

Chapter 6. The Welfare State

The discussion in the previous chapter showed how forming an economic partnership or corporation can allow individuals to overcome some of the collective action problems that the system of property rights and voluntary exchange fail to eliminate. However, the class of market failures that corporations are able to resolve is somewhat narrow. Most of them are concentrated in the area of organizing joint investment and labour strategies. There are a number of other areas in which the corporate organizational form lacks the tools needed to eliminate market failure. There are an additional set of circumstances in which the corporate form generates market failures of its own. As a result, there will still be demand for institutional arrangements that curtail and supplement market patterns. Most of these institutions are constructed and administered under the auspices of the state.

The primary limitations of the corporation stem from the fact that it is able to deploy only a limited range of sanctions. First and foremost, the firm is not able to directly sanction individuals who are not among its members. Even when it comes to dealing with its own employees, the firm has only limited means of exercising control. It can sanction them through manipulation of salary and promotional rewards, but that is about all. And since employees can avoid any particular rule or sanction by quitting, the firm cannot push too hard. As a result, corporations do not always have the means to control free rider problems, either internally or for society generally. More importantly, however, they often do not have an interest in doing so. Firms are created so that agents may engage in a cooperative pursuit of economic rewards. They are able to pursue this goal without concern for the broader social consequences of their actions because, in a properly structured market, this "private" interest coincides with the broader public interest. But in many cases where the market is not properly structured, profit-seeking on the part of firms generates negative externalities. In this case, the private interest of the firm does not fully coincide with the public interest.

Ultimately, the capacity of the corporate organizational form to resolve market failure is limited by the fact that it produces only *voluntary organizations*. Since people cannot be forced to join a voluntary organization, and since they are always free to quit, voluntary organizations exercise very little power over their members (in the larger scheme of things), and their

membership is usually less than universal. Since every market failure amounts to an unresolved free rider problem, it is evident that the corporations will only be able to do so much. A voluntary organization is not able to eliminate all free riders, simply because people can often free ride by not joining that organization. As a result, within every capitalist society, there remains an important place for an institution that is able to dispose over a broader range of sanctions, and whose membership includes the entire population of a territorially defined region – in short, an institution in which membership is *universal* and *compulsory*. This is the state. Only an institution of this type has the power, in principle, to eliminate any and all free riders.

The sociologist Max Weber defined the modern state as simply an institution that exercises a monopoly over the use of force in a particular region (1978, 314). The two most important powers that this institutional structure confers upon the state are the power to *regulate*, and the power to *tax*. The former allows the state to restructure the terms under which competitive exchange will occur. The latter allows the state to "force" exchanges, and thus to make individuals pay for benefits they receive, or costs they impose, in cases where no effective system of property rights can be institutionalized. By virtue of these two powers, the state is capable of engaging in a wide range of economic activities that no other institution in the society is able to effectively manage. In particular, the state is able to provide goods and services that no individual or corporation would be willing to provide. As a result, the state has emerged as the single most important economic actor in our society, precisely because of its capacity to work alongside markets and corporations, correcting market failures, and improving the overall efficiency of the economy.

In order to see where this claim fits into the larger picture being developed here, it is helpful to recall the way that the argument has unfolded to this point. The discussion began with an outline of a moral economy, in which the production and distribution systems of the society are all centrally organized and hierarchically integrated in the state.

1. It was then shown that "decentralization" of certain spheres of interaction, in particular, those in which individuals exchange goods and resources, could lead to significant improvements in efficiency.

2. This led to a proposal for the total decentralization of society, and a replacement of moral constraints by market interactions. This form of market utopianism received aid and

comfort from the invisible hand theorem, which said that complete decentralization of the economy would lead to a Pareto-optimal distribution of goods and resources.

3. By examining the assumptions needed to prove this theorem more carefully, it was found that this Pareto-optimum is impossible to achieve under "real-world" conditions. The ineliminability of market imperfections suggested that not all production processes should be decentralized. There may be efficiency gains to be had from retaining hierarchical or trust-based organizational forms.

4. Finally, it was argued that corporations are formed when an organizational hierarchy would perform better (i.e. be more effective at resolving collective action problems) than a collection of individuals engaging in market exchange.

It will now be shown that there are gains in efficiency to be had from not allowing all sectors of the economy to be integrated through "voluntary" relations and associations – whether these be markets or corporations. By virtue of its special powers of compulsion, the state is often a more efficient provider of certain types of goods and services.

6.1 From the liberal state to the welfare state

Even extreme libertarians recognize that the market exchange system depends, for its stability, upon it not being universal. In order for exchange to occur, individuals must be confident that everyone is going to live up to their side of the bargain. In order for this to be the case, there must be some kind of enforcement mechanism, used to punish those who fail to comply with the terms of the contracts that they enter into. In order for this enforcement mechanism to work properly, the services of these enforcers must not be just one more commodity that can be bought and sold at will. If it were, then anyone could avoid being punished for breaking a contract, simply by "paying off" the police, the judge or the jury. This is what we refer to as "corruption." Since the system of property rights only works if it is enforced, the central enforcement mechanism must not be something over which individuals can exercise property rights. If this constraint is not respected, the market system will simply collapse into the Hobbesian state of nature.

Thus a market economy, no matter how extensive the system of exchange, is never self-sustaining. It requires a state in order to maintain the basic set of institutions that constitute the market. Thus most libertarians admit the need for some kind of state. But they attempt to

minimize the significance of this concession by arguing that the state's only role should be to enforce the system of individual rights. This is sometimes referred to (confusingly) as the *liberal*, or (more perspicuously) the "nightwatchman" state (Nozick, 1973). Like a security guard on the night shift, all that the state should do, according to this conception, is keep an eye on the shop. Government should initiate no positive legislative measure or engage in any economic activities; it should merely ensure that private rights are respected, and that private contracts are enforced.

However, once this first concession is made – once the state is assigned the role of general enforcer of individual rights – it is very difficult to resist the line of argument that seeks to extend the activities of the state into other areas. To see how this argument goes, consider the following: What is the state doing, when it enforces the system of property rights? All it is doing is providing a service, in this case enforcement services. Why don't people just buy these services from other people? Because there is a collective action problem – without a system of *enforceable* property rights, exchanges cannot occur (or transaction costs would skyrocket). So when the state supplies enforcement services, it is responding to a particular type of market failure, in this case a failure in the market for enforcement (although this is a funny way of putting it, since markets would hardly exist without enforcement). The general point is that, when faced with a collective action problem, people develop an institution that will deliver the social good directly. Thus enforcement of rights is a social good that the state provides.

But why should enforcement be the only type of service that the state provides? This service solves one particular kind of collective action problem, and it puts into place a system of rights that solves many others. But what about those that remain? If the state is able to solve other collective action problems, why shouldn't it? If markets will fail to provide other goods and services at an efficient level, and the state can, there is an important efficiency argument that says the state should proceed to do so. Enforcing the system of rights is just one kind of social good that the state provides, no different in kind from building bridges and running schools. Libertarians arbitrarily select this public good as one that they endorse, and oppose the others. But the same logic which dictates that the state should enforce the system of rights also dictates that the state should provide other goods and services whenever it is capable of delivering them more efficiently than the private sector.

