

Nordic Journal of Political Economy

Volume 29

2003

Pages 87-104

Public Provision and Private Incentives

Agnar Sandmo

This article can be downloaded from:

http://www.nopecjournal.org/NOPEC_2003_a06.pdf

Other articles from the Nordic Journal of Political Economy can be found at:

<http://www.nopecjournal.org>

Agnar Sandmo*

Public Provision and Private Incentives**

This paper surveys classical and modern arguments for public production and provision of goods. It reviews the conventional case for public production under conditions of increasing returns and discusses the modifications that have to be made if public production involves a cost inefficiency. It then discusses the causes behind a possible cost inefficiency, such as the difficulty of designing good incentive schemes in agencies with multiple and complex objectives. An alternative to designing better incentives in the public sector is that of contracting out to private firms, and the conditions favourable to this alternative are also discussed. JEL Classification: H11, H40.

At the present time, many people – economists, other social scientists as well as politicians – are engaged in a process of re-thinking the boundaries between the private and public sector. Where that boundary should be located is one of the classical problems of economics in general and of public economics in particular. Today the dominant attitude of both economists and politicians in many countries seems to be that the public sector has become too large. Of those who subscribe to this view there are undoubtedly some who believe that part of what the government does – such as spending on culture, environmental protec-

tion and redistributive transfers – is simply not worth doing. Others take less issue with government objectives but hold that the organization of the public sector is such that the objectives are pursued at much higher costs than necessary. For the government to operate more efficiently, it is claimed, it should reform its internal systems of resource allocation and rely more on private agents and individual incentive mechanisms.

One of the central elements in economic thinking has been the insight that markets have the ability to allocate resources in a socially efficient manner. From the time of Adam Smith to modern welfare economics

* Norwegian School of Economics and Business Administration, Helleveien 30, 5045 Bergen, Norway. E-mail: agnar.sandmo@nhh.no.

** This paper formed the basis for an opening lecture at the 58th Congress of the International Institute of Public Finance in Helsinki, August 2002. I am indebted to Vidar Christiansen, Avinash Dixit and Kåre Hagen for helpful comments.

this insight has been interpreted as having the following important implication: *If you wish to argue that the public sector has the better system for resource allocation, then the burden of proof rests on you.* Hence, at least according to economists' way of thinking, the arguments for public sector provision of goods and services and for the regulation of private markets should be derived from a diagnosis of market failure. This implication has enjoyed wide acceptance in the economics profession.

Some public finance economists may, however, have been guilty of considering this burden to be lighter than it ought to be. One interpretation of it is that if you have identified a case of market failure, you have thereby also established a case for public sector involvement. (Assar Lindbeck likes to tell the story of the song contest where the jury, after having heard the first singer, decided to declare the second one to be the winner.) This interpretation may, as argued e.g. by Shleifer (1998), have been a common one among leading economists in the 1930s and '40s. Without doubt, it has definitely been less common in recent decades. Nevertheless, we should be grateful to public choice theorists like James Buchanan for continuing to remind us that this kind of argument is seriously incomplete. To establish a convincing case for public sector involvement, you should be able both to point to a market failure *and* to argue convincingly that the public sector is able to handle the problems involved in a better way. The last part of this requirement is far from trivial; there is no particular reason to believe *a priori* that actual bureaucrats and politicians will be motivated to take decisions in accordance with the prescriptions of welfare economics and public finance. Recommendations regarding the division of labour between the private and public sectors should take

account not only of market failures but also of the possible failures of policy.

In my view, neither our historical experience nor the developments in economic theory has invalidated the traditional approach to normative public economics that takes its point of departure from welfare economics. It remains important to identify the nature and causes of market failure and to study optimal public policy in the presence of such failures. As a part of the public discourse about the appropriate roles for markets and government, this is a very important input, and there are hardly any other suppliers of this input than the economics profession. At the same time, our analysis and recommendations should be based on realistic assumptions concerning both the functioning of markets and the workings of government, and it is in this area that modern developments to a very significant degree has advanced our understanding. This seems a good occasion, therefore, to try to take stock of where we stand. Accordingly, I wish to re-examine the standard arguments for public involvement in the light of modern developments in the theory of incentives for both private and public agents.

This is an extremely broad area, and there are some topics that inevitably have to be left out. The most important of these is the whole area of redistribution, pensions and social insurance. Thus, I do not discuss the issues of the design of pension and insurance systems and the respective roles of the private and public sectors in this field. The focus is on the division of labour between the private and public sectors in the production and provision of goods and services.

Market failure and government activity

Different taxonomies of market failure have been suggested in the literature. The basic

definition is that there is market failure if the price system fails to establish a Pareto optimal allocation of resources. This could happen either because the market equilibrium is non-competitive (as with natural monopoly) or because there are structural features of the economy that prevent even perfect competition from establishing a Pareto optimum (as with externalities). But since an allocation which is not Pareto optimal implies that there are unexploited gains from trade between individuals, there must be a lack of markets, or more generally, a lack of bargaining possibilities, which prevents individuals from capturing these gains. This is the perspective taken in a famous article by Arrow (1971), in which he identifies market failure with incomplete markets that again are explained by transactions costs. The transactions costs related to the operation of markets explain why it may be socially efficient to use other systems of resource allocation – in particular political and bureaucratic decision-making – as supplements to the market system. In this view, the political system and the public bureaucracy are the most important arenas that individuals can use to overcome the transactions costs connected with private bargaining solutions to the problem of market failure.

