An prosthetic Demand, Public Expenditures, and Full Cost Recovery

Alan Abuecher*

1. Introduction

In spite of the efforts of the last 25 years to establish a rigorous theory of what goods and services the public sector should supply and to analyze the implications of public sector participation for pricing, agreement on the classification and criteria for public sector expenditure is still far off, at least if we look with close and critical eye. When we do, we find several competing theories and classifications, different names being attached to the same concepts by different authors, and different concepts parading under the same name, sometimes even in the hands of a given writer.

It is more than a nice scholastic objection that the theory of public expenditure "has not yet attained rigid orthodoxy," as Buchanan once put it (1967, p. 115), and has failed to develop into a unified whole. The fact is that policymakers and economists have no basic standard by which to evaluate recommendations on specific issues, such as the pricing structure for urban or intercity roads, or to evaluate the petitions of special interests. The result, almost invariably, is that the prices charged to the directly benefiting private groups involve a subsidy of varying amounts. A benefiting group is almost sure to find, among the many public expenditure theories, some one theory that justifies provision at public expense of the particular service that it happens to need at the moment. For example, education is justified by reference to income distribution objectives; the argument is then extended to related activities such as rural roads to permit physical access, so that the prices for these activities are also exempted from the requirements of full-cost recovery from direct beneficiaries. Having developed the theme that secondary roads promote some desirable social goal beyond the mere transfer of goods and people between points, it is a simple step to extend the notion to primary roads (though not, of course, invoking the same specific reasons) since the distinction between secondary and primary roads is never very clear in any case. Or, again, marginal cost pricing of decreasing cost activities is freely invoked for the pricing of almost any aspect of railroad activity without ever bothering to ask, much less answer, the question of the dimension with respect to which marginal costs are supposed to be decreasing—annual flow, length of haul, size of marshalling yard, etc.

Meanwhile, many writers and practitioners in the fields of public finance and social choice, having become disenchanted with arguments over normative justifications for public sector activity, are all too disposed to start their analyses of taxation or pricing policies by invoking the always-at-hand

---

*University of Toronto. Comments by William Baumol and E. Ray Captner are on an earlier version of this paper are gratefully acknowledged. Responsibility for any errors remains with the author.

"I would like to thank William Baumol and E. Ray Captner for their very helpful comments on an earlier version of this paper. Responsibility for any errors remains with the author."
expression "public good" as though there were universal agreement on what that term meant and as though its use were sufficient justification for public sector participation in the first place.

We have said enough to indicate the need to review the theory of public expenditures and to analyze the implications of public sector participation (investment and operations) for the pricing of the activities. Our conclusions are that, in addition to Musgrave's well-known three-way classification of justifications for public sector activity (stabilization, allocative efficiency, and income distribution), a fourth (national political objectives) and a fifth (legal-administrative efficiency) should also be specified. While agreement has never been reached on precisely what is the hallmark for allocative efficiency—what is it about an activity that requires it to be performed in the public sector—all formulations of the efficiency justification include the inability to price to recover full cost while providing an optimal activity level. But, we will argue, for activities performed in the public sector for reasons of legal-administrative efficiency, nothing other than full cost recovery is implied. For these activities the most basic Lerner-type welfare maximizing pricing rule really leads to full cost recovery. Further, the range of activities which fall under this rubric is much wider than many traditional economists would probably be willing to accept and include many of the activities most cherished and used by writers as examples of allocative efficiency-related public sector activities—activities such as fire and police protection or highway services.

We will proceed as follows. In Section 2 we see what pricing policy implications follow from Musgrave's classification and indicate the direction along which his analysis must be supplemented. Section 3 treats the Bowen-Samuelson apparatus which has been advanced as the most demanding classificatory test for those public expenditures which are undertaken for allocative efficiency. As is well known by now, this apparatus is inadequate in itself to distinguish "private goods" from "public goods." Then we examine externalities and in Section 5, cost-based public expenditure criteria. We conclude that the simplifying assumptions made in the service of enlightenment represent a serious distortion of reality and change the problem. Section 6 discusses legal-administrative efficiency as a justification for public sector participation and examines its pricing implications.

