducive to practices that are liable to result in business losses. They have reasoned that a general application of differential cost considerations might mean that firms forget that they ought to recover their overhead in *some* part of their business.

Reasoning of this sort reveals a twofold misunderstanding of the marginal principle. (a) That marginal cost does not "include" fixed overhead charges need not mean that it will always be below average total cost; indeed, marginal cost may equal or exceed average cost. (This will always be true for volumes of output at or beyond "optimum capacity" of the firm.) (b) To use marginal cost as a pricing factor need not mean that price will be set at the marginal cost level. Indeed, this will never be done. In the exceptional case of pure competition, price cannot be "set" at all but is "given" to the firm and beyond its control; and marginal cost will be equal to price not because of any price policy but only because

of adjustments in the firm's production volume. In the normal case of monopolistic competition, the firm will never charge a price as low as marginal cost; it will charge a price at which marginal revenue is equal to marginal cost, and this price must therefore be above both.

It is a stupid misunderstanding to believe that the use of marginal cost in the business man's pricing technique implies an advice that selling price should be set at the marginal cost level. Marginal cost and marginal revenue considerations mean nothing else but what a business man means when he asks himself: "Could I get some more business and would I want it under the conditions under which I could get it"?

The idea, held by some accountants, that pricing on the basis of the marginal principle would sacrifice profits is the opposite of the truth—except in one very special sense: where the average-cost rule has been used as a monopolistic device, resort to the marginal principle might be taken to mean abandonment of a cartel arrangement in the industry and "outbreak" of unrestricted competition.

Average-Cost Pricing as a Cartel Device

In times of depression business men often discover that it is wiser to lose only a part rather than all of their overhead cost; that it is better to sell at prices below full cost than to stick to prices which would cover all costs but at which they cannot sell. They usually deplore these deviations from the full-cost principle of pricing and argue that nobody would *have* to sell below cost if nobody *did* sell below cost.

Price fixing among producers or official price codes may in such situations succeed in the maintenance of a monopolistic level of price in spite of strong temptations for competitive price cutting. Tacit understandings about the observation of average-cost rules of pricing sometimes constitute an alternative way of achieving price maintenance in a declining market. Moral suasion in the direction of "good accounting" and of "sound pricing" on the basis of "full cost" may be an effective device of domestic price cartels (through trade associations or in the form of tacit understandings).

Outright price fixing, just as any other cartel agreement, is a device to affect the estimates of demand conditions for the products of the individual firms. Only if demand as seen by the individual seller is effectively changed through his anticipations of serious reactions on the part of his competitors and fellow cartel members will he find it advantageous to restrict his output to the extent necessary for the maintenance of the agreed price. The essential effect of the agreement is upon the elasticity of the expected demand. As a rule, elasticity becomes absolutely zero (that is, the demand curve breaks off abruptly) at the largest volume of output which the individual cartel member thinks he can sell at the

fixed price. If he considers price cutting in contravention of the agreement as a practical alternative, the demand curve will not break off but continue downward with reduced elasticity—reduced because of the risk of penalizing or retaliatory actions.

The general adoption of an average cost rule is in effect a price agreement among the members of the particular industry. Where a trade association announces a representative "average cost," the announced value need not tally at all with the average cost of an individual firm. Where cost conditions are believed to be very similar throughout the industry, the understanding may be informal and tacit. It may be made entirely a matter of "business ethics" not to sell below average cost plus fair profit margin. For the firm which strictly observes this ethical code the demand curve breaks off abruptly at the output it can sell at that price. The average cost calculation of that firm takes the place of the fixed cartel price and is the essential determinant of its demand and marginal revenue considerations.

If a business man believes that the best policy for him in the long run is to stick to the cartel, this does not mean that he disregards the marginal principle. On the contrary, the feared consequences of breaking away from the cartel, its probable effects upon long-run demand and revenue, dictate his continued adherence. Likewise, if violations of the ethical code of average-cost pricing are feared to have adverse consequences, continued membership in this "ethical cartel" is not a departure from the marginal principle. The average-cost rule and the sanctions for violating it have the same sort of effects upon demand elasticity and marginal revenue which other types of price agreements have been shown to have.