This is an illuminating general perspective on the problem, but there is also a need to identify more specific causes of market failure. A standard classification consists in ascribing market failure to the existence either of increasing returns, public goods, externalities or asymmetric information. Obviously, these categories are to some extent overlapping. A communications network with high fixed and low variable costs (and consequently with increasing returns) has also elements of a public good, externalities caused by manufacturing or

transportation cause environmental public goods to deteriorate, and the generation of new information has positive externalities for other agents in the economy. Still, it is useful to think of market failure in terms of these four separate categories.

In addition, of course, one may wish to correct the market outcome in terms of the personal distribution of income. The possible inequity of the market-determined distribution of income is not usually classified as a market failure. The reason for this is simply that the market mechanism ideally promises to deliver efficiency, but it offers no promise of justice and fairness. Nevertheless, redistribution is both in theory and practice an important reason for public interference in the market mechanism. Indeed, some of the inefficiencies arising from distortionary taxes and regulations are difficult to explain except as the side effects of attempts to redistribute income between individuals and social groups.

Obviously, I do not claim much in the way of originality for this classification of the sources of market failure, and some would no doubt prefer to use different concepts and classifications. But if there is lack of originality, this must imply that there are many other economists who think about the problems of market failure in this way. Accordingly, this is a strong point of the classification, since my aim is to structure the discussion along the lines in which most economists think about market failure and its implications for the appropriate balance between the private and public sector.

Public sector involvement may take place at several levels. At one level there is public ownership and production, which is clearly the «heaviest» kind of public involvement (national defence, the police force, public schools etc.). At an intermediate level there is public provision of particular services but

without public production (as when garbage removal, while a public responsibility, is contracted out to private producers). At lower levels of involvement production and provision remain in private hands but is subject to government regulations, taxes or subsidies. Clearly, in a modern economy there is hardly any private producer who is not subject to some kind of government regulation, does not pay taxes or does not receive subsidies. What I have in mind here are regulations, taxes or subsidies that are targeted on specific objectives, so that they are designed with a view to make these activities conform to some kind of public objective.

The argument for public sector involvement may draw on multiple sources of market failure. Thus, the argument for public transportation is usually seen as mainly involving increasing returns, but externalities such as the relief of traffic congestion also play a part. The provision of social insurance can be justified from considerations of asymmetric information that lead to market failure in private insurance, but redistributive arguments and increasing returns have also played important roles.

Increasing returns and public production: The received wisdom.

In the history of economic thought, increasing returns has been considered as providing the strongest argument for public production. When we speak about increasing returns as a source of market failure, we are clearly referring to the case where average cost is decreasing throughout the range relevant for market equilibrium (or at least through a substantial portion of that range). In that case, an equilibrium with a number of firms sufficiently large to justify price-taking behaviour, at least as an approxi-

mation, is unsustainable, and equilibrium prices will be above marginal cost. Because of the cost advantages of large-scale production, the equilibrium will converge to monopoly, hence the term «natural monopoly». But this implies that the equilibrium price will be above marginal cost, so that we have a clear case of market failure and therefore a case for government action. One alternative for a government that wishes to overcome this market failure is to convert the private monopoly into a public utility and set prices equal to marginal cost, thus ensuring a socially efficient allocation. This alternative was e.g. strongly recommended by James Meade, writing in 1944:

«Where a community needs only one gasworks, or electricity station, or railway network, monopoly must obviously exist. In these cases, socialisation in one form or another, of the industries concerned, is the only radical cure to ensure that they are run in such a way as to equate marginal costs to prices of the product produced (or the prices of the factors of production to the value of their marginal products) rather than to make a profit» (Meade and Fleming 1944, p. 322).

An early version of this argument goes back to Dupuit (1849), and it was later restated and analyzed by a number of prominent economists, e.g. Hotelling (1938), Vickrey (1948) and Johansen (1965), to mention but a few. The basic argument is that since a public utility is not constrained by the market to make a profit in order to survive, it should use its liberty to set prices that lead to a socially efficient adjustment of production to consumption. Since this implies that the marginal willingness to pay should equal marginal cost, consumer prices should be equal to marginal cost, and the resulting

deficit should be covered by a transfer from the government. In the special case where variable unit costs are constant, the transfer should simply be equal to the amount of fixed cost.

Let me add for completeness that the optimality of marginal cost pricing presumes that this policy is better than the alternative of not producing at all, thus saving both the fixed and variable costs. For the level of production corresponding to marginal cost pricing to be optimal it must be the case that the resulting consumers' surplus exceeds the fixed costs. Another extension of the analysis is to the case of multiple products, where the case for marginal cost pricing obviously applies to each individual product. Here the comparison between private and public production would have to take account of the fact that the private monopoly would not only charge higher prices for each product, but would also apply stricter criteria for a particular product to be produced. Instead of requiring consumers' surplus to exceed fixed costs, which would be the criterion to be applied by a welfare-maximizing public firm, the private monopoly would demand that each product's contribution to profits exceed its specific fixed costs. So not only would the private natural monopoly produce too little and charge too high prices; it would also tend to produce too few products.

However, there are some rather fundamental objections to the marginal cost pricing argument for public production, and this was also pointed out by some of the earlier writers on the subject. Thus, Johansen (1965) notes that for the welfare theoretic

argument to support the alternative of public production, it is necessary that the public sector really implements the optimal pricing rule, and that it is less than obvious that this will in fact be the case. He also emphasizes that if the deficit has to be financed by distortionary taxes, the efficiency losses that are avoided by the optimal price policy will have to emerge somewhere else in the economy, and that this weakens the case for marginal cost pricing¹.