2. Musgrave's Classification and Its Pricing Implications

The most fruitful contribution of the last 20 years to the understanding and analysis of the public sector is without a doubt Musgrave's three-way classification of public sector activities: under the headings of stabilization, income distribution, and allocative efficiency (1950). We believe that two additional functions should be added. The first is the attainment of national political objectives, and the second is legal-administrative efficiency.

What pricing policy implications follow from this classification? For stabilization activities, there is a strong presumption that less-than-full-cost recovery should be tolerated. If it is decided that the government must undertake a project to promote employment, it must be because the private sector, left to its own, would not find it profitable to do so at the particular stage of the business cycle, and this means that full cost could not be recovered. However, the question of pricing really is not relevant to most stabilization activities, which are fiscal and monetary rather than tax policies (including special incentives such as investment credits), interest rate policy, and control over the money supply which the government undertakes for stabilization purposes.

Income distribution policies are designed to influence the income shares accruing to various socioeconomic classes. Welfare payments, progressive personal income taxes, and progressive excise taxes represent specific income distribution policies. Education, not always recognized as an income distribution objective, also falls under this rubric, since it enables children of low-income families to aspire higher than they could if education were to depend solely on their parents' ability to purchase education. Clearly, for such activities full-cost recovery would be inconsistent with achievement of more equal income distribution.

Activities undertaken for national-political objectives must be paid for by the public at large. Given the usual openness of the tax structure and assuming reasonable success in implementing it, we can assume that in most countries a rich person pays much more than a poor one for each such activity. But the welfare economist must not relax his vigilance with this recognition. Indeed, this public expenditure justification may exercise him the most since it is so frequently invoked for activities to which it really is irrelevant. For example, it could be argued that the construction of the Canadian Pacific in the 19th century served the national political objective of confederation and, consequently, the investment cost and possibly even part of the operating expenditure should not be laid at the shipper's door. But it does not at all follow that the continued operation of the railroad serves a national-political purpose rather than an economic purpose and, therefore, there should be no presumption that anything other than a Lerner-type, incremental social cost-equivalence guideline should be imposed. The term "incremental social cost" seems preferable to "marginal social cost" because the latter is too closely associated with a simplified two-dimensional world which assumes away both the complexity of the real world and the mechanism by which efficiency generating prices can be imposed. This is examined further in Section 5 below.)

Resource allocation objectives, that is, the efficient use of resources, also may justify public sector activity. For, although the principles of perfect competition hold under certain classical conditions to efficient resource allocation, various obstacles to the fulfillment of these conditions do arise in practice, and public sector activity may help to overcome these. The questions that remain to be answered are: What circumstances give rise to such obstacles? and What are the implications for pricing policy? The traditional view, which is not under dispute here as long as it is properly restricted, is that for one of several reasons we cannot price to recover full cost and at the same time attain the optimal activity level so that efficiency requires public sector participation and subsidy. Three different, but not entirely independent, reasons have been the focus of attention here—vertically additive demand curves (which we will prefer to call "anaprosthetic demand"), externalities, and decreasing cost.

3. Anaprosthetic Demand and Joint Products

Soon after its introduction by Bowen (1948) and Samuelson (1954) the vertical addition of demand curves was thought to constitute a sufficient argument for assigning activities to the public sector for efficiency purposes, although the term in original or modified form—e.g., "pure public," "semi-public"—soon came to be applied to other constructs. Today, it is probably safe to say, the latter expression conveys the notion of an "efficiency-related public sector expenditure" although in its original form it was intended
to relate to a very special kind of activity.\(^1\) On the other hand, it was quickly perceived that the apparatus had a much wider applicability outside the public sector, its first application concerned the problem of peak pricing (Stie- ner, 1957). Samuelson himself used it to handle joint products in general and argued that public goods could be distinguished from joint private goods through their equilibrium welfare conditions (1969a).

To divorce the concept of vertical additivity from the notion of public sector expenditures, with which it has been almost exclusively associated in the literature, we call such demand relations "aneprosthetic" or upward additive. It may be worthwhile at the same time to denote by "paraprothetic" those demand relations that are literally added in the customary manner.

Typical examples of "public goods" cited in the literature include flood control, police activity, defense expenditures, television broadcasting. In all of these the quantity available to one person is not diminished by availability to others. The value of some given amount of the good is equal to the sum of the prices that people are willing to pay for that quantity, and differs from the horizontally additive case in which a given quantity represents the sum of the quantities that such a person would wish to consume at that price.