Average Cost as a Clue to Demand Elasticity

Even without any ethical or unethical code prescribing an average-cost rule of pricing, average cost may be the most important datum for the estimate of demand elasticity. The elasticity of demand for any particular product is determined by the availability of substitutes. In order to estimate how much business a firm may lose if it raises its price, it will consider whether existing or potential competitors can supply competing products at the particular price. The elasticity of supply from competing sources determines the elasticity of demand for the firm's product. The supply from competing sources will depend on their actual or potential cost of production. And usually the best clue that a firm has to the production cost of competitors is its own production cost, corrected for any known differences of conditions.

Assuming that competitors have the same access to production factors, materials and technology, their production cost can not be much

different from that of a particular producer who may just be weighing the chances of a price increase. In the absence of any cartel arrangements he will have to count on his competitors to expand their business at his expense if he ventures to raise his selling price above average cost. Where he need not fear the capacity of existing competitors, but entry into the industry is relatively easy, he will have to reckon with newcomers' competition if he makes the business too attractive by allowing himself too generous a profit margin above average cost. Under such circumstances he will know that he stands to lose too much business and had better stick fairly closely to a price based on average cost.

Notwithstanding any rationalizations of this price policy, the reasons for it lie in the competitiveness of the industry resulting in a high elasticity of demand visualized by individual sellers.²⁶ To "explain" this price by reference to some emotional attachment to the average cost principle is to miss the mark. The rôle of average cost in the firm's pricing process in this case is to aid in gauging the elasticity of the long-run demand for its product.

Reasons and Variables

Seeing how many different rôles average cost may play in the pricing process without in the least contradicting the statement that marginal cost and marginal revenue determine output and price, one should realize the dangers of attempts to use utterances of business men as evidence against the correctness of marginal analysis.

Business men's answers to direct questions about the reasons for charging the prices they are charging are almost certainly worthless. Every single fact or act has probably hundreds of "reasons"; the selection of a few of them for presentation to the inquirer is influenced by the prejudices or old theories which the informant had impressed upon him by school, radio, newspapers, etc.

Except in the case of a genuine decision leading to a recent change of policy, one may say that an approach much more fruitful than that of asking about reasons *for* some policy is to ask about reasons *against* its alternatives. Instead of asking for explanations of the price actually charged or the output volume actually produced, questions about "why not more" and "why not less" are likely to yield more revealing results. But even these answers must be checked and double-checked through a network of cross-examination, segregating and isolating certain variables in a manner familiar to the scientist working with the calculus of variations and with the determination of partial derivatives.

²⁶ Where the average-cost rule is a cartel device, the elasticity of demand will be small or zero from the actually realized point on *downward*. When average cost is a clue to size up potential competition, the elasticity of demand will be high from the actually realized point on *upward*. The former prevents price reductions, the latter price increases.

Research on Actual Pricing Methods

On the basis of marginal analysis of the firm and the industry, we should expect for most industries that price in the long run would not deviate too much from average cost, yet that the firm would attempt to get better prices when it could safely get them and would not refrain from cutting prices when it believed that this would increase its profits or reduce its losses.

Now let us compare with this the findings of one of the empirical research undertakings which shook the researchers' confidence in the marginal principle and convinced them that business men followed the "full-cost principle" of pricing regardless of profit maximization. Inquiry was made through interview of 38 entrepreneurs.²⁷ "A large majority" of them explained that they charged the "full cost" price. Some, however, admitted "that they might charge more in periods of exceptionally high demand"; and a greater number reported "that they might charge less in periods of exceptionally depressed demand."²⁸ Competition seemed to induce "firms to modify the margin for profits which could be added to direct costs and overheads."²⁹ Moreover, "the conventional addition for profit varies from firm to firm and even within firms for different products."³⁰

This is precisely what one should have expected to hear. Do these findings support the theory of the average-cost principle of pricing? I submit that they give little or no support to it. The margins above average cost are different from firm to firm and, within firms, from period to period and from product to product. These differences and variations strongly suggest that the firms consult other data besides or instead of their average costs. And, as a matter of fact, the reported findings include some that indicate what other considerations were pertinent to the price determinations by the questioned business men.