These insights had in fact been stated in the literature a good deal earlier. Thus, almost fifty years before Meade, Knut Wicksell (1896), after a compact analytical statement of the marginal cost pricing rule, commented that many public utilities in his time did not understand the principle of efficient pricing and attempted instead to run their operations with an accounting surplus:

«The existence of such a surplus as, for example, the spectacular profits of the Prussian State Railways, may be a shining testimony to the efficiency of the administration and to the prosperity of the industrial and commercial life of the country; but at the same time the surplus also indicates that the enterprise is far from its optimum degree of utilization both in national and in individual terms. The passenger and freight traffic of the Prussian State Railways would probably increase very substantially with an appropriate reduction in rates. Everyone would gain thereby and no-one need lose, provided only that the ensuing deficit be

1. Interestingly, both Hotelling (1938) and Johansen (1965) state that this problem is of little or no importance if the deficit is financed by direct taxes or (in Johansen's case) by a uniform tax on all commodities. The more or less implicit assumption underlying this assertion seems to be that the distortionary effects of either of these taxes would be negligible, an assumption which few present-day economists are likely to support.

financed by taxes in a suitable manner.» (Quoted from Musgrave and Peacock (1958, p. 103)).

The economists that were cited above would probably all agree that if the policies adopted by the Prussian State Railways could be interpreted as a case of monopolistic pricing, it would be inconsistent with the argument for organizing them as a public utility in the first place. This would also be the case if the policy took the less extreme form of aiming at just covering total costs with profits set equal to zero.

Wicksell's story of the Prussian railways can be seen an early example of the public choice criticism of the more naïve form of normative public economics; to prove that a particular policy is optimal does not ensure its implementation. Moreover, his cautionary remark about taxes being levied «in a suitable manner» provides some of the motivation for the literature on second best price policies for public utilities, where the problem is to design an optimal price system, given a revenue requirement in excess of variable costs. This is the problem that was formulated and solved by Boiteux (1956). The price structure that he derived is closely related to the Ramsey optimum commodity tax structure; e.g., with independent demands (zero cross price elasticities), price markups above marginal costs should be inversely related to the direct price elasticities of demand. This is a second best solution to the problem of the optimal price system for public firms, which accordingly does involve distortions. In fact, the structure of prices is in important respect similar to the one that would be chosen by a profit-maximizing private monopolist. The difference between the two structures is primarily that the level of prices is higher under private monopoly, relative to competitive prices elsewhere in the

economy. The difference in pricing and social efficiency between the Boiteux type of public utility and a private monopoly is therefore less than suggested in the first best arguments of James Meade and others.

Other justifications for public production

There are a number of actual examples of public production which can hardly be said to involve increasing returns to any significant degree. In many countries, governments have a big ownership role in regular manufacturing companies as well as in the financial sector. There is a variety of historical explanations for this, but it is widely acknowledged that the theoretical support for this kind of government ownership is weak, certainly much weaker than alleged by a number of leading economists in the 1930s and '40s (Shleifer 1998). There is also substantial – although not unequivocal – empirical evidence that privatization in such cases has led to substantial efficiency gains; see e.g. the survey by Megginson and Netter (2001). However, there are other areas in which government ownership is in fact substantial even in predominantly market economies. In particular, governments are heavily involved in institutions of education and health care, although most countries also have a substantial share of private production in these sectors. For the moment I abstract from the important issue of whether the government should produce these services itself in contrast to providing them on the basis of a contract with a private producer. I simply assume that the choice is between public and private provision and production.

Why does the government involve itself with the provision of goods and services that could in principle have been allocated

through the market? One reason has to do with distributive justice; removing these services from the market system makes their availability to individual citizens less dependent on income. Economists' instincts might be instead to recommend redistribution of income which would allow the poorer individuals to buy more of educational and health services, but the tax-transfer mechanisms that are available may not be seen as sufficiently fine-tuned to achieve this goal on their own². In addition, many people in society may be concerned with what Tobin (1970) called «specific egalitarianism». They are less concerned with the overall distribution of income and welfare than with the distribution of specific goods like health and education. One reason for such an attitude might be that equality of access to such goods is important for ensuring equality of opportunity, especially for the young, while e.g. progressive taxation is mainly designed to achieve equality of outcomes. There is an element of paternalism in this, but should we non-paternalists forbid people to have paternalist preferences? In any case, the line between paternalism on the one hand and a concern for equality of opportunity on the other is not an easy one to draw.

Elster (1992) has emphasized that there are different conceptions of justice or fairness that apply to different sectors and institutions of society. While the allocation of ordinary market goods is allocated on the basis of willingness and ability to pay, educators think it just that scarce educational resources should be allocated according to ability, and doctors believe that medical

resources should be allocated on the basis of need. In holding these beliefs, moreover, both doctors and educators seem to be in line with the mainstream of public opinion. Neither of these criteria is well suited for implementation through the market mechanism without any form of public regulation, because it would be inconsistent with the objective of profit maximization. This may be at least a partial explanation why governments are so heavily involved in the provision of health and education.

A common argument in political debates about provision of services like health and education is that private production could lead to too low quality, since competition among private producers may cause them to cut costs in such a way that quality will fall below its socially optimal level. The reduction of production costs in itself is obviously a good thing and may reflect the competitive pressure to develop new and better technologies of production; on the other hand the desire to reduce labour costs in particular may lead to a lower quality level. This problem has been analyzed in an important contribution by Hart, Shleifer and Vishny (1997), to which I will return below.

One aspect of quality that is missing from this analysis – or at least is treated very implicitly – is the problem of service reliability. For you as a consumer it is always a good thing to have a permanent supplier to serve you, particularly if the good in question is a individualized service where quality depends on the supplier's knowledge about your personal characteristics. For most ordinary consumer goods this is a matter of little importance. If your hairdresser or local pizza supplier decides to close down or goes

2. Blomquist and Christiansen (1995, 1998) view the public provision of private goods as means to overcome the informational problems that restrict redistribution through the tax system.

bankrupt, it does not seriously upset your life. The situation is likely to be different with your school, hospital or retirement home³. In some areas of life we might attach some value to *institutional stability*, and this may more easily be ensured by government than by private ownership.

