Product Differentiation and Welfare

Michael Spence

The American Economic Review, Vol. 66, No. 2, Papers and Proceedings of the Eighty-eighth Annual Meeting of the American Economic Association (May, 1976), 407-414.

Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28197605%2966%3A2%3C407%3APDAW%3E2.0.CO%3B2-T

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

Product Differentiation and Welfare

By MICHAEL SPENCE*

One of the functions of the market system is to select the commodities that are produced and sold. This process is variously referred to as product differentiation, product selection, and monopolistic competition. Product differentiation involves a set of real economic choices because there are increasing returns or declining average costs in the development, production, marketing and distribution activities of firms. The full range of possible products is neither feasible nor desirable in the presence of increasing returns to scale.[1] Product differentiation is also an important component of imperfectly competitive strategic interaction, both static and dynamic. Some would argue it is, in many industries, the most important part of the dynamics of competition.

My main purpose here is to discuss some of what has recently been learned about the welfare aspects of product differentiation and monopolistic competition in a market system. In order to do this, it is necessary to measure welfare, and I shall adopt the strategy of measuring welfare with the multiproduct net surplus, the difference between the benefits to consumers, gross of revenues paid, minus the costs of production.[2] A product's marginal contribution to surplus is the additional gross surplus it adds, minus the cost of producing it. That in turn is the area under the inverse demand function minus the costs of production. The total surplus, of course, is not the sum of these marginal surpluses; typically the total net surplus will be greater for products which are substitutes.

Everyone who is sensitive to the intricacies of market structure is aware that product differentiation has interesting information aspects, and that there is a collection of activities associated with marketing that bear on the demand for and sales of a product. I do not deal with these here, though they are important. But so are the limits of the price system in selecting products and product characteristics. These are the subject of this paper.

If firms are not perfect price discriminators, then the revenues of a firm selling a differentiated product are not equal to that product's contribution to gross surplus. Similarly profits are not equal to its net contribution to surplus. Revenues and profits are less than the gross and net contributions to surplus respectively. But profitability is the criterion by which products are selected or rejected in a market system. In the light of the divergence between benefits and profits, it seems reasonable to suppose that the criterion of profitability may not always lead to de-

* I am grateful to Richard Caves, Richard Nelson, and Oliver Williamson for their comments. Support for this research from the National Science Foundation and the John Innon Guggenheim Foundation is gratefully acknowledged.

[1] One can think of products being points in a continuous spectrum of attributes. The market selects a finite subset of points in the continuum, these being the products actually produced. For recent contributions to this subject, the reader can refer to Robert L. Bishop, Edward Chamberlin, Kelvin Lancaster, Bruce Owen and Spence (1975), and Spence (1975).

[2] For a discussion of the accuracy of this measure, see Robert D. Willig (1973).

sirable results. Indeed, that is the case.

This paper divides into three sections. In the first, I have attempted to summarize the forces that generate welfare problems in product selection, and the nature of the problems generated. There are several of these forces, and some of them work against each other. It is nevertheless useful to identify them and to understand the direction in which they push market outcomes. The second section deals with the magnitudes of welfare losses. In the past, some of the costs of imperfect competition have been measured by the cost of the nonmarginal cost pricing of the existing set of products. Using some numerical examples, I have tried to show that a significant fraction of the cost of imperfect competition may be due to the currently unmeasured cost of having too many, too few, or the wrong products. This analysis is far from decisive empirically, since it is based on numerical examples. It is however suggestive, that over a range of assumptions about the structure of demand and costs, product selection failures may be significant components of welfare costs.

In the third section, I have tried to outline, somewhat briefly, how the empirical study of product differentiation in an industry might proceed. The problem here is that the welfare analysis requires the estimation of the demands for products that currently do not exist. This line will be pursued more fully in subsequent papers.

The reader may find it useful to read this paper with the concrete case of television programming in mind. Pretend that pay television on the cable makes per program charges feasible.

I. Welfare Problems in Product Selection

There are several market forces operating in the selection of products.

A. Profits and Net Surplus

In the absence of price discrimination, the profits of the seller of a product will fall short of the net surplus generated by the product. Therefore, a product can have a positive potential net surplus and at the same time, be incapable of generating positive earnings. The exception to this principle is the case of constant or diminishing returns to scale in production. The fact that profits do not capture all of the net surplus is a conservative force, tending to eliminate products that should be made available. I hasten to add that this is one of several forces. It does not warrant the conclusion that there are too few products, or too little variety.