Of 24 firms which gave reasons for not charging higher prices, 17 were tabulated as admitting that it was "fear of competitors or potential competitors" and a "belief that others would not follow an increase." Another two stated that "they prefer a large turnover."³¹ To me the 19 answers indicate that these business men were estimating the risk of losing business if they raised prices or, in other words, that they were concerned about the elasticity of demand.

Of 35 firms which gave reasons for not charging lower prices, 4 firms explained that they were members of price-fixing combinations; 2 stated

²⁷ R. L. Hall and C. J. Hitch, *op. cit.*, p. 12.

²⁸ *Ibid.*, p. 19.

²⁹ *Loc. cit.*

³⁰ R. L. Hall and C. J. Hitch, *op. cit.*, p. 20.

³¹ *Ibid.*, p. 21.

that it was "difficult to raise prices once lowered"; and 21 referred directly or implicitly to their estimates of demand elasticity. (Nine firms: "Demand unresponsive to price"; one firm: "Price cuts not passed on by retailers"; eleven firms: "Competitors would follow cuts.") Only 8 firms gave reasons other than monopolistic price fixing or monopolistic elasticity considerations; these 8 were listed as having "quasi-moral objections to selling below cost."³² Unfortunately the interviewers did not find out what these conscientious objectors to price cutting thought about the responsiveness of demand; and whether they would remain adamant if they were sure that a small price concession would produce a large increase in sales. I suspect that a cross-examination would have brought out the fact that the moral or quasi-moral views on price maintenance were regularly coupled with a very strong opinion that a price reduction would not produce sufficiently more business and, thus, would constitute useless sacrifice of profits.

In any event, there is little or nothing in the findings of this inquiry that would indicate that the business men observed an average-cost rule of pricing when such observance was inconsistent with the maximum-profit principle. On the other hand, there is plenty of evidence in the findings that the business men paid much attention to demand elasticities — which to the economist is equivalent to marginal revenue considerations.

The Absence of Numerically Expressed Estimates

Why should others in the face of this evidence have come to the conclusion that the marginal principle was not applied and profit maximization not attempted by the group of business men studied? How could others have failed to be impressed by the facts just recited?

It seems that their confidence in the conventional analysis was lost when they found to their surprise that the business man had no definite numerical estimates of the magnitudes relevant to the application of the marginal principle. They had assumed that a business man should "know" the elasticity of demand for his product, and now they were shocked to find "that the great majority of entrepreneurs were in profound ignorance with regard to its value."³³ A student who had expected to find exact estimates must indeed have been disappointed when most of his informants "were vague about anything so precise as elasticity."³⁴

The inquirers found the same vagueness with regard to marginal cost estimates. While the entrepreneurs usually computed direct cost and

³² *Loc. cit.*

³³ R. F. Harrod, *op. cit.*, p. 4. Concerning this discovery Mr. Harrod remarks emphatically: "This, indeed, must be regarded as one notable result of our inquiry."

³⁴ R. L. Hall and C. J. Hitch, *op. cit.*, p. 18.

total overheads "with some pains at accuracy,"³⁵ they could not furnish any data on marginal cost. He who expected that marginal cost and marginal revenue were equated on the basis of precise calculations must feel stultified. The student who had to do homework computing marginal cost and revenue figures to the second or third decimal point may feel befooled when he learns that the business man does not do anything of the sort. But to conclude from the absence of definite numerical estimates that the magnitudes in question were irrelevant in the conduct of the firms is a *non sequitur*. On the basis of the previous discussion of this subject (see above pp. 534 ff.) we should understand that the construction of a pattern for the analytical description of a process is not the same thing as the actual process in everyday life; and we should not expect to find in everyday life the definite numerical estimates that are part of the scientific pattern.