So far I have been mainly concerned with the choice between private and public production and ownership. But public provision – of health services, education and communication services – does not necessarily imply public production. The government can provide the service in question by paying a private contractor to produce it. The choice between public production and contracting out will be considered further below.

Some questions about costs

Some textbook discussions of public utility pricing assume, sometimes without stating it explicitly, that costs are independent of organizational form. This is an extremely strong assumption in the present context. We usually think of cost functions as having been derived from a process of cost minimization for every level of output, and of the set of technologies that enter into this process to be selected from considerations of technological efficiency. Both of these assumptions are in turn based on the assumption of profit maximization, which is the objective of the firm that conforms to the private interests of the owners. With private economic objectives being supplanted by social objectives, it can no longer be taken for granted that the managers of the public firm will be

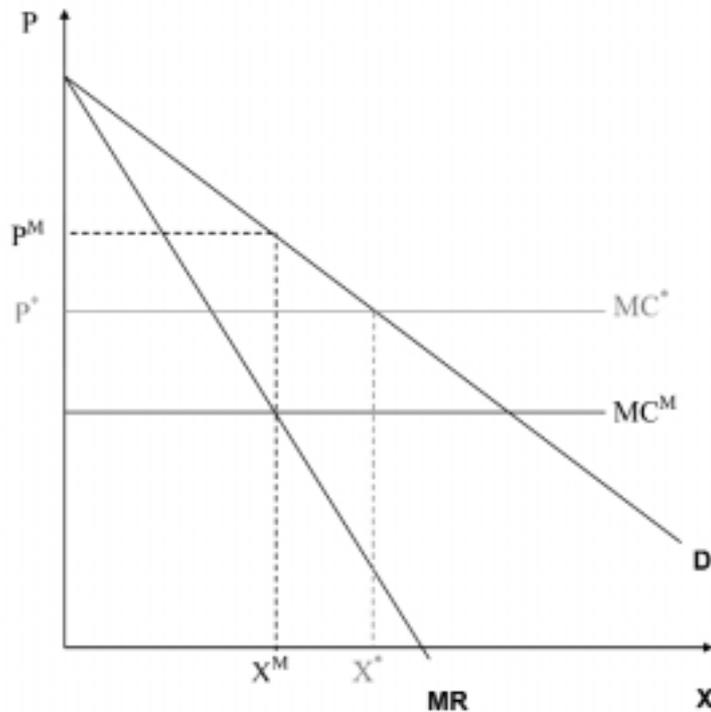
motivated to operate at minimum costs.

Even more implicitly, standard expositions assume that the nature of the product in qualitative terms is independent of the way in which production has been organized. This too is an objectionable assumption, particularly since what the government produces are often services where the quality of the product is difficult to separate entirely from the technology of production. Moreover, as I have already pointed out, the range of products offered is also likely to depend on the nature of ownership and the objectives of the firm or organization.

Leaving these aspects of the problem aside for the moment, let us reconsider the choice between private and public production in the presence of cost inefficiencies with public production. Figure 1 shows a case with constant marginal costs and declining average cost (not shown in the diagram). With a private monopoly, the price-output combination would be (P^M, X^M) . Public ownership raises unit costs to the level MC^* . As the curves have been drawn, consumers' surplus with marginal cost pricing is still higher than under the private monopoly regime. However, it is easy to see that by changing slightly the positions of the cost and demand curves, the public sector optimal output X^* could actually end up to the left of X^M . The price P^* is «right» in that it corresponds to the marginal cost of production, but the marginal cost in the public firm reflects a cost inefficiency. Figure 1 leaves fixed costs out of the picture, but the problems underlying the high marginal costs may also cause the fixed costs – which are not fixed in the long run – to be higher than they

3. Shleifer (1998, p. 139) mentions private nursing homes as an example of institutions where “consumers can switch suppliers if they are dissatisfied with the service.” He clearly has a point, but the argument seems to me to underestimate the cost of switching and consequently the importance of institutional stability.

Figure 1.
Marginal cost pricing with public cost inefficiency



otherwise would have been. If so, it may no longer be true that public ownership makes it more likely that a positive level of output is better than no output at all. Summing up, it is not clear that the market failure associated with a private monopoly is improved by public ownership. Instead, the result may be that we substitute a cost inefficiency for a price inefficiency.

This is a simplified picture in at least two ways. First, a private monopoly might not have the strongest incentives to produce at minimum cost either. As Hicks (1935) remarked, «the best of all monopoly profits is a quiet life». Second, if we believe that the public firm is not run in a way which is cost efficient, why should we believe that it sets the right price, relative to that inefficiency? Still, at least for the moment, I shall stick to

the simplified picture and concentrate on inefficiency in public production as the main counter-argument to the position taken in my quotation from James Meade.

Is public production inefficiency an established fact? Time and space constraints do not allow me to go into the extensive empirical literature that exists in this area, but let me offer some fairly general remarks. One of the difficulties about comparing efficiency in private and public firms lies in finding lines of production where the nature and quality of the product is the same, like garbage collection (which, for that reason, may be a relatively over-researched area). In the empirical comparisons that have been made, there seems to be a majority of studies that show a cost disadvantage for public firms. However, it is sometimes open to

question whether this disadvantage is a true inefficiency, or whether it is the result of particular constraints that the public firms face and that may reflect socially desirable aspects of their mode of operation. Moreover, there are enough counterexamples to the superiority of private production that should make us wary of sweeping empirical generalizations. Still, it remains a fact that the public sector is much more sheltered against the forces of competition than the private sector, so that the economic incentives to cost minimization are much weaker. This in itself should provide sufficient motivation for theoretical studies of what these incentives in fact are.