In distinguishing between private joint goods production and public goods production, Samuelson used the example of mutton and wool. Individuals' demand curves for mutton and wool are summed horizontally. But the demand for sheep is derived from vertical summation of the wool and mutton demand curves (the product curves are assumed to relate to the amount of product per sheep). The demand for sheep in Samuelson's example is an aneprosthetic combination of industry demand curves for each product.

Evidently, then, the aneprosthetic demand curve has much wider applicability than was originally thought. As Steiner showed by example and Samuelson showed pedagogically, it can be used to analyze pricing and output in many situations in both the private and public sectors. Clearly, additional conditions must be dropped before it can be used to classify goods into public and private—the construct is powerless by itself to do so. Samuelson attempts to circumvent the impasse through a set of optimality conditions whose essence is that private joint goods can be ultimately resolved into paraprothetic demand relations and public goods cannot. The pioneer multipurpose water resource project, the TVA, for example, can be resolved into joint products (national defense and flood control), since each activity is itself aneprosthetic. But this is still not enough, since many goods, such as a symphony concert or a baseball game, whose privacy would not be questioned also have an aneprosthetic demands not further reducible into paraprothetic curves! The ultimate reason is to the Samuelson pure theory why some goods with aneprosthetic demands must be publicly provisioned is that the subdemand curves will never be revealed, making it impossible to know what prices should be charged to individual users. Therefore, they should be paid out of centrally collected funds.

Two major objections can be raised to Samuelson's pure theory.\(^2\)

The first concerns internal consistency. If the subdemand curves, i.e. Samuelson's "pseudo demand curves" (1969a, p. 28) cannot be known, it is not possible to know total demand. Therefore, no public agency would know the right amount to supply, and it would not be known whether the actual provision attained an allocative optimum. Is public sector overprovision superior to private underprovision with monopolistic and possibly inefficient price discrimination? How much over, how much under, and how much better? This difficulty has been acknowledged by Samuelson: "If you know enough to use the Lindahl pseudo prices, you don't have to use them" (1968b, p. 117). We will argue in Section 6, however, that we still may wish to undertake flood control projects or provide police service in the public sector, but the justification is more to be found in administrative-legal arguments than in the joint nature of the activity and in this case no subsidy to users is necessarily implied. For example, if the demand for flood control will have differential impacts on land use (including crop use), the income or property tax mechanism should both reflect people's demand curves and provide a method to collect payment. To be sure, some uses will escape the tax and be beneficiary to an externality—e.g., duck hunting—but this need not trouble us anymore than many private sector external benefits in consumption arising from joint supply need provide an objection to performing them in the private sector.

A second objection concerns descriptive adequacy. As a descriptive device the classification breaks down when we try to apply it. Consider the case of television broadcasting, one of the most often cited examples of pure public goods. The demand is certainly aneprosthetic. It is, however, privately supplied in many countries. The aneprosthetic approach is appropriate for examining public sector activities engaged in to achieve national-political objectives such as national defense, or, at least, it would be if we could figure out what to put on the x-axis. We would argue, however, that this has nothing to do with economic efficiency. It is, in particular, important to stress this because both the aneprosthetic mechanism and the term "public goods" are frequently taken as the hallmark of—or even as being synonymous with allocative efficiency, a point to which we return in Section 6.

4. Consumption Externalities

Since public sector activities are often observed to give rise to external economies, many writers believe that they necessarily involve externality. These externalities are supposed to represent something more basic than the adventitious consequence of an imperfect pricing mechanism which settles for less than full-cost recovery. But less than full-cost recovery is accepted in the first place because of the belief or expectation that conditioned by the expectation that a price mechanism cannot be designed to transfer enough of the producer or consumer surplus to reimburse fully the agents providing the activity.

The case of consumption externalities may best be regarded within the framework of aneprosthetic demand.\(^3\) We may define the

\(^{1}\)Again we remind the reader that what is meant here is the Samuelson pure theory of allocative-efficiency-related public expenditure. The only way to neutralize the term would seem to be through the introduction of the neoclassical expression proposed here.