The primary question, then, becomes not *whether* the state should be involved in the economy, but precisely *when* and *where*. The answer that has emerged over the course of the 20th

century in most liberal democracies is roughly the same as that given for the corporation. The state should assume a particular economic role just in case doing so generates a more efficient outcome than would otherwise be achieved in the private sector, i.e. through markets and corporations. Thus the state is given the task of supplying goods and services that markets will not provide, or that they will not provide at an efficient level. Again, an equilibrium will be achieved when the benefits of state provision of goods and services become outweighed by the costs associated with its additional expansion. When the state assumes an economic role of this type, it is referred to as the **welfare state** (or the social-democratic state).

The logic of having such a role assigned to the state is best illustrated through an example (a more complete survey of state economic activities will follow in the next section). One of the earliest and most comprehensive "welfare state" programs to be implemented in capitalist societies has been the public education system. Universal and mandatory primary education was introduced in England in the late 19th century, and was motivated precisely by the concern that, when left to their own devices, individuals would select an inefficiently low level of education for themselves. In part, this is due to the fact that education is an "experience good" – it is only once you are educated that you can fully appreciate the value of education. This fact led John Stuart Mill, as early as 1848, to argue that markets would tend to underproduce education (1848 [2004], 869).

The other major reason that markets for education will tend to fail is that education generates significant positive externalities. A contemporary example of this phenomenon can be seen in debates over job training programs. Consider the game shown in Figure 6.1, which illustrates a somewhat simplified, but highly typical situation. Two rival firms are in a position where they can each gain an extra \$2000 in output by training one of their workers. Since it costs \$1000 to train a worker, the firm's net gain would be \$1000. However, instead of training a worker, the firm can instead hire a headhunter to find a similarly trained worker elsewhere. Doing this would only cost \$500, and so the firm would be looking at a net gain of \$1500. Naturally, if the firm decides to train a worker, but then some other company's headhunter hires this worker away, the \$1000 invested in that worker's training will be lost. As a result, the firm is better off hiring a headhunter than it is training workers. Since every firm is in the same situation, no workers will be trained – thereby guaranteeing low productivity in that sector and useless expenditures on headhunters' fees.

		Firm 2	
		Headhunt	Train
Firm 1	Headhunt	(-5,-5)	(15, -10)
	Train	(-10, 15)	(10,10)

6.1 Whether to train

The problem is that the firm can train its workers, but it cannot exercise a property right over the skills that any worker acquires. As a result, the firm's training expenditure generates a positive externality that it may or may not recapture. This generates the temptation to free ride off the training efforts of other firms (i.e. to headhunt). To the employees, training increases their value on the market, tempting them to leave the firm. There is no way around this, other than having people sign very rigid contracts (which would in turn reduce labour-market flexibility). This is why the private sector, left to its own devices, will tend to train its workers at an inefficiently low level (particularly with regard to highly general skills, which are the most portable).

The state, however, has the power to eliminate this collective action problem. Notice that both companies would be willing to spend the money on training, as long as their workers are not going to be headhunted. If the state comes in, imposes a \$1000 fee per worker on every corporation in the sector, then supplies training to these workers, the collective action problem is eliminated. Since all workers will be trained, no firm has to worry about its workers being headhunted. And since everyone must pay the tax, the free-rider incentive is eliminated (in fact, it is reversed – as long as you're paying for the training program, you might as well make sure some of your workers attend). Note also that the taxes are not "passed along" by the corporation to the consumer, they are paid for by efficiency gains that each corporation enjoys.

(This explains, incidentally, why the public education system is often caught between the private sector, which is constantly pushing for an emphasis on *basic skills*, and students, who are

often more interested in vocational training. Corporations want to hire employees with generalizable skills because these are the ones that they cannot provide without attracting free riding from rival firms. Very specific technical skills tend not to be as readily transferable from one workplace to another, and so corporations can teach new employees these skills without having to worry about them leaving. This is why, for instance, no corporation will pay to train its employees to type, or do basic math, but they will train them to use specific software applications, or manage proprietary manufacturing processes.)

What goes for corporations goes for society at large. Training workers creates a large positive externality. The benefits of primary education tend to be even more widely dispersed. Being literate and numerate benefits the individual, but it also benefits all those with whom she must interact. (We get so used to dealing with educated people that it is easy to forget how much trouble their ability to communicate through writing saves us, e.g. it allows us to post signs with the assumption they will understand them, it allows us to exchange letters and email, etc.) But literacy does not occur automatically – this is why most countries run campaigns to increase literacy every few years. People are inconvenienced by their own inability to read, but this is nothing compared to how much they inconvenience others. But since they don't bear this cost, they have diminished incentive to reduce it. This is why they must often be offered free or heavily subsidized education in order for society at large to achieve an efficient level of general education.

6.2 Economic activities of the state

The provision of educational services is an example of a classic welfare-state activity. In this case, the state compensates for a market failure by providing a particular kind of commodity directly to its citizens. Instead of charging them for the amount that they consume, revenue to support provision of the good is secured through taxation. Thus the costs are imposed upon *all* those who benefit, not just those over whom an effective property right can be exercised. In cases like this, the state is engaging in a positive economic activity. Like a corporation, it is hiring inputs, organizing production, and delivering a service. The key difference is just that it does not rely upon the mechanism of voluntary exchange in order to organize several key transactions. The goods are not paid for at the point of delivery, but in a more indirect way, through the

taxation system. In a sense, the administrative hierarchy is extended downward to include the consumer within the organizational structure.

Positive economic activities of this type can be organized into three general categories:

1. Public and quasi-public goods. **Public goods** are often defined, following Paul Samuelson (1954), as commodities that are nonexcludable and nonrival. Nonrivalry means that one person's consumption does not decrease anyone else's consumption. Nonexcludability means that if the good is provided for one person, no one else can be excluded from consuming it (within perhaps a certain geographic region). In some cases this will be because it is physically impossible to stop people from getting a share, but in other cases it will just be impractical to do so. If the cost of charging people who have benefited from a good exceeds the amount that could be collected, the good will be, for all practical purposes, nonexcludable.

There are very few examples of pure public goods. What makes them an important special case, however, is that no one has any reason at all to buy them – an individual can get just as much by free-riding as she can by purchasing. As a result, markets will not provide these goods *at all*. Examples of goods that come close to being entirely public are national defence, basic infrastructure (like lighthouses, traffic signals, some roads, etc.), parks, sewage disposal and pest control. A good example of this, from Joe Stevens's *The Economics of Collective Choice* (1993) is mosquito control:

If I decide to spray the breeding grounds in the ditch near my house, the service that I provide has strong public good characteristics. It is nonrival; my dead mosquito is your dead mosquito. High transactions costs, however, would likely discourage me from identifying other beneficiaries and creating a market by which I could tap the gains from my actions. My neighbors might not know that I have sprayed, they might not believe me when I tell them so, or we might simply disagree on whether my spraying has helped them. Thus, spraying is basically nonexcludable because of its physical nature and high transactions costs. The market would underproduce mosquito control because of public goods and externality problems; a more efficient level of output would require collective action. Mosquito control thus seems to be a prime candidate for public provision, and

indeed it has been a recognized function of local governments for many years (Stevens 1993, 69).