It ought perhaps to be emphasized that while this discussion refers explicitly to the case of decreasing average cost, the following analysis of the possible causes of public cost inefficiency has a wider applicability.

Costs and incentives.

To a large extent, public production is production of services. The crucial input in service production is labour. Unlike some popular applications of the theory of production, cost minimization in service production is not a problem that can easily be formulated as a programming problem and solved satisfactorily using numerical methods. The reason is obviously that the supply of human effort, both in quantitative and qualitative terms, is essential for the nature of the output itself, and workers in service organizations are not inputs that are perfectly controlled by the top management. This is true both for private and public production, but may create special problems in the public sector.

A framework for studying this kind of problem is principal-agent theory. A principal has certain objectives that he wishes to

pursue, and he enters into a contract with an agent to carry out the task for him. However, the actions of the agent are imperfectly observable by the principal, who can only observe them indirectly and imperfectly by the outcome of the process of which the agent is in charge. The outcome is a function of the agent's effort and some exogenous factors; the principal's problem when it comes to the design of incentives is that he cannot in general observe how important the agent's effort has been for the actual outcome. This is obviously a very general formulation, and particular versions of the model have been applied extensively to business problems. But it is also a fruitful way of thinking about a number of problems in the public sector. In some applications one could think of the principal as being a government ministry and the agent as being the manager of a public utility; in other applications the manager could be the principal with his division heads as the agents and so on. Tirole (1994) and Dixit (2002) are excellent discussions of the relevance and application of the theory to problems of the public sector.

How should the reward scheme of the agent be designed? On the one hand one has to take account of the agent's incentive to perform in the interests of the principal; this calls for tying the reward closely to the observable outcome. On the other hand, if the agent is risk averse, this exposes his income very much to exogenous forces over which he has no control. He would therefore like to exchange some of his outcome-based income for a fixed income, even if the latter were lower than the expected value of the former. Under certain assumptions (Holmström and Milgrom 1987) it can be shown that a linear reward scheme, in which a fixed income component is combined with a variable component that depends on the

outcome, is optimal. Following Dixit (2002) I shall refer to the coefficient of the variable component as the bonus coefficient. A natural interpretation is then that the bonus coefficient, which corresponds to a fraction of the outcome, ensures that the agent has the right incentives at the margin, while the fixed component provides some insurance against the risk involved and ensures the satisfaction of the agent's participation constraint⁴. However, one also has to take account of the principal's risk aversion; if this is high, it calls for a higher bonus coefficient, since this implies that more of the risk is borne by the agent.

In order to implement such an incentive scheme it is obviously of central importance to define the outcome to which the bonus coefficient is to be applied. In a business context the reward of the company's CEO could be tied to profits or changes in stock market value or other indicators of company performance that could easily be measured and monitored. The top management could in turn devise schemes for the lower echelons of the staff that are related to division output or profitability. However, the typical public sector organization or firm – «agency» for short – has objectives that cannot in a simple manner be brought down to simple scalar measures of performance. Garbage removal and power supply may be examples of cases where performance measures are not too difficult to invent; with collective transportation, schools, universities, hospitals and cultural institutions it becomes considerably more complicated. From the point of view of welfare economics, we might suggest that the

correct indicator of performance should be the agency's contribution to social welfare or the social surplus. This is a good point at which to start thinking about the problem, but for several reasons these are not indicators that can easily be made both measurable and politically and socially acceptable.

A complex set of objectives makes it difficult for the principals to monitor the agency's performance. This enables the employees of the agency to pursue their own goals to a larger extent than would have been possible in a private firm with a simple objective function⁵. These goals may be selfish – enjoying leisure and consumption on the job. But they need not be. Many employees of government agencies like hospitals and child care institutions differ from their political and bureaucratic principals in having a higher estimate of the benefit of the agency's output. Indeed, the reason that they value that output so highly may have been the reason why they chose to work for the agency in the first place. The result may be that they tend to promote quality to a higher extent than their principals may desire, with the unavoidable result that costs increase.

Following Dixit (2002), two reasons for these complexities are that public agencies may have multiple principals or they may have multiple tasks or objectives. These reasons are closely related to each other, so I concentrate on the problem of multiple objectives. A hospital should take care of people who are in immediate need of medical assistance, e.g. as a result of accidents. It

4. The participation constraint requires that the agent receives some minimum expected utility in order to stay with the organization.

5. Niskanen's (1971) theory of bureaucracy can be seen as a principal-agent model for the public sector, while Baumol's (1958) sales maximization model is a theory that is formally similar to Niskanen's, although the application is to a business firm which is imperfectly controlled by its owners.

should also treat the sick who need operations or other kinds of treatment for regular illnesses, and take care of those who, following treatment, are unable to take care of themselves. Health authorities in many countries have experimented with incentive schemes that are designed to make hospitals perform better, typically by introducing bonus coefficients to encourage better performance. However, since each type of bonus has to be tied to one particular performance measure it tends to distort local priorities in favour of the activity being measured, sometimes with unfortunate effects for other aspects of performance. E.g., if there is a bonus for each operation performed, this alone creates an incentive to perform as many operations as possible (which may be good) while limiting the number of days each patient is allowed to stay in the hospital (which may be bad). Many European university departments of economics are introducing rewards for international publications; this may increase the overall quality of the research done (which is good), but lead to a neglect of national economic problems (which may be bad)⁶. Evaluations of performance that focus on only one aspect of output (the number of operations or international publications) may conclude that the bonus system has improved efficiency, while a broader evaluation may conclude in the opposite direction.