The tendency to lose desirable products is greater, the more substantial are the increasing returns or fixed costs.

B. Falling Average Costs

In a monopolistically competitive market, without entry barriers beyond those imposed by the usual profitability requirement, prices are set above marginal costs, and entry occurs until profits are driven to zero, or equivalently until prices are equal to average costs. These statements technically apply to products that are symmetrically different in the way that it has become customary to assume in the theory. Since price is at average cost and above marginal cost, marginal cost is below average cost and average costs are declining at that point. It does not follow from this fact that there are too many products. It is true that there are more products than there would be if nonnegative profits were required and price were equal to marginal cost. It is also true that prices equal marginal costs in the first best optimum. But it is not true that the first best or second best outcome is achieved when profits go to zero. Profits can be above or below zero when the

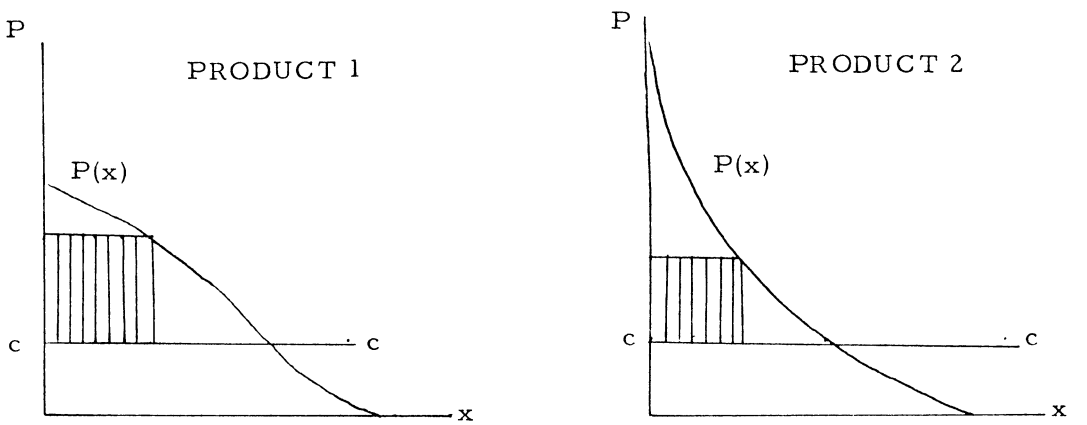

FIGURE 1.

marginal contribution of a product to the surplus falls to zero.[3]

This point is most easily seen for the case in which average costs decline indefinitely. In that case, if price equals marginal costs, profits are always negative. But it is not true in general that the optimal number of products is zero.

C. Revenues, Gross Surplus, and Biases Against Products

Products are not equally subject to the conservative force discussed under (1) above. There are biases in product selection. Although revenues, in the absence of price discrimination, do not capture all of the gross surplus generated by a product, the fraction of the gross surplus that is or can be captured varies from product to product. The ratio of revenues to gross surplus depends upon the properties of the demand functions. Rather than analyze this in the abstract, let me illustrate the point with the case of constant elasticity demands. For the time being, I suppress explicit consideration of interproduct interactions. Let the price elasticity of demand for a product be η. It is easy to establish that at any price, the ratio of revenues to gross surplus (the latter being the area under the inverse demand function) is

$$\beta = 1 - 1/\eta. \tag{1}$$

Therefore, as η rises, the fraction β rises. The fraction of gross benefits that appear as revenues to the selling firm is larger, the higher the price elasticity of demand.

Since products vary in elasticities, it is quite possible for a product with a low price elasticity to have a higher net surplus and lower profits than a product with a high price elasticity of demand. This has two implications. The tendency to lose valuable products, discussed earlier, is stronger for low elasticity products. And second, the process of selection of products in a differentiated product industry will be biased against the lower elasticity candidates.[4]

The nature of the bias is illustrated in Figure 1. The first product has a lower

[3] The content of the so-called excess capacity theorem is therefore limited to the statement that at the equilibrium, average costs are still falling.