As this example suggests, pure public goods are just a limit case of a more general class. Any good that generates a significant positive externality will tend to be underproduced by markets, because its price will not reflect its true value to those who enjoy it. As a result, the price system will send producers the wrong signal, suggesting that "society" wants less of these goods than it actually does. As a result, such goods will be produced at inefficiently low levels. Public goods, which are almost pure externality, will get produced at the most inefficient levels – perhaps not at all. However, there are a large number of other goods that will be produced, but just not in sufficient quantities to meet people's real needs. In these cases, governments may decide not provide the good directly, but merely to subsidize its production. When they do so, they are in effect "correcting" the price so that it sends a more accurate signal to producers and consumers. When people take mass transit, for example, it generates significant positive externalities. Not only does everyone in the area benefit from a reduction in vehicle emissions, but everyone who drives a private vehicle also derives a sizable benefit from a reduction in the level of traffic congestion. Governments usually allow mass transit riders to "recapture" a portion of this externality by subsidizing their fares.

In both cases, government action has the effect of eliminating a free rider problem. It is worth noting, however, that talk of a "free rider problem" is somewhat ambiguous. In the case of pure public goods, the problem is that everyone tries to free ride, and so no one provides any of the good. So even though everyone tries to free ride, no one actually succeeds doing it, because no one allows himself to be suckered. The "problem" is therefore that free-riding generates a total failure of cooperation. In the case of quasi-public goods, however, or goods with significant positive externalities, some people may actually succeed in getting a free ride. For example, only a fraction of the audiences of "listener-supported" radio stations – like National Public Radio in the United States – actually donate any money to support these stations. In this case, they are able to free ride on the backs of more avid or conscientious listeners. The problem here is not that cooperation fails, but that some people are not paying their fair share, and exploiting those who are. In some cases, this type of exploitation will be sufficiently objectionable that the state will intervene in order to eliminate it.

Consider, for instance, the case of welfare. In almost every known society, there has been a widely recognized social obligation to contribute to "poor relief." A certain percentage of any population will either lack the competence or skills necessary to sustain themselves, or will suffer some tragedy that impairs their ability to do so. This gives rise to a potential free rider problem. Most properly socialized individuals are disturbed by the sight of human suffering, and so derive some benefit from its alleviation. But if care of the poor is left to charity, i.e. voluntary contribution, then those who have a greater tolerance for the sight of human suffering can free ride off those who do not. This generates the highly perverse consequence that individuals with greater moral sensitivity can be exploited with impunity by those of lesser sensitivity. But since subsidizing immorality runs contrary the more general goals of socialization and social integration in any community, most societies have instituted some form of *mandatory* poor relief. In pre-industrial European societies poor relief was handled by the church, but this should not generate the impression that it was based on voluntary or charitable donation. Churches used to have the quasi-legal power to tithe, which was essentially a taxation power. With the development of the modern nation-state, all taxation rights passed to the state, and so the responsibility for poor relief went along with it.

2. *The "social safety net."* What is commonly referred to as the social safety net in Canada is essentially a set of insurance programs administered by the state. Such programs include worker's compensation, unemployment insurance, health insurance, deposit insurance, loan guarantees, disaster relief, and retirement pensions. These insurance services are really no different in principle from public and quasi-public goods. Insurance markets are especially prone to market failure, and markets simply don't exist for many different kinds of insurance products. As a result, individuals are unable to insulate themselves from a variety of different kinds of risks. Thus the state steps in and provides various kinds of insurance that markets would otherwise not provide. Even though this is just another kind of good, no different in principle from a road or a park, people tend to think of them as a different category of government activity. This is largely a misperception. However, some aspects of these programs are distinctive, due to the type of conditions that generate market failure in insurance markets.

One of the government insurance programs that has done an enormous amount to stabilize the market system is the provision of insurance on bank account deposits (Moss, 2002).

When banks receive deposits, they keep only a small percentage of the cash "on hand." The rest is lent out at interest. This is the primary mechanism through which the surplus wealth in society is invested in replacing and improving its productive technology, and so is an essential component of any well-functioning capitalist economy. The entire system, however, requires that depositors retain confidence in the solvency of their bank. No more than a small percentage of deposits can be "paid back" in any short period of time (usually between 5 and 10 per cent). In order to get more cash, the bank would have to "call in" its loans. Since many of these loans will have been used to purchase equipment, machinery, buildings, vehicles, etc. it will often not be possible for the bank to get its money back in a short time. This means that if too many people try to make withdrawals, the bank will become insolvent. If any depositors come to suspect that this may occur, it generates a classic collective action problem – if the bank is going to go bust, it is better to try to get your money out right away, but doing so increases the chances that the bank will go bust. The result is a "run" on the bank (though which depositors destroy the value of their own assets). In the movie "It's a Wonderful Life," the bank manager was able to persuade his depositors that they were engaging in collectively self-defeating behaviour, and so headed off the run. In real life, runs were much more likely to turn into stampedes. These were genuinely terrifying for depositors, in part because the mere *expectation* that there would be a run could be enough to produce a run (and so runs would tend to occur in irrational and random patterns).

One way of solving this problem is to have some third party offer deposit insurance. The terms of such an insurance policy would be simple – you pay a certain premium, and if you lose your deposits because your bank goes bust, insurance will pay out the full value of the assets. Notice a key advantage of this insurance – the very fact that depositors are insured makes them less likely worry about runs, and so less likely to cause runs. As a result, having insurance against the unfortunate event makes that event much less likely to occur. Unfortunately, no corporation is in a position to offer this kind of insurance, for the simple reason that insurance companies, like banks, can go bankrupt. This means that private deposit insurance would simply transform a run on the bank into a run on the insurance company. The easiest way to solve this problem is to have an insurance company that cannot go bankrupt. This is where the state fits in. By insuring deposits, then using tax revenues to pay out claims, the state can create a *credible* commitment to compensate any losses. This has had a massive impact on the banking system,

and has been a crucial contributing factor in the stability the market economies achieved during the 20th century.

In this case, the reason that the government is able to provide insurance where private markets fail is that it is the only agency capable of making a credible commitment. This is ultimately because it has the power to tax. An insurance company on the other hand, can only pay out on claims as long as its creditors and investors don't pull the plug (and everyone knows that as soon as they face unrecoverable losses, they *will* pull the plug).

A more common reason for market failure in the insurance sector, however, is the problem of adverse selection. Insurance contracts are extremely sensitive to information asymmetries. Because they are designed to transfer assets in response to uncertain events, individuals who can exercise control over these events are able to take advantage of insurance. This is the problem of moral hazard. However, it is not necessary that they be able to exercise control. Even if an individual has better information about the probability of an uncertain event, she may be able to use it to gain an advantage out of an insurance contract. This is the phenomenon known as adverse selection. To take a crude example: if an individual knows that he is going to drop dead of a heart attack within the year, he may choose to take out a huge life insurance policy. Unless the company is able to detect this condition before signing the policy (i.e. restore information symmetry), it will have been set up for a significant loss.

The term "adverse selection" refers to the fact that any company which offers an insurance policy will automatically attract the kind of customers that it least wants – viz. those who pose the greatest risk. Consider how this works in the case of pensions. So-called “defined benefit” pensions are essentially a form of insurance that compensates people for not dying. If you knew exactly when you would die, then you would know exactly how much money you needed to save for your retirement. But since people generally don't know this, they may run out of retirement funds before they die (or they may oversave, piling up a mass of wealth that they will never have the opportunity to enjoy). But while individuals are uncertain when they will die, the law of large numbers allows us to predict with great certainty what portion of a large group will die at what age. This means that individuals can benefit by joining a retirement fund which, for a fixed premium, offers a fixed payout during each year of retirement. The "losers" under any such arrangement are those who pay premiums, but then die young, while the "winners" are those who live for a long time – they will take out more than they put in.