There are two main solutions to this type of problem. When confronted with a set of imperfect and distorted incentives, many economists would react instinctively by recommending the agency: Get the incentives right! Another solution is to transfer the

activity in question to the private sector, either by outsourcing or by outright privatization, combined with regulation.

Getting public sector incentives right

In the competitive model, the justification of the assumption of profit maximization is twofold. First, since it is in the interests of the owners to achieve a maximum of profits or present value, this is what they will try to motivate the managers to do. Second, even if owners or managers were to have other objectives than profit maximization, the discipline imposed by a competitive environment will force them to maximize profits, simply in the interests of economic survival. Similarly, we could think of a government that wishes to motivate its agencies to maximize its contribution to social welfare or the social surplus, either to provide them with explicit incentives to do so, or to expose them to outside competition. Both strategies have been tried in recent attempts to improve on the performance of the public sector.

I have already discussed the difficulties associated with providing efficient bonus incentives in organizations characterized by multiple principals and multiple objectives. It may not in fact be administratively possible to design an incentive scheme that imposes the right bonus for any conceivable action that one may wish to reward. The set of activities may simply be too large for this to be possible, or some activities have results that are too hard to measure for bonuses to be feasible.

Nevertheless, there is a frequently voiced concern that the reward structures for public

6. See Stigler (1963) for an amusing and thought-provoking fantasy about incentive schemes in a university context.

employees provide too weak incentives; people are mainly on fixed salaries, and there are no explicit bonuses. However, this impression may be a bit superficial. A bureaucrat who works hard and does a good job for the agency may be promoted to a better paid and more attractive position, either in the agency itself or in some other part of the public sector⁷, and the opportunity to climb the bureaucratic ladder may in many cases be a close substitute for explicit bonus payments. The design of career opportunities may be a better strategy for the improvement of incentives than the introduction of bonus payments, although the relative benefit of each type of incentive is likely to vary considerably among agencies.

Principal-agent theory usually assumes that the agent is motivated solely by his material reward. As a general theory of motivation, this is obviously unrealistic. Many employees of public agencies – schools, hospitals, environmental agencies, the courts – see themselves as working for organizations with an idealistic objective, and this is at least in part both their reward and their motivation. When this intrinsic motivation to perform in the interests of the principal is strong, there may actually be a cost associated with the introduction of explicit material incentives in the agency. The agents may feel that they are not expected to perform well unless they are explicitly paid for it, and this weakens their intrinsic motivation.⁸

The motivation to work for the agency's objectives may not only be the result of the

exogenously given preferences of the workers. The leadership of the agency may consciously try to foster a corporate spirit, an *esprit de corps*, among the workers, leading them to identify with the social goals of the agency and taking pride in their work. This kind of motivation is difficult to analyze by means of the standard analytical tools of economics. Nevertheless, there is reason to believe that it is of great practical importance, and that this kind of collective motivation may explain why many public sector agencies actually perform well in spite of apparently weak incentive structures in the usual economic sense.

Exposure of public agencies to outside competition is an alternative route towards encouraging cost efficiency on the part of public agencies. This may work well in cases where the nature of the output is reasonably well defined; studies of competition between public agencies and private firms in areas like garbage removal tend to show that competition has a beneficial effect on efficiency. Where the nature of the output is more complex the situation is more difficult. The private firm may adopt a strategy of cream-skimming, whereby it concentrates on the more profitable segments of the market. The principals in the public sector may react to this by regulating the private competitor, but if the number of regulations becomes too large, the incentives of the private firm may weaken so much that it becomes more like the public agency (Dixit 2002). Alternatively, if the private competitor is not constrained in

7. Or indeed in the private sector. In Norway and presumably also in other countries, young lawyers with some years' working experience in the Ministry of Justice may thereby become very attractive for private law firms or other private companies. Economists who have done well in junior positions in the Ministry of Energy, have gone on to fill top positions in private oil companies.

8. See the discussion of this effect in Frey (1997), who applies it to the problem of constitutional design at the political level. A more formal treatment with emphasis on individual incentives in organizations and personal relationships is in Bénabou and Tirole (2002).

this way, the public agency may feel forced to become more like the private firm in order to survive. It is not difficult to find examples of both of these outcomes. Exposure to private competition may work well in some cases, but it is not a universal solution to the ills of the public sector.

A major difference between private firms and public agencies lies in the exposure of the former to the risk of bankruptcy. It is true that public agencies may also face a risk of being closed down, but on the whole public agencies and their employees are more protected from this particular risk than private firms are⁹. This could be positive when institutional stability is important, as discussed above. But it clearly also has a negative side, to the extent that public employees feel that there is no connection between the effort that they supply to the agency and the safety of their job. Combining the concern for institutional stability with adequate incentives on the part of the workers is a major challenge for public sector reformers.¹⁰

Contracting out to the private sector

There are no precisely drawn limits between the alternative ways in which the public sector can utilize the private sector for its own purposes. Exposing public agencies to outside competition from private suppliers is a form of contracting out. However, there is a point in considering separately the case

where contracting out is seen as a direct alternative to in-house production. When is it rational for the public sector to abandon ownership and instead limit its role to that of a provider of public services?

In many cases of public supply, the essential feature about the publicness is not production itself, but the *provision* of the good. It is important that basic education is provided for free, but this does not necessarily imply that schools should be owned by the government and that teachers should be civil servants. Instead, schools could be operated by private organizations under a contract with the government who would cover the cost of providing children with education. The same could be said for hospitals, public transportation, nursing homes, prisons and (once again) garbage removal.