\(^{2}\)Many apparently different formulations of consumption externalities may be cast into this form. For example, the Buchanan-Kafkalis (1962) case of non-reciprocal
case of pure consumption externalities as an intersection of the supply curve and the anoprosthetic demand curve at some point where the latter is coincident with the demand curve on only some subdemands. The situation is shown in Figure 1. The two individual demand curves (labeled D1 and D2) intersect the X-axis at different points, Q1 and Q2.

The anoprosthetic demand curve D3 is the vertical summation of the component curves. In this case, it has a bend when the demand price for the first component becomes zero at K, and coincides thereafter with the demand curve D1 until intersecting the X-axis. As drawn, the demand function D3 is everywhere below D2 but this is not necessary. D3 could intersect the Y-axis at a point higher than the D1 intersection.

If the supply curve intersects D3 to the right of K, the total cost of the product could be paid by the demand represented by D2. In this case D2 will be satisfied at zero price. This follows in a straightforward manner, either from the diagram or from the Bowles-Samuelson analytic conditions for anoprosthetic demand. On the other hand, a price mechanism might be found to recover part of the cost through the first component. From an allocative standpoint it makes no difference as long as the average cost curve is horizontal. If the average cost curve is declining, allocative efficiency requires that part of the cost be recovered from the first user and that the activity level be at the intersection of MC and D0, which will, of course, occur farther out and, as before, in the range of coincidence of D1 and D2.

Now, what have we just said for public goods applies with equal force to the public or private sectors. This may not be immediately apparent, since in Samuelson's example the diagrammatic treatment shows the supply-demand intersection at positive prices for both products (1969a, p. 28). But it is an obvious possibility and one that has prevailed at various times in many industries. For example, precisely this situation existed in the meat market in our parents' generation, when some consumers were regular beneficiaries to the street to the nearest butcher shop if asked to pay for it at the first shop. Eventually, as demand for liver and other by-products caught on, a positive price came to be feasible.

The existence of the foregoing externalities in the meat market was due to the inability to impose a positive price on some of the subcomponent demands. This was simply a private market manifestation of the "nonexcludability" principle that is often invoked today as justification for public sector activities without specific charges. Roads, for example, are said to be nonexcludable because it is difficult to charge in proportion to benefit, and so externalities arise. But the point should be emphasized that these externalities are adventitious. They are not inherent in anoprosthetic activities. And, as noted, under certain demand configurations it can also be very difficult to devise a properly discriminating pricing system for goods such as meat, the propriety of whose supply by the private sector is almost self-evident. The problem is to find the right pricing mechanism. Until this is done some sacrifice in social welfare is implied if (1) the marginal cost curve is falling (see the next section) and (2) efficient discrimination cannot be practiced. But the reason is to be found in the decreasing cost and not in the vertical additivity of the underlying demand curves. If the cost curve is horizontal, it does not make a difference. Therefore, whether resource allocation can be made more efficient through the influence of the public sector will not depend on the generation of externalities, but on the slope of the cost curve.

5. Decreasing Costs

The traditional efficiency justification for public sector involvement in economic activities stems from the assumed decreasing cost nature of the activity. The development of this theory is usually associated with the names of Dupuit (1844) and Hotelling (1939) and reduces essentially to the following argument. For activities with large capital outlays which have relatively low average short-run marginal cost (variable cost may even be zero), if we are unable to discriminate among consumers, total cost recovery implies pricing at average total cost which exceeds marginal cost. This means that consumers whose value is below average total cost but above short-run marginal cost will be kept out of the market. Society as a whole suffers a loss on each one of these consumers equal to the difference between his forgone utility per unit and short-run marginal cost.

Although the argument appears persuasive as usually presented, in most cases it cannot withstand the attempt to answer the most basic question: how should the output be determined? With respect to which costs are said to be marginal, are defined? For example, if we have in mind a highway network, would we think of marginal cost as the cost of expanding the network, the differential cost of building wider rather than narrower road beds in the network, the differential cost of a more ramified rather than less ramified network, the differential cost of building more durable roads, the incremental cost of operations this year (snow removal, repaving, administration, etc.), or of some other...
subset of these items, the cost of one more vehicle mile, or, although many transport economists resolve to recognize even the possibility of some, the disproportionately much greater damage imposed by heavier loads, a fact amply documented by the American Association of State Highway Officials (AASHO) in its 1962 report on the road tests it conducted. Clearly, the usual attempt to squeeze this complex set of relationships into a two-dimensional diagram cannot be said to illuminate truth by simplifying the fundamental relationships, rather, it must be said to permit a neat conclusion to be reached by assuming away the problem.