The problem of adverse selection with pension funds arises because people are not completely bereft of information about when they will die. Suppose (to simplify enormously) a pension fund starts up that offers to pay out \$30,000 a year after retirement, in return for a contribution of \$3000 per year for 20 years. Since the average life expectancy in Canada is 79 for men and 85 for women, this plan will have an expected payout of \$420,000 for men, and \$600,000 for women. Assuming that the plan is designed to exactly break even, this means that every man who joins to the pool gets an expected loss of \$90,000, while every woman can expect to gain \$90,000 (in effect, the pension fund redistributes money from men to women). As a result, if a private company offered a pension scheme of this type, only women would sign up for it (and the fund would go bankrupt).

Thus clients of insurance companies have a tendency to self-select in a way that is disastrous for the insurer. So even if it would be profitable to offer insurance to a whole class of people, insurers know that if they offer it, only some subset will take it, and this subset will be formed through self-selection of those most likely to be the largest liabilities. This is a recipe for losing money. The insurance company might respond to this situation by partitioning their clients, so that premiums are adjusted according to age or gender, made dependent upon medical examinations, etc. All of this increases transaction costs, however, making private markets increasingly less feasible. One solution is for the government to provide universal insurance protection. It can do so, again, because it is able to make participation in the scheme *mandatory*. By eliminating consumer choice, it eliminates the problem of adverse selection.

Many aspects of the “safety net” have insurance-like functions, even though they are not always recognized as such. For instance, the single most significant economic risk faced by most women in our society is divorce – especially in the years immediately following the birth of a child. Rising divorce and abandonment rates have been the single most important factor driving the so-called “feminization of poverty” in our society over the past three decades. Thus any economically rational woman should want to buy “divorce insurance.” Yet there are no private markets for such insurance, for fairly obvious reasons. Such insurance would be subject to extreme moral hazard and adverse selection problems. Thus one of the functions that the welfare system has increasingly taken on has been to provide a (very low) baseline level of divorce insurance. (This is especially true in the United States, where most welfare recipients are single mothers.)

3. *Natural monopolies.* Both of the previous two cases involved instances where private markets will either fail to produce a particular good or service at all, or else produce it at an inefficiently low level. The problem of natural monopolies is slightly different. Here markets fail not because they suffer from externalities, or information asymmetries, but from an absence of competition. Any supplier who is able to secure a monopoly position in a market is able to dictate prices by controlling supply. As a result, it will almost always be in its interest to restrict supply, and thereby raise prices. (The equivalent is true in cases of monopsony, when one buyer controls the demand side. Monopsony is relatively rare, compared to monopoly, because there are often significant economies of scale on the production side, but seldom are on the consumption side.) In some cases, government is able prevent this by ensuring that markets do remain competitive (anti-trust action is the most prominent strategy). However, in some cases the most efficient way to produce a good or service is to have only one supplier. This will be the case whenever there are significant economies of scale, or when there is the need for a single standard (e.g. in file formats, communication protocols, etc). Whenever the physical characteristics of a particular production or distribution process make it more efficient to have a single supplier, this is referred to as a **natural monopoly**.

Most of the classic examples of natural monopolies involve services delivered to the home: telephone, water, cable TV, sewage. These tend to become monopolies because once a company is delivering the service in a particular region, the cost of hooking up one more client is miniscule compared to the amount that it would cost a competitor to come in and offer rival service. For example, it might cost several thousand dollars to run a telephone line from a company's central exchange to a client's home. However, once that home is hooked up, it may cost only \$50 to hook up their neighbour. A rival firm would have to bring in an entirely new line.

Natural monopolies do not always occur because of physical distribution constraints. Certainly the distribution of electricity is a natural monopoly. But production of electricity is also subject to increasing returns, depending upon the technology used to produce it. So in Quebec, for instance, all of the electricity in the province comes from hydroelectric dams (which either produce a *lot* of electricity, or none at all). Most of it comes from a single dam at James's Bay. It is also the cheapest electricity in North America. In this case, it is because there are only so many

rivers that can be dammed up that the production of hydroelectric power is a natural monopoly. Similarly, transportation networks like roads and railway lines tend to be natural monopolies, because there is only one feasible route through a particular region. Someone who owns a road is very unlikely to have effective competitors, and so is under no pressure to set prices at a level that will generate an efficient use of the resource.

Natural monopolies can also arise through the entrenchment of standards. This is usually a consequence of what are known as **network externalities**. These are essentially positive externalities that are cumulative in effect. For example, everyone who uses Microsoft Windows creates a slight positive externality for every other user of Microsoft Windows. Not only can any other Windows user sit down in front of the computer and figure out its use quite quickly, but files can be exchanged more easily, software providers are more likely to support that operating system, etc. Even if there are better systems around, if everyone you know is using Windows, it *becomes* in your interest to use Windows as well. Thus the market for desktop computer operating systems is widely regarded as a natural monopoly (with some room for niche players).

In cases where it is not feasible to create genuine competition in these markets, there are two ways of handling natural monopolies. The first is to regulate pricing, so that private corporations continue to provide the service, but are forced to do so at a rate that approximates the market clearing price. This is how electricity distribution, cable television, along with water and sewage systems work pretty much throughout Canada. A different strategy is simply for the government to set up its own company (a Crown corporation), and instruct its managers to produce and price its goods at a level that minimizes deadweight losses.

One thing that should be noted about all three of these types of government activities is their dynamic character. Whether a property right can be exercised over a particular good is in part determined by the current state of technology. Similarly, whether or not a particular production process is a natural monopoly depends upon the technology used in that process. This means that the proper economic role of the state is not set in stone, but must be adjusted in response to changing conditions. For instance, it used to be the case that the only way to collect money from motorists for the use of a road was to install a set of physical toll booths. But these cost an enormous amount to run, and generate significant losses to motorists in the form of delay and inconvenience. Transaction costs of this type generally made it more efficient to pay for the

roads through a gasoline tax. However, new "transponder" technology makes it possible for vehicles to register electronically when they enter a roadway, after which they are billed automatically (as on Ontario's highway 407). This reduces the transaction costs significantly, making it much easier to exercise a property right over a roadway (therefore making it less of a public good, although perhaps still a natural monopoly).

New manufacturing and distribution processes may also make it possible for firms to compete in areas that were once natural monopolies. Most obviously, mobile communications technology (cellular, PCS), by eliminating the need for physical wires, eliminates also the vastly increasing returns to scale associated with traditional telephone technology. This makes it possible for the market to sustain more than one firm. Similarly, competition among long-distance carriers only became possible after the development of fibre-optic lines and computerized billing. Thus deregulation of long-distance services in the late 80's was largely a success. But had the government deregulated the long-distance market in, say, 1975, prices would have gone up, not down, because there would not have been any competition. (Fibre-optic lines have such higher capacity that rival firms need not lay out rival lines, they can rent or "cost-share" one.)

As a result, one would not expect the "size" of the state relative to the private sector to remain fixed. On the one hand, different countries will strike a different balance, depending upon the preferences people happen to have for goods and services of one type or another. There is no "correct" size for the public sector, independent of the preferences of its citizens. Just as there is no "correct" number of shoes, independent of the demand for shoes, there is no "correct" amount of education, infrastructure, knowledge, poor relief, etc. outside of demand, and so no correct size for the welfare state. A country in which people have a strong preference for goods that can only be supplied efficiently by the government will tend to have a larger government. The long-term trends are also difficult to ascertain. On the one hand, a lot of technology has made it easier to acquire and manipulate information, and this has had some effect on extending the range of property rights that people can exercise (e.g. you can charge people for every "hit" you generate for their web site), and has also had a presumptive effect on transaction costs. However, new technology also has the effect of "rationalizing" manufacturing processes, and so may generate increasing returns to scale, thereby creating natural monopolies where there used to be none.