The basic arguments for contracting out are, first, that it is likely to lead to lower costs, since the private supplier has a clearer interest in cutting costs than the public agency has. On the other hand, cost reduction can in some cases be expected to have an adverse effect on the quality of the service. Of course one could argue that with a sufficiently detailed contract, the government could impose very precise quality specifications on the private contractor. But in many cases quality is such a complex concept that contracts must necessarily be incomplete, leaving the private contractor with considerable leeway in choosing the

9. Kornai (1980) coined the term «soft budget constraint» to characterize the situation of public firms in the former socialist countries, in which decreased sales or increased costs were expected to be followed by increased transfers from the government.

10. From the point of view of social design it might be natural to think that if the bankruptcy threat to job security is removed, other aspects of job security ought to receive less emphasis in the labour contract. Someone who works for an agency where there is no risk of bankruptcy ought perhaps to run a higher risk of being fired for unsatisfactory performance than one who works for a private firm. In reality, of course, the opposite is the case. A possible explanation for this is that the absence of bankruptcy risk means a strengthening of workers' bargaining power, which can be used to obtain greater job security.

cost-quality tradeoff. The question is then when this is a net disadvantage from the point of view of the government, and when the gains from private production outweigh the possible disadvantage.

In the analysis of Hart, Shleifer and Vishny (1997) the main benefit of contracting out is not that one reaps the benefit of competition, but that the private contractor has a stronger interest than the public agency both in improving quality and in cutting costs. In their model the public manager devotes too little effort (relative to the social optimum) both to cost reduction and quality innovation. Under private ownership, by contrast, where the manager collects more of the surplus from the operation, the effort devoted to cost reduction is too high while the effort devoted to quality innovation is too low, although higher than under public ownership. This makes the choice between the two organizational forms non-trivial; it depends on which kind of innovation is more important from a social point of view. The central insight that is derived from the theoretical analysis is that public ownership is likely to be better than private ownership if the adverse effect of cost reduction on quality is large. But this alone is not decisive unless it is also true either that quality improvement is unimportant, or that government employees have stronger incentives for quality improvement.

Applying their analysis to a number of examples, the authors conclude that the case for government production is strong in areas like foreign policy, police, the armed forces and (probably) prisons, while the case for contracting out is strong in the cases of garbage collection, arms production and (probably) schools. It is also worth noting that the first group of cases may be such that it is difficult for the government to know in advance exactly what it wants to be done,

while this is less true for the cases in the second group. The more difficult it is to make the terms in the contract with the private producer reasonably precise, the stronger is the case for government ownership; this has also been emphasized by Shleifer (1998).

The last point may be related to my previous remarks about the development of an *esprit de corps* in a public agency. A shared understanding among the agency's employees of its social goals, responsibilities and professional code of conduct may act as a substitute for a complex contract that sets out in detail what should be done in different situations that the private contractor might face.

Private and public goods

The application of incentive theory to the choice between private and public provision appears either to neglect the distinction between private and public goods or to assume more or less implicitly that it is limited to the case of private goods. But this is actually not a correct understanding of the literature. Even when free-rider incentives create a strong case for government provision of the public good, this does not necessarily imply that the good should be publicly produced. In the original Samuelson (1954) formulation of the theory of public goods, he makes a strong case for public provision, but not for public production. So governments can provide public goods to consumers while contracting with a private agent to produce them.

Obviously, however, public goods are of many types, and the possibility to write a contract with a private agent to produce them varies enormously between types. It is fairly easy to draw up a contract whereby a private firm commits itself to lay out a public

park, to be paid for by the government and put at the disposal of citizens free of charge. It is considerably more difficult to write a contract to provide law and order in a city centre. This conclusion is therefore perfectly in line with that of the previous section. When it comes to the choice between private and public production, the main considerations are similar whether the good in question is private or public, although the case for public provision is stronger in the latter case.¹¹

Controlling externalities

Is the existence of externalities a strong argument for public interference with the market mechanism? A reader of the older literature, with its emphasis on quaint examples like the interaction of apple-growing and bee-keeping, might be inclined to say no. With the discoveries in recent decades of the threat to the natural environment from economic activity, however, attitudes have changed, both among economists and the general public. Nevertheless, there are those who hold that externality problems are best solved by private agents with only minimal interference from the government.

The argument is as follows. Externalities come from the uncompensated effects of some agents' actions on the costs or utilities of others. But if property rights are well defined, the fact that the market equilibrium is not efficient holds out the promise of mutual gains from negotiations to alter it. If there are no obstacles to negotiations, competitive equilibrium plus side transactions by the parties affected by the externality should lead to efficiency after all. It is natural to refer

to this as the Coase (1960) perspective, although, as pointed out e.g. by Stiglitz (1994), Coase was careful to emphasize that the result holds in the strict version only if negotiations are costless.

Even if there is a cost to negotiations, however, we would expect a number of externalities to be solved on a voluntary basis simply because the costs of negotiation are fairly small. But it is reasonable to assume that the costs of negotiation increase with the number of affected parties. Even when the number of parties to the negotiation is fairly small – say twenty or fifty – it is hard to imagine that the negotiation can be carried out without some kind of broker or specialized organization. The formation of such organizations has in fact been observed in a number of studies of private solutions to externality problems like «the tragedy of the commons». If there are many negotiations of the same kind going on in the economy, there will be increasing returns from having just one or a few brokers who specialize in this kind of negotiation. Moreover, if the number of parties involved becomes very large, the government – either the central government or local governments – becomes a natural broker for «negotiations» about externality issues. Then one can view the results of environmental legislation, regulations or Pigouvian taxes as outcomes of negotiation processes in which the government acts as a broker between the parties affected by the externality. This could be seen as one justification of the analysis of environmental policy within the framework of welfare maximization (Sandmo 1990). This line of reasoning provides arguments for an active role for the government in this area, although one that relies heavily on private incentives.