[4] This conclusion appears to run counter to the intuition developed in countless students of microeconomics, that low elasticities are nice for sellers because they permit higher markups. I don't wish to try to overturn this truth, but rather to point out nevertheless, that the surplus is harder to capture with the price system for low elasticity products. Let me put this more positively. I would expect the incentive for sellers to try to price discriminate would be higher for lower elasticity products.

surplus and higher profits than the second. The profits of both products in Figure 1 are positive. However, if there were a fixed component of costs, then the first product could be profitable and survive, while the second would not, even though the second one contributes more to the surplus. If the two products are substitutes, and if the market mechanism selects the more profitable, then again the first product will be taken.

If one does this analysis more rigorously, it turns out that it is *not* exactly the elasticity that matters, but rather what fraction of net potential surplus for a product is capturable by the selling firm. This depends upon both the structure of demand and of costs. But the qualitative nature of the bias is clear. If the potential consumers of a product have a highly variegated set of willingnesses to pay for it, so that there is a small group with a high willingness to pay, and then rapidly declining reservation prices after that, then the selling firm will have difficulty capturing the surplus, and the product will have difficulty being supplied under the market system. With some caution, one might refer to such products as special interest ones. They tend to be supplied by clubs or other institutional devices that do not use the pure price system.

D. Interactions Among Products

Thus far, I have discussed forces that affect a single product. Let me turn now to interactions among products that are imperfect substitutes. Apart from biases in selection, interactions on the demand side determine whether there are too many or too few products, too much or too little variety. In what follows, I shall assume that products are substitutes, that they have fixed and variable costs, the latter being linear or convex, and that market equilibria are Nash equilibria in quantities. From a social point of view, the fixed component of cost can be regarded as an entry fee for having an additional product.

Perhaps the best way of approaching product interactions is to describe the circumstances that lead to too many products. They are two. The first is high own price elasticities, for these cause firms to contract output further away from the level at which price equals marginal cost. And that leaves "room" for further entry, room which would be absent with marginal cost pricing. But high own price elasticities are insufficient. The cross elasticities must also be high.

The entry of an additional product has several effects. It increases the surplus from the new product, but lowers the demand for existing products and causes them to contract output. In terms of the surplus, there are gains and losses. The gains are the profits and the consumer surplus from the new product. The losses are reductions in the profit and surplus from the existing products. When the products are close substitutes and the cross elasticities are high, the extra surplus created by the entering product is lost through contractions of existing firms. The familiar metaphor of expanding and slicing the pie is appropriate. A new product expands the pie and causes it to be sliced into more pieces. If the cross elasticities are high, the expansion is not large. But since there is a cost (the fixed cost) of adding a slice, the costs may outweigh the benefits. These forces will be illustrated in the calculations of the next section.

E. Half a Loaf is Better than None

While profitability is not the correct criterion for deciding on the entry of a product, it is, for practical purposes, the only criterion we have. Nor is it likely that the information required to provide

the correct subsidies will ever be available. One can reasonably accept profitability as a constraint, and pose the problem of product selection as that of determining the right set of products subject to that constraint. The solution to the problem includes specification of not only the products, but also the prices. The prices will typically be above marginal cost, since that may be required to increase the profitability of products to permit the entry of products that are not profitable under marginal cost pricing. In short, the solution to the second best problem will include a trade-off between numbers of products on the one hand and the inefficiency due to nonmarginal cost pricing on the other.

Roughly, that is what occurs in a monopolistically competitive market. Therefore, the monopolistically competitive equilibrium has the qualitative features of the solution to the problem of maximizing the surplus subject to the condition that all products are profitable. It can happen that the equilibrium and the second best optimum coincide.[5] But in general, this coincidence is unlikely, and the best that can be said is that the equilibrium approximates the constrained optimum. The calculations that follow will give some insight into the question of how good the approximation is.

F. Complements

I want to comment briefly on complementary products. Complementary products tend to be undersupplied by imperfectly competitive industries. That is to say, the prices are too high and the number of products is too small. The reason is that the tendency to hold back quantities reduces the demand for other products and makes entry more difficult, not easier, as in the case of substitutes. This means that all forces work in the same direction. The failure of firms to price at marginal cost reduces options for other firms, and entry is retarded, as well as expansion by existing firms. One would therefore expect complementary products to be supplied by multiproduct firms.

[5] See Avinash Dixit and Joseph Stiglitz (1974) and Spence (1976).