There is also the fact that people's preferences change over time. Some theorists have argued that the historic tendency of the public sector to expand can be explained by the fact that once individuals find that their basic material needs have been taken care of, they start to look for more ephemeral forms of satisfaction. And since ephemeral forms of satisfaction are harder to exercise property rights over, then tend not to be readily supplied by markets. A more affluent and technologically advanced society is one in which there is greater demand for education, greater demand for health services, and a greater demand for insurance against a thousand-and-one types of risk. As a result, there is a steady increase in demand for the type of goods that only government can supply efficiently. In short, since governments are good at solving problems-for-happy-people, as people get happier they tend to look more to their government to solve their problems.

6.3 Regulation

The other major area of welfare state activity in the economic sphere is **regulation**. The state already lays down the "ground rules" under which economic competition will take place. These rules include, first and foremost, the system of property rights. However, in circumstances in which these rules alone are not enough to guarantee "healthy" competition, i.e. competition which generates positive externalities for society at large, the state has the option of imposing a set of more specific rules, designed to correct market imperfections in a particular sector.

The analogy to sports is helpful here. For a long time, rules governing sports simply specified how the games were to be played. The competitive structure of these games then motivated players to constantly improve the skills necessary to succeed, thereby generating a high level of athletic excellence – along with the qualities of character that come with it: perseverance, hard work, determination, etc.. However, when performance-enhancing drugs became available, it became possible for players to excel without cultivating these qualities, and in a way that generated a variety of unhealthy side-effects. This meant that the basic "rules of the game" were no longer sufficient to guarantee a healthy competition. As a result, new rules had to be invented, ones that restricted the range of competitive strategies available to athletes.

Much of the state's regulatory activities can be understood in the same way. Consumer protection laws are a classic example. The system of property rights lays down the ground rules, so that two firms producing a given product can only compete with one another by attracting

more customers. Since they must induce these customers to voluntarily purchase their products, they can only better one another by offering a superior bargain, e.g. a lower price, higher quality goods, etc. However, under conditions of asymmetric information, one firm may be able to fool customers into believing that their product is of higher quality than it in fact is. It will therefore be able to conclude a "voluntary" agreement, but one whose terms the customer would not agree to if fully informed (such an agreement need not be fraudulent, it may simply fall under the "buyer beware" category, where the value realized from a purchase is not as high as expected). When this asymmetry of information states is a structural feature of the market, it means that the system of property rights alone will systematically fail to ensure "healthy" competition between these firms. They may begin to compete with one another to find better ways of concealing the defects in their products. As a result, the state has the capacity to improve this outcome, if it is able to design a rule that will eliminate the set of competitive strategies that generate this negative outcome for consumers. So, for example, by making minimum warranties mandatory, the state makes it possible for consumers to be "undo" contracts they have entered into when products have defects that show up after the point of purchase.

The other key area of regulatory intervention is pollution control. "Pollution" is almost a generic name for negative externalities generated by production processes. Consider the following definition of pollution:

It is the result of (1) the production of a product with no commercial value because it is technologically tied to the production of a product that has economic value, (2) the dumping of the unwanted by-product into the environment to dispose of it, and (3) the consumption of the by-product, with a negative value to anyone forced to consume it, simply by being exposed to it. It is the final characteristic – the externality or negative impact – that explains why pollution is so undesirable (Burgess 1995, 402).

The definition captures the sense in which we speak of noise pollution, light pollution, along with the more usual forms of industrial pollution – carcinogenic chemicals, greenhouse and smog-producing gases, etc. A good example of the impact of regulation on the production of such pollutants is the case of lead. For a long time, lead was a very popular additive to fuel. Lead functions as a catalyst that makes gasoline burn more easily, thereby making it perform as if it

had a higher octane level. Since it costs less to add some lead to gasoline than to further refine it, almost all gasoline in North America contained lead until the mid 1970s. Unfortunately, when lead is burned it forms various compounds that remain suspended in the air – it generates atmospheric pollution. Individuals are then "forced" to consume this lead by inhaling it, with a variety of adverse health consequences that include kidney damage and impaired mental development in young children. However, a company that sold "clean" unleaded fuel could not compete with one that sold "dirty" leaded fuel, because unleaded fuel costs the consumer more, but generates no tangible health benefits for that individual. Thus companies selling leaded gasoline enjoyed a competitive advantage over those that tried to stop.

In response to this situation, the state created a number of rules that changed the structure of the competition. In particular, the United States and Canada passed rules that required all new vehicles to be "unleaded-only," so that eventually the consumer market for leaded gasoline would disappear. This simple rule change generated a dramatic improvement. In the United States, lead emissions from transportation were reduced from 122.6 million tonnes in 1975 to 3.5 million in 1986. Concentrations of lead in the atmosphere dropped by more than 85 per cent (Sunstein 1990, 78). This kind of regulation is extremely effective, in part because the costs it imposes upon consumers in the form of higher gasoline prices are completely outweighed by the benefits they receive from the overall reduction in atmospheric pollution. Because the regulation corrects a market failure, it eliminates a collective action problem, and so generates a Pareto-improvement. Effective regulation makes the economy more efficient.

Naturally, not all regulatory interventions are successful, and it can be very tricky to design rule-changes that will have precisely the desired consequences. However, it is important not to underestimate the improvements in the quality of life that have been brought about through regulation. Despite our ongoing concerns about air quality and pollution, for instance, the level of pollution in most urban centers in North America is significantly lower now than it was in the 1960s. When assessing the impact of a particular regulation, it is important to recognize that the existence of a regulation does not change the underlying incentive structure of a particular market. It remains in the interest of every firm to circumvent regulations that are imposed, even when these regulations are in the general interest. So while it may appear that the regulation is interfering with the profitability of the firm, it is actually just preventing the firm from free riding. As a result, companies in regulated industries always have an incentive to

complain about government interference, regardless of whether this interference is actually having its intended effect, and promoting the general welfare of the community.

In some cases, regulation represents an alternative to direct state provision of goods and services. If a particular market fails, the state can respond by suspending that market and providing the good directly, or it can attempt to restructure the market through regulation, in order to make it function better. When there is a natural monopoly, for instance, the state can create a crown corporation and order its managers to provide the good at market-clearing prices. But it can also allow a private corporation to acquire a monopoly position and then regulate it by, for instance, specifying the maximum price at which it may sell its products. This strategy has both advantages and disadvantages, and so most welfare states use a mixture of regulatory intervention and positive economic provision. In general, regulation is more easily used to control **public bads** – it can prevent firms and individuals from externalizing certain types of costs. Positive welfare-state measures, on the other hand, are usually favoured when it comes to providing public goods. The rationale here is simple – it is much easier to prevent people from harming each other than it is to induce them to help one another. In the United States, however, where there is a strong preference against direct state provision of goods and services, regulation has been used far more aggressively than in any other industrialized democracy. Some have even suggested that the American state is sufficiently unusual in this regard that it should not be classified as a "welfare" state, but rather as a "regulatory" state.