Apart from the absence of numerical estimates of marginal revenue and marginal cost it is difficult to see what other findings of the inquiry could have persuaded the researchers that they had disproved the theory of marginalism in the conduct of the firm. There is not a single proposition in the tabulated results of the inquiry that can not be fully harmonized with marginal analysis. The "Analysis of Replies to Questionnaire on Costs and Prices," which the researchers presented as an appendix to their report,³⁶ contains a wealth of illustrative material—illustrative, as I see it, of the application of the marginal principle to business decisions of the single firm.

B. *Marginal Productivity and Wage*

Empirical research designed to verify or disprove marginal productivity theory in the analysis of input of the individual firm is beset with difficulties. Few systematic endeavors have been made and none has led to any suggestion, however vague or tentative, of an alternative theory. Whereas in certain price research projects those who felt compelled to reject the marginal theory have advanced the average-cost theory of pricing as a substitute, no substitute theory has been forthcoming from those who decried marginal productivity theory.

Statistical Research

Empirical research on cost, price and output of the individual firm has resulted in several interesting attempts to derive marginal cost functions from statistical data; and also in one or two attempts to derive price elasticities of demand for a firm's products. But nobody, to my knowledge, has ever undertaken to construct from actual data a

³⁵ R. F. Harrod, *op. cit.*, p. 4.

³⁶ R. L. Hall and C. J. Hitch, *op. cit.*, pp. 33-45.

marginal net revenue productivity curve for a given type of labor employed by a firm. The difficulties are formidable and, since the raw material for the calculations could not come from any records or documents but merely from respondent's guesses of a purely hypothetical nature, the results might not be much more "authentic" than the schedules made up by textbook writers for arithmetical illustrations.

Statistical studies of the relationship between wage rates and employment in large samples of individual firms or industries would be nearly useless because we have no way of eliminating the simultaneous effects of several other significant variables, especially those of a psychological nature. An increase in wage rates may have very different effects depending on whether the employer (1) (a) has foreseen it, (b) is surprised by it; (2) (a) reacts quickly to it, (b) reacts slowly to it; (3) (a) expects it to be reversed soon, (b) expects it to be maintained, (c) expects it to be followed by further increases; (4) (a) assumes it to be confined to his firm, (b) assumes it to affect also his competitors, (c) believes it to be part of a nation-wide trend; (5) connects it with an inflationary development; or is influenced by any other sort and number of anticipations. Most of these moods and anticipations can be translated by the economist into certain shapes or shifts of the marginal productivity functions of the firms; but since the researchers cannot ascertain or evaluate these conjectural "data" for the large number of firms contained in a representative sample, statistical investigations of the wage-employment relation of individual firms are not likely to yield useful results.

Questionnaire on Employment

It has been pointed out above (p. 538) why the method of mailed questionnaires without supporting interviews is hopelessly inadequate for empirical studies of business conduct. Even the most intelligently devised set of questions would not assure reliable and significant answers. Questions designed to achieve the necessary separation of variables would be so complicated and call for so high a degree of "abstract thinking" on the part of the questioned business men that questionnaires of this sort would be too much of an imposition, and coöperation would be too small. Although the questions in Professor Lester's research project on employment did not even approach these standards, he received only 56 usable replies from 430 manufacturers whom he had asked to fill out his questionnaires.³⁷

Professor Lester's questionnaires suffered not merely from the inherent weaknesses of the method but also from defects in formulation. These defects were so serious that even the most complete, reliable and intelligent

³⁷ R. A. Lester, *Am. Econ. Rev.*, Vol. XXXVI, No. 1, pp. 64-65.

answers could not have yielded significant findings. The business men were asked to rate the "importance" of several factors determining the volume of employment in their firms. No explanation was given whether this importance of a variable—that is, I presume, its responsibility for changes in the employment volume—should refer to (a) the frequency of its variations, (b) the extent of its variations, or (c) the effects of its variations. Surely, the variable rated as least important—perhaps because it varied less frequently than the others—may be just as strategic as any of those with higher importance ratings. What we really need to know, however, is not the *comparative* importance of several factors but rather the effects of variations of each factor separately while the others remain unchanged.