II. Numbers and Variety of Products: Numerical Examples

In this section, I have taken some quite simple numerical examples, in which demand functions are linear and products are symmetrically different, and calculated the market equilibrium and two different optima, for a variety of the values of the parameters. The purposes of this are twofold. First, the interaction effects discussed above can be seen to work in the examples. Second, I want to show that the fraction of the total welfare loss that is attributable to the nonmarginal cost pricing of the equilibrium products, varies considerably, and is frequently less than half of the total welfare loss. The calculations also show that the equilibrium is often a reasonably good approximation to the constrained optimum, where the constraint is that profits be nonnegative.

Briefly, the basis for the calculations is the following. The quantity of the ith product is x_i. The inverse demand for the ith product is

$$p_i = a - 2bx_i - 2d \sum_{j \neq i} x_j. \tag{2}$$

The cost function for the ith firm is

$$F + cx_i. \tag{3}$$

In the calculations, c is taken to be one. The remaining parameters, a, b, d, and F vary. F is the fixed cost, d is the interaction effect; a and b are the intercept and slope of the inverse demand for each product, when there are no other prod-

TABLE 1

Group	Case	a	b	d	f	T_1	ΔT_2	ΔT_3	N_1	N_2	X_1	X_2	X_3
I	1	10	1	.5	1	28.8	2.8	1.9	5.4	9.7	1.4	.71	.84
	2				2	24.5	3.5	1.8	3.5	6	2	1	1.29
	3				4	19	4.3	1.2	2.2	3.36	2.82	1.41	2.06
	4				6	15.3	4.8	.6	1.59	2.19	3.46	1.73	2.81
	5				8	12.5	5	.2	1.25	1.5	4	2	3.6
	6				10	10.3	5.1	0	1.01	1.02	4.47	2.24	4.4
II	7	10	2	1.5	2	8.2	2.6	1.4	1.17	2.58	2	.71	1.03
	8				4	6.3	3	.87	.73	1.3	2.8	1	1.8
	9				6	5.1	3.2	.33	.53	.78	3.46	1.22	2.7
III	10	10	1	.1	2	99.7	8.7	1.7	21	26	1.49	1	1.3
	11				5	56.6	13	.1	10	9.4	2.36	1.58	2.4
	12				7	39.6	14.8	2.2	7.1	5.1	2.79	1.87	3.2
	13				9	27.3	16.3	7.3	5.2	2.2	3.16	2.5	4.0
	14				10	22.5	16.8	11.3	4.5	1.1	3.3	2.12	4.4
IV	15	10	.7	.5	2	29.9	4.8	3.5	2.4	5.7	3.2	1.2	1.5
	16				4	26	6.2	3.8	1.6	3.5	4.5	1.7	2.3
	17				6	23.2	7.1	3.5	1.2	2.5	5.4	2.0	3.1
	18				8	20.9	7.7	3.0	1	2	6.3	2.4	3.8
V	19	10	.3	.2	2	82.1	7.6	6.2	4.5	10.3	4.5	1.8	2.1
	20				6	69	12	7	2.4	5.1	7.7	3.2	4
	21				10	61	14	7.5	1.75	3.5	10	4.1	5.6
	22				15	53.6	16.1	6.8	1.3	2.5	12.2	5	7.5
VI	23	10	.3	.05	10	170	23	2	9.2	11	6.3	4.1	5.6
	24				20	102	29	0	5.1	4.6	8.9	5.8	9.4
	25				30	62	32.9	10	3.2	1.72	11	7.1	13.3

ucts.[6] Note that because of the symmetry of the example, an equilibrium is described by x, the output per firm and n, the number of firms.

In Table 1, T is the surplus, N is the number of firms, and X is the output per firm. The table shows the equilibrium, the optimum, and a constrained optimum. The constrained optimum takes the equilibrium number of firms as given and prices the goods at marginal costs. The difference between the surplus here and the optimal surplus is the loss due to the nonmarginal cost pricing of existing products. The remainder of the welfare loss is due to product selection; in this context, having the wrong number of products.

A word about the table is in order. There are six groups of calculations, distinguished by own and cross elasticities. Within each group, fixed costs increase. Thus in group I, the own slope is -1, the cross partial -.5, and the fixed costs increase from 1 to 10. The numbers reported for each case are as follows. The subscript refers to the following cases:

1 — the optimum
2 — the market equilibrium
3 — the equilibrium number of firms with marginal cost pricing.