Sometimes clever regulation is able to act as a very effective substitute for direct provision of goods. Most cities, for instance, try to reduce traffic congestion by providing subsidized mass transportation. Motorists "pay" for the reduction in congestion through gasoline and other taxes. In many parts of the United States, however, motorists have traditionally refused to purchase reduced congestion in this way, because they are unwilling to license the state as their agent. One clever regulatory solution that was developed in response to this situation is the HOV (or "high-occupancy vehicle") lane. Most people do not want to share transportation, e.g. by car-pooling, or taking the bus to work, because it is slower and less convenient than taking their own car. This is because everyone has to wait while the vehicle stops to pick up other passengers. However, if everyone takes their own car, it generates enormous congestion that makes the overall commute slower than it would have been if everyone took the bus, or car-pooled. One solution is for the state to provide subsidized mass vehicles, so that they will be

cheaper than private ones. An alternative, however, is to regulate the use of private vehicles. One way of doing this is to pass a law stating that a certain lane on the freeway can only be used by mass-transit vehicles and cars with more than one person in them. When traffic flow is light, this makes no difference. However, as traffic flow increases, the travel time of a multiple-occupant vehicle becomes much lower than that of a single-occupant vehicle. This offsets the additional time that it takes to pick up and drop off these passengers, and so the collective action problem is resolved. This is shown in Figure 6.2. What makes the regulation clever is that it specifically targets congestion – the rule only imposes costs (in the form of delay) on drivers of single-occupancy vehicles if they drive at peak hours.

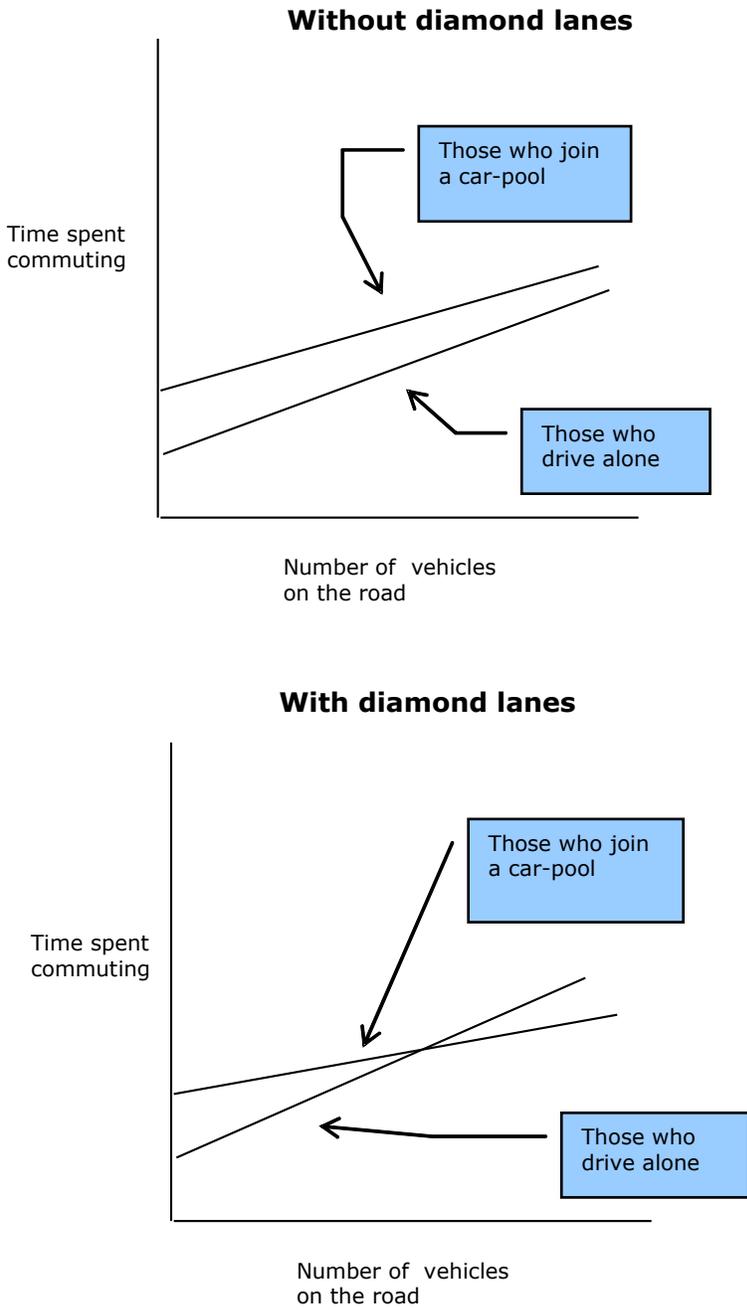


Figure 6.2 Regulation and traffic congestion

One of the attractions of regulation in the American political context is that the population is extremely sensitive to the level of taxation. Regulation allows the state to largely externalize the costs of its policies, and so it need not directly fund them through taxation. But this is very much a mixed blessing. While it is often more efficient to order someone to stop doing something than to do it for them, there are many other cases in which it is merely politically expedient to do so. People who might notice and protest an increase in their taxes are much less likely to either notice or protest an increase in consumer prices due to regulations imposed upon corporations. Ultimately someone has to bear these costs, but regulation makes it hard to see how large these costs are, and where they fall. For this reason, regulations are not always subject to the same kind of careful cost-benefit analysis that programs financed through taxation must undergo. This makes it much harder for people to tell whether or not a regulation is efficient. Furthermore, the regulatory state may have a tendency toward increasingly inefficient interventions over time. According to one analyst:

The trend is unfavorable. As easy social gains (safer cars, for example) have been exhausted, regulations focus more on marginal problems. Ever larger costs are imposed for ever smaller gains. One analysis by the [U.S.] Office of Management and Budget examined the effectiveness of fifty-three health and safety regulations covering everything from toxic emissions to aircraft safety, issued between 1967 and 1991. The study sought to show how much in regulatory cost was required to prevent one premature death. Seventeen of the rules extended a life for less than \$1 million in regulatory costs. Of these, all but five were issued before 1986. By contrast, eighteen rules extended a life for more than \$25 million. Of these, twelve were issued since 1986. The general problem is the same as with the formal budget: How are public benefits and costs to be balanced? But the practical difficulties are greater. Regulatory costs are less visible than taxes. Every regulation seems to advance some worthy cause, and opposition is often led by unpopular corporations (even if the final costs are paid by the public through higher prices). Regulation becomes politically expedient: a substitute – outside the budget – for direct spending, which would either require higher taxes or higher deficits (Samuelson 1997, 170).

Thus many theorists caution against an over-reliance on regulation of the economy, on the grounds that its consequences in efficiency terms are difficult to assess. When the government provides services directly, it must finance those services through taxation. This generates public accounts which can be directly scrutinized and evaluated. The costs imposed by regulation are more difficult to determine. This might generate a presumption in favour of taxation. Unfortunately, taxation has its own peculiar difficulties.

6.4 Taxation

One of the primary uses of taxation is to correct market failure. In private markets, every transfer of wealth between individuals is effected through voluntary exchange. When markets fail, this system of voluntary exchange generates a Pareto-suboptimal outcome. The state, however, can use the power of law to redistribute wealth, i.e. to transfer wealth between individuals directly. It is the only agency in society that has that power to circumvent the system of voluntary exchange in this way. As a result, it has the capacity to achieve outcomes that would not be the equilibria of voluntary exchanges. Thus the state has the capacity to achieve efficient outcomes in areas where the system of voluntary exchange would generate a suboptimal outcome. For example, whenever production of a particular good generates significant positive externalities, private markets will underproduce that good, because its price will not reflect its true value to those who consume it. If the state is able to intervene and subsidize the good by transferring additional resources from those who consume it to those who produce it, the result will be a Pareto-improvement – its intervention will lead to a higher, more efficient level of that good being produced.