If I want to know by how much an increase in the price of spinach may affect its consumption in an individual household, I shall not get very far by asking the householders to give a percentage rating to each of several listed factors that are believed to be "important" influences on spinach consumption. If it were tried, we should not be surprised to find changes in family income, the number of children and guests at dinner, and the notoriety of Popeye the Sailor's gusto for spinach, receiving much higher percentage ratings than changes in the price of spinach. (In a number of households price may not be a factor at all.) Nobody, I hope, would conclude from such a poll that price is an unimportant factor in the consumption of spinach.

Yet Professor Lester followed just this procedure when he wanted to find out how important wage rates were in determining the volume of employment in the individual firm. He asked the executives of the companies to "rate" the following factors "in terms of the percentage of importance of each":

- a. Present and prospective market demand (sales for your products, including seasonal fluctuations in demand).
- b. The level of wage rates or changes in the level of wages.
- c. The level of material costs and other non-wage costs and changes in the level of such non-labor costs.
- d. Variations in profits or losses of the firm.
- e. New techniques, equipment, and production methods.
- f. Other factors (please specify).

Of these items the first unquestionably excels all others in frequency and extent of variations. That it won first prize in Professor Lester's importance contest is therefore not surprising. If several respondents gave ratings to item d (variations in profits or losses) and at the same time also to other items, they obviously did not realize that this variable comprised all the others. Professor Lester does not explain why he listed it when he knew that it was not "completely independent" and

that "for example, wages affect profits."³⁸ Nor does he state whether the 43 firms which failed to mention changes in wage rates as an important factor meant that they would continue in business and continue to employ the same number of workers regardless of any degree of wage increase. If this is what they meant, they can hardly be taken seriously. If they meant something else, then it is not clear just what the replies should indicate about the probable effects of wage increases upon employment.

The strangest thing about Professor Lester's list of possibly important variables is that all—except f, the unspecified, and d, the all-inclusive profit-and-loss item—are essential variables of the very analysis which he means to disprove. The prize-winning item, a, the demand for the product, is certainly a most crucial determinant of marginal productivity. (See above pp. 529 and 531.) Items c, non-labor cost, and e, production techniques, are two other determinants of marginal productivity. How Professor Lester came to think that the results of this poll would in any sense disprove or shake marginal productivity analysis remains a mystery.

Questionnaire on Variable Cost

Professor Lester asked his business men also some questions on unit variable costs and profits at various rates of output. The information obtained in answer to these questions might have been useful had it not been based on an undefined concept of "plant capacity." Unfortunately, it must be suspected that not all firms meant the same thing when they referred to "100 per cent of capacity."

Economic theorists use different definitions of capacity. One widely-used definition marks as 100 per cent of capacity that volume of output at which short-run total cost per unit is a minimum; another definition fixes the 100 per cent mark at the output at which variable cost per unit is a minimum. The former definition implies decreasing average total cost, the latter decreasing average variable cost, up to "100 per cent capacity." Professor Lester after painstaking empirical research arrives at the following finding:

The significant conclusion from the data in this section is that most of the manufacturing firms in the industries covered by this survey apparently have decreasing unit variable costs within the range of 70 to 100 per cent of capacity production. . . .³⁹

Has Professor Lester asked himself whether this is not merely a self-evident conclusion implied in the definition of capacity used by his respondents?

³⁸ *Ibid.*, p. 66.