In less complex situations, it is never very clear whether it is long-run or short-run marginal cost-price equalization that writers intend. Let us imagine the simplest case—a unique technology and zero variable costs. This is precisely the situation that Dupuit was assuming—although he devoted no extended diatribes or apologies for it. But, as showed in an earlier paper (1976), the fact is that he was discussing a footbridge which is not part of an intercity road network, and this does imply uniquenesses of design and homogeneity of output (pedestrian crossings), which severely restricts the range of its applicability.

If the bridge were a carriageway bridge, we would expect it to be part of a network; while the other end of town might be the ultimate destination for pedestrians, a carriageway bridge would be tied in to an interurban network. Thus, whereas we can concentrate on a homogeneous output (pedestrian crossings) in the footbridge case, output in the carriageway-network case would have so many dimensions that it is not obvious what should go on the X-axis—trip-lengths, vehicle-miles, trips, tons, miles, or any of these subject to specified vehicle size distribution, not to speak of the different cost dimensions that we may associate with expansion of the network. What is marginal cost?

Few people have recognized the severe limitations of Dupuit's analysis—a partial equilibrium analysis—and very restricted at that, a two-dimensional framework which cannot accommodate the much more complex real world. For example, what is marginal cost in the U.S. Interstate Highway System which grew almost 30 times, over a period of nearly 20 years, with a population growth of 30% or more spreading over vast and relatively new regions of the country, and carrying an extremely diverse variety of traffic.

The foregoing considerations do not mean that the decreasing cost-justification for public sector provision without full cost recovery from users can never be applied, but they surely restrict its scope much more narrowly than practitioners usually allow. It would certainly hold for a bridge like Dupuit's but most such examples today would probably be small walkways in parks and playgrounds rather than roads of communication properly speaking. The parks and playgrounds themselves, to the extent that no rising cost curve due to congestion is encountered and no materials are becoming consumed, should also be provided under this heading, i.e. as allocative efficiency public sector activity paid for out of general tax revenues (originating locally, of course, unless there is some desired interregional transfer, serving as proxy for an income transfer).

6. A Legal-Administrative Theory of Public Sector Efficiency-Related Participation

We have examined the major arguments commonly advanced to explain or justify public sector resource allocation activities and showed that none of those furnish an unambiguous descriptive or normative criterion. But this does not mean that direct public sector economic activities must be restricted to only two of Musgrave's three branches, those that discharge income distribution and stabilization functions. Public sector resource allocation activities can also be justified on grounds that, for want of a better term, we call "legal-administrative efficiency." In such cases there is no predisposition for not imposing full-cost pricing. That we have not referred to a traditional economic-theoretic notion suggests that it is easier to explain this viewpoint by reference to individual examples rather than through a theoretic generalization. The following examples might be considered.

1. An intercity highway network may be viewed as a public resource allocation activity in the public sector because land assembly is much easier for the public sector than for private developers, thanks to eminent domain rights. Any controversies regarding individual parcels of land can be settled in court if they cannot be decided by direct negotiation and the prospects of early solution and early project completion are much better when this procedure is followed. Spoilers cannot then begin negotiations by holding out for high prices, artificially inflated by the belief that the assembler has already put together a ribbon of lots and is willing to purchase the missing lots at almost any price. This enables the public sector to execute the whole project relatively fast, reduce the period of resource immobilization, and minimize social costs.

But the fact of public sector participation should not itself imply anything other than a full-cost user pricing approach, which should be pursued unless there is some reason for wishing to redistribute income, such as providing access to rural schools as part of an indirect income transfer.

We note that the right of eminent domain can be used often is delegated by the public authority to a private firm. However, because there are economies of scale in road network construction and administration, it would be unnecessarily costly to delegate parcels to different firms. On the other hand, no single firm is likely to wish to take responsibility for the whole network, so that it is administratively more convenient and less costly for the public sector itself to retain responsibility for this activity.