When the state redistributes wealth in this way, those who experience a monetary loss in the transfer are paying a **tax** (those who experience of monetary gain from the transfer are often said to be receiving an **entitlement**, or a subsidy). The state's taxation power can be exercised in one of two ways. In the presence of negative externalities, taxation can be used to impose costs upon producers in such a way that their costs of production will be moved closer to the true social cost. When goods that generate positive externalities are not being produced, taxation can be used to finance the subsidization or direct provision of these goods. The former are referred to as **Pigovian taxes** (named after the economist Arthur Cecil Pigou). Such taxes are designed to penalize activities that generate significant negative externalities (when such taxes are imposed

upon consumption goods, they are sometimes referred to as "sin taxes." For example, alcohol and tobacco have special taxes imposed upon them designed to lower the level of consumption.) A lot of pollution-reduction, for example, is achieved through a combination of regulation and Pigovian taxes – firms are forced to dispose of hazardous waste in designated areas, for instance, and then charged special disposal fees for every unit of waste that they bring to the disposal site; or else they have to buy an expensive permit in order to generate a certain volume of atmospheric pollution, effluent, etc.

Pigovian taxes are usually popular with the public at large, because the underlying principle seems sound. The slogan "make polluters pay," has intuitive appeal because it is inappropriate for the firm to profit from its capacity to externalize costs. It is sometimes objected that the polluters will not really pay, but will simply "pass along" the costs to their customers. But this is not really to the point. Whether the firm's shareholders or its customers ultimately take the hit, the advantage of the tax comes from the fact that it lowers the level of the good being produced to a more socially desirable level. Whenever a firm is externalizing costs, this means that it will be producing too much of the good in question, and consumers will be buying too much of it, from the standpoint of efficiency. Taxing the firm raises its costs of production, forcing it to raise its prices, and thereby reducing demand for its products. This then prompts the firm to produce less of the good in question.

The other major form of taxation can be referred to as **direct taxation** – this is where the government simply raises money from a broad stratum of the population in order to supply public and quasi-public goods. Income taxes and consumption taxes are of this type (since the state does not intend to discourage either working or shopping – although such taxes may have this effect). Direct taxation tends to be less attractive than Pigovian taxation, despite that fact that it provides the same kind of benefits. In order to understand the politics of direct taxation, it is important to look a bit more carefully at the psychology of it.

The basic function of direct taxation is to eliminate free-riders (since whenever there is a positive externality, someone has a chance to free-ride). By taxing everyone, then providing a public good, the state is doing no more than requiring that people pay for the benefits they receive. Thus the system of taxation, like the system of property rights, is a generic mechanism for internalizing externalities. In the case of property rights, the state eliminates free riders by forcing people to pay those whose property they use or consume. This makes people willing to

build houses, plant crops, and so on, because they can be confident they will not be exploited. However, the system of property cannot always provide such an assurance, because it is impossible to assign a property right to some goods. In such cases, the state can only eliminate free riders by directly seizing wealth from those who benefit, and transferring it to those who supply the good, i.e. through a taxation-entitlement system. In the former case, private individuals conduct the transaction, and the state merely enforces the rules of the game. In the second case, the state conducts the actual transaction on behalf of the parties. The former is just a decentralized version of the latter.

This means, however, that governments will constantly get resistance from citizens to taxation, and citizens will seek to avoid paying their taxes – for the same reason that any private property system generates a problem of theft. Both are nothing but an expression of the underlying free rider incentive that motivated the state's intervention. People want government services, but they would much rather get them without paying for them. People know that even if they cheat on their taxes, government services will probably still be provided (just as they know that if they shoplift, the store will probably not go bankrupt). Naturally, if everyone does this it generates a suboptimal outcome, but that doesn't change the structure of the incentive. That is precisely why governments must *force* citizens to pay taxes. Government provision of public goods is not a strategic or market equilibrium, which means that people will constantly be trying to defect. This of course generates deadweight losses in the form of enforcement costs. However, the private property system also has enforcement costs, since it's always cheaper to steal something than to pay for it. There is no reason to think that the deadweight losses generated by tax avoidance are higher than those generated by theft.

In any case, the fact that citizens often resist and seek to avoid taxes should not be taken as evidence that they do not derive benefits from the tax regime. It can much more easily be understood as the reflection of the free rider incentive. When the state's direct taxation policy is efficient, all of the monetary losses are fully compensated by the value of the externalities that are generated from the activities that are funded by these taxes. It is just that because the benefits come in the form of externalities, an opportunity is created for free-riding. The important thing to keep in mind is simply that, as Thomas Nagel put it, the government "is not really forcing people to do what they don't want to do, but rather enabling them to do what they want to do by forcing them to do it" (1987, 215).

Setting aside the free rider problem, there is one further complication. Precisely by virtue of the fact that the benefits people receive from their taxes take the form of externalities, they often find it difficult to assess the value of what they "get back" from the state, unless it also takes a monetary form. Precisely because these externalities have no market value, it is difficult to state how much benefit people get compared to what they pay. For instance, having a good judicial system generates benefits for everyone in society – low levels of corruption, for instance, mean that the guilty are less likely to go unpunished, but also that the innocent are less likely to be wrongly convicted. Most citizens of Canada suffer very little anxiety over the possibility of being wrongfully convicted of a crime. It is very hard to pin down, however, precisely how much a "benefit" they get from this peace of mind. They pay for it, however, in very tangible ways – from the salaries of judges right down to the subsidization of legal aid services. So one of the reasons that taxation issues are controversial is not just that they can be difficult to evaluate, but because the structure of taxation requires that agents be forced to suffer a tangible monetary loss, in order to receive a *less tangible* set of benefits. This creates a cognitive bias against taxation, which can often be seen in popular discussion of such issues. Hobbes noticed this a long time ago:

For all men are by nature provided of notable multiplying glasses, (that is their passions and self-love,) through which, every little payment appeareth a great grievance; but are destitute of those prospective glasses, (namely moral and civil science,) to see afar off the miseries that hang over them, and cannot without such payments be avoided (1651, 129).

When you pay for something in a market, you see what you get. When government provides a good or service, however, everyone appears to get it "for free." As a result, the taxes paid seem like a rip-off. For example, people often think that only those actually drawing welfare cheques derive any benefits from the welfare system, because they fail to perceive the indirect benefits that poor relief confers upon the entire society (reduction of crime, improvements in quality of urban environment, stabilization of rural populations, improved educational level of population, decreased infant mortality, etc.) It is hard to assign a value to these benefits, precisely because they occur in areas of market failure. For the same reason, there is no way to determine how much these benefits "contribute" to the economy, because they escape

measurement. (As a result, economic indicators like the GDP, while very important, are inherently biased against the public sector.)

The problem is such that even when people are *receiving* a redistribution from the state, they sometimes fail to realize it. Often this is simply because they don't know the value of the benefits they receive. Many people consider it normal to spend their entire lives paying for their house. But considering the fact that many serious medical procedures cost about as much as an average Canadian home, people should also not be too upset if they must spend their entire lives paying the taxes that will be used to supply this medical care. However, many people find it unreasonable when their tax bill is larger than their mortgage. Often this is just because the house comes with a price tag, while the medical care does not. Unfortunately, the problem is difficult to correct, because instituting the kind of accounting system that would be needed to assign a price tag to all medical services received would impose transaction costs, and eliminating these costs is a large part of what makes public health insurance systems more efficient than private ones.