³⁹ *Ibid.*, p. 71.

The steepness of the reported decline in unit variable cost, however, would be an interesting observation—if the data were reliable. (Few of Professor Lester's firms had "constant unit variable costs," or anything approaching this situation, over a considerable range of output.⁴⁰) It is rather peculiar that unit variable costs should decrease steeply (at an increasing rate!) down to a certain point and then abruptly start rising—as one must infer from the term "100 per cent capacity." Where equipment is not utilized for 24 hours a day, the steep decline and abrupt rise of the unit cost is somewhat questionable.

Professor Lester, nevertheless, has sufficient confidence in his findings to draw conclusions—conclusions, moreover, which could not even be supported if the findings were of unquestionable validity. He states:

If company output and employment policies are based on the assumption of decreasing marginal variable cost up to full capacity operations, much of the economic reasoning on company employment adjustments to increases and decreases in wage rates is invalid, and a new theory of wage-employment relationships for the individual firm must be developed.⁴¹

This deduction simply does not follow from the premises. There is no reason why decreasing marginal costs should invalidate the conventional propositions on factor cost and input. Professor Lester could have found dozens of textbook examples demonstrating the firm's reactions under conditions of decreasing marginal cost.

Professor Lester may have been deluded by a rather common confusion between related concepts: from decreasing marginal cost he may have jumped to the assumption of increasing labor returns,⁴² and from increasing physical returns he may have jumped to the assumption of increasing marginal productivity of labor. Both these jumps are serious mistakes. For instance, the very conditions which may cause a firm to restrict the employment of labor to a volume still within the phase of increasing physical productivity per unit of labor are likely to result in decreasing marginal net revenue productivity of labor. These conditions are:

- (a) an indivisibility of the firm's physical plant facilities,⁴³ combined with either (or both),
- (b) a low elasticity of the demand for the firm's products⁴⁴ or (and)

⁴⁰ *Ibid.*, p. 70.

⁴¹ *Ibid.*, p. 71.

⁴² *Ibid.*, p. 68.

⁴³ *I.e.*, the firm cannot adjust the number of machines or production units to smaller production volumes but must instead produce small outputs with an inefficiently large productive apparatus.

⁴⁴ *I.e.*, the firm realizes that it can charge much higher prices for smaller outputs or cannot dispose of larger outputs except with substantial price reductions.

(c) a low elasticity of the supply of labor to the firm.⁴⁵

The first condition, (a), makes a phase of increasing physical productivity of labor in the firm a practical possibility; the other conditions, (b) or (c), make that phase relevant for actual operations by providing the pecuniary incentive to operate the plant inefficiently. Condition (b), the low elasticity of demand for the product, will cause marginal net revenue productivity of labor to be diminishing in a range of employment in which average or even marginal physical productivity of labor are still increasing.

It is not possible from Professor Lester's exposition to find out whether his failure to see these relationships was at the bottom of his faulty theorizing on this point. In any event, his findings on variable costs contain nothing that would even vaguely bear on the validity of marginal analysis.

Questionnaire on Adjustments

Professor Lester's fact-finding and theorizing on substitution between labor and capital and on other adjustments of the firm to changes in wage rates are also marred by inconsistencies and misunderstandings.

After trying to make the most of increasing returns to labor and only a few lines after referring to "unused plant capacity," Professor Lester asserts that "most industrial plants are designed and equipped for a certain output, requiring a certain work force. Often effective operation of the plant involves a work force of a given size."⁴⁶ To operate within the phase of increasing returns is to operate inefficiently, that is, with an employment of less labor with a given plant than would be compatible with efficient operations. (Because an increase in employment would raise output more than proportionately.) "Effective operation," on the other hand, logically implies employment at or beyond the point where diminishing returns set in. Professor Lester does not seem to be clear which way he wants to argue.⁴⁷

Professor Lester seems to think that substitution between capital and labor can occur only in the form of installation of new or scrapping of existing machinery⁴⁸ and that it is supposed to occur "readily" and would, therefore, be "timed" with the wage changes.⁴⁹ These are rather common but nevertheless mistaken views.