[6] The total surplus here is

$$T = \sum_{i=1}^{n} \int_0^{x_i} p_i(x, \ldots, x_{i-1}, s_i, 0, \ldots, 0)ds_i$$

$$= \sum_{i=1}^{n} \left(a - 2d \sum_{j=1}^{i-1} x_j \right) x_i - bx_i^2.$$

When $x_i = x$ for all i, the surplus is

$$T = n[ax - bx^2 - d(n-1)x^2].$$

The variables are

X = output per firm
N = the number of firms
T = surplus
ΔT = the difference between the surplus for that case and the surplus at the optimum.

Notice that $N_2 = N_3$ so only N_2 is reported. ΔT_2 is the welfare loss in the market equilibrium. ΔT_3 is the welfare loss if the equilibrium goods were priced at marginal cost.

The calculations reported in the table illustrate a number of points. First the equilibrium number of products can be above or below the optimum. Cases of too few products tend to occur when cross elasticities are low relative to own elasticity, and fixed costs are high (see Groups III and IV). I should say here that linear demands have a strong tendency to produce too many products because the profit maximizing quantity is half of that when price equals marginal cost.

The welfare loss from the nonmarginal cost pricing of existing products (ΔT_3) varies considerably. Often it is less than half the total welfare loss (ΔT_2). Moreover, the cases where a substantial fraction of the welfare cost is associated with the wrong number of products vary. When cross elasticities are high (Groups I, II and V), they occur when fixed costs are low. When cross elasticities are low, they occur when fixed costs are high. These are also the situations where the optimal and equilibrium number of firms differ by a large percentage.

With nonlinear demands, similar effects hold; I do not have the space to report on them. The excessively strong tendency to get too many products can be avoided with other kinds of demands.

III. Hypothetical Products and Product Characteristics

A natural starting point for theoretical welfare analysis of monopolistic competition is the multiproduct structure of demand. For welfare purposes, preferences for actual and potential products are the proximate determinants of welfare problems. What the underlying attributes are, that differentiate the products, are not by themselves directly relevant. In terms of location problems, locations by themselves have no welfare consequences until the distribution of purchasers is specified. The same applies to products located in higher dimensional spaces of attributes.

When one turns to the task of implementing the welfare analysis of product differentiation empirically, the importance of the attribute space in which products are located, is elevated considerably. The empirical study of product differentiation requires the estimation of the demand for hypothetical or potential, as well as actual, products. This is a difficult problem, to say the least. The most natural way to approach it seems to me to attempt to estimate the structure of preferences for attributes rather than products. The rationale for taking this approach is that if the distribution of preferences for attributes in the consuming population can be estimated, then the demand for any set of products (not just the existing set) can be computed from that distribution. Attributes then, provide one way of dealing with the demand for potential products.

There is a considerable literature on this general subject, much of it centering on probabilistic discrete choices in transportation demand.

The goal of the theory is to derive functional forms for the demands for products as a function of their attributes. Attributes include price, physical and perform-

ance characteristics, aspects of service and so on.[7]

Using the attributes approach, one can in principle, compute the demand for any set of products, not just the existing set. In that sense, the conceptual apparatus permits one to compute the demand for hypothetical products, and therefore with a certain amount of cost data to draw welfare conclusions about the product differentiation aspects of industry performance. In future work, I hope to report in greater detail on the attributes approach and its application to industry and product data.

[7] For a discussion of probabilistic discrete choice models, see McFadden (1973).

REFERENCES

R. L. Bishop, "Monopolistic Competition and Welfare Economics," in R. Kuenne (ed.) *Monopolistic Competition Theory,* New York 1967.

E. Chamberlin, "The Product as and Economic Variable," *Quart. J. Econ.,* 1953.

A. Dixit and J. Stiglitz, "Monopolistic Competition and Optimum Product Diversity," Institute of Mathematical Studies in the Social Sciences, technical report no. 153, Stanford 1974.

K. Lancaster, "Socially Optimal Product Differentiation," *Amer. Econ. Rev.,* Sept. 1975, *65,* 567–85.

D. McFadden, "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka (ed.) *Frontier of Econometrics,* New York 1973.

B. Owen and M. Spence, "Television Programming, Monopolistic Competition and Welfare," *Quart. J. Econ.,* 1975.

M. Spence, "Monopoly, Quality and Regulation," *Bell J. Econ.,* Fall 1975.

———, "Product Selection, Fixed Costs, and Monopolistic Competition," *Rev. Econ. Studies,* forthcoming 1976.

R. Willig, "Welfare Analysis of Policies Affecting Prices and Products," research memo no. 153, Center for Research in Economic Growth, Stanford Univ. 1973.