11. Besley and Ghatak (2001) discuss the role of the voluntary sector in public goods provision and financing.

Concluding comments

The division of labour between the public and private sector is being subject to critical scrutiny by the economics profession. The trend of the times is to advocate privatization, stronger incentives for private agents who work for the government and general withdrawal of the government from many of the areas where it has traditionally played an important role. On the whole I believe that this trend could generate a number of benefits for society, although their magnitude depends on the initial position of the country in question with respect to the size of its public sector. Presumably, the gain from privatization and deregulation should be larger, *ceteris paribus*, the more extensive public ownership and government regulations are to begin with.

But although the trend may be beneficial, one should proceed with some caution. It is easy to listen to the song of Big Government and decide that the Market is the winner. But a better way is to apply in reverse the principle that I mentioned in the Introduction: If you observe a significant area of policy failure, that in itself provides a case for deregulation or privatization. But before making your final recommendation to policy makers, you should also be able to argue convincingly that the private sector is likely to handle the problems involved in a better way.

References.

- Arrow, K. J., 1971, «Political and economic evaluation of social effects and externalities,» in M. Intriligator (ed), *Frontiers of Quantitative Economics*, North-Holland, Amsterdam.
- Baumol, W. J., 1958, «On the theory of oligopoly,» *Economica*, 25: 187-198.
- Bénabou, R. and J. Tirole, 2002, «Intrinsic and extrinsic motivation,» unpublished, Princeton University.
- Besley, T. and M. Ghatak, 2001, «Government versus private ownership of public goods,» *Quarterly Journal of Economics*, 116: 1343-1372.
- Blomquist, S. and V. Christiansen, 1995, «Public provision of private goods as a redistributive device in an optimum income tax model,» *Scandinavian Journal of Economics*, 97: 547-567.
- Blomquist, S. and V. Christiansen, 1998, «Topping up or opting out? The optimal design of public provision schemes,» *International Economic Review*, 39: 399-411.
- Boiteux, M., 1956, «Sur la gestion des monopoles publics astreint à l'équilibre budgétaire,» *Econometrica*, 24: 22-40. Translated as «On the management of public monopolies subject to budgetary constraints,» *Journal of Economic Theory*, 1971, 3: 219-240.
- Coase, R. H., 1960, «The problem of social cost,» *Journal of Law and Economics*, 3: 1-44.
- Dixit, A., 2002, «Incentives and organizations in the public sector: An interpretive review,» *Journal of Human Resources*, 37: 696-727.
- Dupuit, J., 1849, «De l'influence des péages sur l'utilité des voies de communication,» *Annales des ponts et chaussées*. Translated in part as «On tolls and transport charges,» *International Economic Papers*, no. 11, Macmillan, London, 1962.
- Elster, J., 1992, *Local Justice*, Russell Sage Foundation, New York.
- Frey, B., 1997, «A constitution for knaves crowds out civic virtues,» *Economic Journal*, 107: 1043-1053.
- Hart, O., A. Shleifer and R. W. Vishny, 1997, «The proper scope of government: Theory and an application to prisons,» *Quarterly Journal of Economics*, 112: 1127-1161.
- Hicks, J. R., 1935, «Annual survey of economic theory: The theory of monopoly,» *Econometrica*, 3: 1-20.
- Holmström, B. and P. Milgrom, 1987, «Aggregation and linearity in the provision of intertemporal incentives,» *Econometrica*, 55: 303-328.
- Hotelling, H., 1938, «The general welfare in relation to problems of taxation and of railway and utility rates,» *Econometrica*, 6: 242-269.
- Johansen, L., 1965, *Public Economics*, North-Holland, Amsterdam.
- Kornai, J., 1980, *Economics of Shortage*, North-Holland, Amsterdam.
- Meade, J. E. and J. M. Fleming, 1944, «Price and output policy of state enterprise: a symposium,» *Economic Journal*, 44: 321-339.
- Meggison, W. L. and J. M. Netter, 2001, «From state to market: A survey of empirical studies on

- privatization,» *Journal of Economic Literature*, 39: 321-389.
- Musgrave, R. A. and A. Peacock (eds), 1958, *Classics in the Theory of Public Finance*, Macmillan, London.
- Niskanen, W. A. Jr., 1971, *Bureaucracy and Representative Government*, Aldine, Chicago.
- Samuelson, P. A., 1954, «The pure theory of public expenditure,» *Review of Economics and Statistics*, 36: 387-389.
- Sandmo, A., 2000, *The Public Economics of the Environment*, Oxford University Press, Oxford.
- Shleifer, A., 1998, «State versus private ownership,» *Journal of Economic Perspectives*, 12 (no. 4): 133-150.
- Stigler, G. J., 1963, «An academic episode,» G. J. Stigler, *The Intellectual and the Market Place and Other Essays*, Collier-Macmillan, London.
- Stiglitz, J. E., 1994, *Whither Socialism?* The MIT Press, Cambridge, Mass..
- Tirole, J., 1994, «The internal organization of government,» *Oxford Economic Papers*, 46: 1-29.
- Tobin, J., 1970, «On limiting the domain of inequality,» *Journal of Law and Economics*, 13: 263-277.
- Vickrey, W., 1948, «Some objections to marginal-cost pricing,» *Journal of Political Economy*, 56: 218-238.
- Wicksell, K., 1896, *Finanztheoretische Untersuchungen*, Gustav Fischer, Jena.