2. In Section 4 it was argued that many public services commonly thought not to be susceptible to benefit-taxation principles could indeed be paid for in this way while maintaining welfare efficiency. For example, the greater the value of the house, the greater, on the whole, is the demand for police protection and the more do people in fact pay for it. But even in the private sector provision of police services, therefore, cannot be justified on the grounds of "nonexcludability" or "free riders." Why then, should police services be provided publicly instead of by a private army contracting its services to all members? The problem of pricing is essentially a matter of determining the demand functions of individual house owners, which, it has just been argued, need not present a formidable obstacle. In this case a private police force could simply be given the right to bill each house owner for its services each year according to real estate value and to take whatever legal remedies were necessary to ensure compliance. However, public sector provision of police services can be justified on several other grounds.

First, public provision is an administrative convenience. Public buildings must be protected, public functions supervised, and certain laws enforced that are less obviously related to home and personal protection. It is more convenient to handle these through a
police force directly responsible to the government than through a private force. Once it is decided to have a public sector police force, some efficiencies may derive from a larger single force. But we emphasize that it is not the fact of the scale economies that alone determines the public sector subordination of this activity —scale economies do not lead inexorably to public sector subordination of, for example, water works, electric power, or gas supply, although in some instances these are provided by the public sector. Furthermore, greater care can be exercised over weapons when police searches are provided publicly.

Finally, even though in principle it would be easy for a private firm to receive a mandate to levy police charges, collection costs upon refusal to pay might be high. The private police force would have no resort comparable to that of an electric utility whose explicit threat of discontinued service and disconnection reduces its collection costs. There is no immediacy to the service provided and all that the police firm could do is threaten not to provide the service when the occasion arises. Legal action to ensure payment is costly. All these legal-administrative arguments thus support the consolidation of the police charges with other charges, all of which are then imposed through the central tax collection system.

3. A similar argument could be made about many other goods and services furnished in the public sector. Consider fire protection. May not we assume that the lesser risk of closer proximity to the station is reflected in a higher property value, which then yields a higher tax revenue? We do not suggest that an exact relationship must exist (just as private sector insurance premiums are scaled stepwise to belts rather than continuously in distance); but we would certainly expect to find consistency with the principle that people pay according to their expected utility from the service.

What of scale economies in provision of fire protection? That fire protection should be provided by a single company is undoubtedly consistent with the scale economies of this activity, just as a single electric utility usually services an area. But why should the fire station monopoly not be a private monopoly just like the power company if, as we claim, both services can be paid for according to demand? The explanation is to be found in administrative convenience of collecting payment. As noted, the power company can turn the power off to delinquent customers without serious consequence. The private fire company, even given the mandate to tax individual homeowners, might incur serious collection costs from some recalcitrant individuals. Not that they would dissemble about whether or not they really want the service, as in the usual game-theoretic approach to public allocation theory —they would simply refuse to pay. But unlike the power company the fire company has nothing to turn off until the need actually arises, and legal remedies are costly. And ethical norms would not permit society to withhold the services of the fire station from the recalcitrant homeowners and their families, in case of fire. When the state provides the service, however, the scale economies of collection can again be exploited, and the fire-fighting charges consolidated with taxes intended to cover other activities. The cost of legal remedy on refusal to pay the entire consolidated tax would be much smaller in relation to the amount which would now be involved.

7. Conclusion

This survey of the main principles advocated for public sector activity shows that economic theory does not provide unambiguous descriptive or normative criteria for deciding (1) whether the public sector should participate and (2) when it does participate for resource allocation purposes, what the pricing structure should be. In particular, we should keep the following guidelines clearly before us when it is necessary to make pricing decisions.

1. Extending the distinction made by Musgrave, the public sector may participate directly for stabilization, distribution, allocative efficiency, national-political objectives, or administrative-legal efficiency. Activities undertaken for the first motive, and almost certainly for the second as well, will entail non-compensatory prices, i.e. full costs will not be recovered from users.

2. The efficiency justifications for public sector participation are essentially of two types: the first involves activities in which the added collection costs or the forstalling of the use of nondepleteable resources would signal a real social cost and for these a non-user contribution, i.e. a subsidy out of general (but local) tax revenues is indicated. But the second efficiency justification refers to activities which are merely provided in the public sector for purposes of administrative and legal convenience and carry no indications of other than full cost recovery.

References


---, "(1977) Transportation Economics and Public Policy, with Urban Extensions (New York: Wiley Interscience),


Dupu, I., De L'utilité et de sa mesure. Edited by M. de Bernardi Turin, 1933.