⁴⁵ *I.e.*, the firm realizes that it can enjoy much lower wage rates at lower employment levels or cannot obtain more labor except with substantial wage increases.

⁴⁶ *Amer. Econ. Rev.*, Vol. XXXVI, No. 1, p. 72.

⁴⁷ Absolutely fixed proportions between factors of production would imply that short-run marginal productivity of labor drops precipitously to zero at the full capacity level of employment.

⁴⁸ *Am. Econ. Rev.*, Vol. XXXVI, No. 1, p. 73. See my comments, above, pp. 530-31.

⁴⁹ *Ibid.*, pp. 73 and 74.

Professor Lester does not discuss a glaring contradiction in his findings: On the basis of replies to one questionnaire he states that his data indicate "that industry does not adapt its plant and processes to varying wage rates in the manner assumed by marginalists."⁵⁰ But on the basis of another questionnaire about adjustments to increases in relative wages, he reports that the introduction of "labor-saving machinery" was given the highest rating in relative importance by the questioned firms whose labor costs were more than 29 per cent of total cost.⁵¹

The last-mentioned questionnaire apparently was designed to show that wage increases had no important effects upon employment. Six alternative adjustments to increases in relative wages were listed and manufacturers had to give percentage ratings for relative importance. In this popularity contest an item called "deliberate curtailment of output" got the bobby prize. Quite apart from the fact that the words were loaded against this item, the result is not in the least surprising. For it is a well-known fact that where competition is not pure (as it rarely is in industrial products), output adjustments to higher production costs take place by way of changes in selling price. Price and product adjustments were another of the alternative items and scored rather well in the poll. If all employment-reducing adjustments—labor-saving machinery, price increases, and deliberate output curtailment—are taken together, they clearly dominate in the importance ratings by the firms.⁵² This, or anything else, may not mean much in such an "opinion poll," but it certainly does not prove what Professor Lester wanted to prove. Nevertheless, he contends that "it is especially noteworthy that deliberate curtailment of output, an adjustment stressed by conventional marginal theory, is mentioned by only four of the 43 firms."⁵³ And he concludes that marginal analysis is all but done for, that "there can be little doubt about the correctness of the general results" of his tests⁵⁴, and that "a new direction for investigations of employment relationships and equilibrating adjustments in individual firms" is indicated.⁵⁵

C. Conclusions

I conclude that the marginal theory of business conduct of the firm has not been shaken, discredited or disproved by the empirical tests discussed in this paper. I conclude, furthermore, that empirical research on business policies cannot assure useful results if it employs the method

⁵⁰ *Ibid.*, p. 73.

⁵¹ *Ibid.*, p. 78.

⁵² *Ibid.*, p. 78.

⁵³ *Ibid.*, p. 79.

⁵⁴ *Ibid.*, p. 81.

⁵⁵ *Ibid.*, p. 82.

of mailed questionnaires, if it is confined to direct questions without carefully devised checks, and if it aims at testing too broad and formal principles rather than more narrowly defined hypotheses.

The critical tone of my comments on the research projects discussed in this paper may give the impression of a hostile attitude towards empirical research as such. I wish to guard against such an impression. There should be no doubt that empirical research on the economics of the single firm is badly needed, no less than in many other fields. The correctness, applicability and relevance of economic theory constantly need testing through empirical research; such research may yield results of great significance.

Sharp criticism of bad research can be constructive in two respects: it may save some of the waste of time which the published research findings are apt to cause if they remain undisputed and are allowed to confuse hosts of students of economics; and it may contribute to the improvement of research. The chief condition for improved research is a thorough understanding of the theories to be tested. Supplementary conditions are a certain degree of familiarity with the technological and institutional peculiarities of the fields or cases on which the research is undertaken and a grasp of the research techniques employed.