The suggestion being made here is that the welfare state should be seen as a mechanism through which people purchase goods, just like a market. The way people purchase goods in a market is through private exchange. The way people purchase goods through the state is by taxing themselves. Some goods will be supplied only if everyone agrees to pay, others will be supplied to each individual who agrees to pay. The welfare state provides the former, markets provide the latter. Thus, as Sunstein puts it, there are cases where "the force of law is necessary *in order to allow people to obtain what they want*" (1990, 43). Because some goods can only be purchased through taxation, consumer choice can sometimes only be exercised through the medium of law.

Because of this, citizens often have what has been called a "schizophrenic" attitude toward the public provision of goods. This often shows up in opinion polls about taxation: if asked whether they want their taxes lowered, a majority will usually say yes. But this reflects, at least in part, a free-rider incentive, because when asked a "tied question," such as whether they would be willing to have their taxes *increased* in order to provide more education, or more health care, etc. a majority will also often say yes. In this case, the willingness to pay higher taxes simply reflects their desire to purchase these goods, along with the collateral benefits they provide. As a result, one cannot infer from tax resistance that people are consuming an adequate

level of public goods, because it is very difficult to separate out the free-rider incentive from the effects of genuine satiation.

For the same reason, people's voluntary choices in markets may be inconsistent with their political choices. For example, some people may not give any money to charity, but will support the use of tax money to supply welfare. This is because, while they support the goal of helping those in distress, they are not willing to do so unilaterally. They don't want to be suckered, and so they are willing to assist the poor only if everyone else does too. The same phenomenon has been noticed in the case of regulation, where people sometimes support the institution of rules that would prohibit actions that, in the absence of such a rule, they would choose to perform:

Some government controls commonly regarded as unjustified paternalism can be understood instead as an effort to facilitate the satisfaction of private desires. The phenomenon suggests one explanation for the striking fact that people's preferences, in their capacities as voters or citizens, are quite different from what they are when people act in their purely private or market capacities. A majority of citizens might support regulation that would prevent them from engaging in the very conduct which, in an unregulated system, they are led to choose because of a prisoner's dilemma or a collective action problem. People might prefer a system in which all are prevented from (for example) littering or driving without seat belts even if they would personally rather litter or drive without belts than be the only nonlitterers or beltbucklers (Sunstein 1990, 51).

Thus individuals will often engage in private behaviour that undermines public policies that they support. For example, even if people support public provision of health insurance, given the opportunity to opt out and join a private plan they may do so, simply because it confers a private benefit upon them. Their actions, however, generate an adverse selection problem that may undermine the public system. So if given a chance to vote upon the matter, they might decide to make the public system universal and mandatory. In so doing, they are denying themselves the opportunity to buy something that, given the opportunity, they would actually choose to buy. They are *choosing* to suspend their individual "consumer sovereignty." But this a rational decision, for the same reason that, in a prisoner's dilemma, it may be rational to commit

oneself to not confessing (even though, in so doing, one is preventing oneself from doing something that, given the opportunity, one would in fact do).

6.5 An example: socialized medicine

The structural context that motivates the public sector to provide a particular set of commodities is well illustrated in the case of health care. There never has been, and there never will be, a competitive market for the provision of health care services. In fact, the doctor-patient relationship requires so much trust that it is better characterized as a fiduciary relationship than as a market transaction. This is simply because of the massive information asymmetries. In the same way that you hire a lawyer to act on your behalf, not to sell you legal advice (because you have no way of evaluating the advice you are given, short of hiring another lawyer to do the same thing all over again), you also hire a doctor to conduct a medical examination or procedure on your behalf. This trust relationship was traditionally secured through a code of professional ethics, and enforced by medical associations.

Unfortunately, changes in medical technology have placed serious strains on this relationship. Once medical care became very expensive, most patients found themselves unable to pay for it "out of pocket." This led naturally to the development of health care insurance. Most of the problems in the health care sector now involve the relationship between health care providers and insurers. This relationship, when conducted in the private sector, is beset by insuperable moral hazard and adverse selection problems:

1. Moral hazard. The basic problem with health insurance, to put it bluntly, is that doctors and patients conspire to extract a maximum payment from the insurance provider. Patients, obviously, have an incentive to get the most expensive treatment options available. Since their health insurance premiums are a sunk cost, they generally want to get as much as they can for their money. The cost of their treatment is entirely externalized, i.e. it is displaced onto the other members of the insurance scheme. Doctors, at the same time, have no incentive to minimize costs. Furthermore, doctors often associate the "best quality care" with "the most expensive care," and so feel an obligation *not* to economize. They might also expose themselves to lawsuits by not providing the most expensive treatment options available. Thus even if one completely ignores the fact that doctors can derive enormous personal gain from prescribing expensive care,

it is a structural feature of the doctor-patient relationship that both doctors and patients have an incentive to maximize costs.

From the patient's perspective, this cost inflation creates a classic collective action problem. (Recall that multilateral moral hazard problems generate a race to the bottom among participants in an insurance scheme.) Every patient has an incentive to select the most expensive treatment options, because the cost is paid for largely by the premiums of *other* policy holders. However, when everyone does this, the insurance company is forced to increase the premium level. As a result, everyone winds up purchasing more health care than they actually want to purchase. (And to make it worse, when the premiums go up, people often respond by trying even harder to "get their money's worth" out of their insurance, by increasing their consumption of medical care. This is, of course, collectively self-defeating, as it just leads to another increase in premium levels the following year.) Once this race to the bottom gets started, it can have extraordinary consequences. Health insurance premiums in the United States, for example, increased at an annual rate of approximately 15 per cent throughout most of the 1990s.

The only way to control the escalation of health care expenses, in this sort of situation, is for insurance companies to monitor the claims that are submitted, and to discipline either patients or doctors who opt for unreasonably expensive treatment options. This is where the first weakness of the private sector shows up. Because of information asymmetries, it is very difficult for health insurance providers to monitor doctors and patients. For example, various diagnostic tests can be quite expensive (ultrasound US\$295, chest X-Ray US\$115, CAT scan US\$950).^{*} There is no real way to tell if a doctor is justified in ordering an of these scans, because they are just tests, not medical procedures. It never hurts to have the information that a CAT scan reveals, since at very least it excludes certain possibilities of serious injury. But since hospitals can only pay for their machines by using them as much as possible, there is a predictable overuse of this sort of technology in the United States. For example, doctors routinely order MRIs on sprained ankles and twisted elbows, and there is not much that insurance companies can do about it.

The more important point, however, is that when health insurance is provided by the private sector, insurance companies suffer from a collective action problem when it comes to disciplining health care providers. If insurers suspect that a particular hospital is overusing MRIs, it can decide to conduct an inquiry. But such an inquiry would be very expensive. Since the

* Data from 1998. John L. Cameron (ed.), *Current Surgical Therapy*, 6th edn. (St. Louis: Mosby, 1998), p. 590.

insurance provider is unable to tell which tests and procedures are medically indicated and which are not, it has no choice but to hire its own doctors to go in and reexamine the patients. Such an inquiry, however, generates positive externalities for other insurance firms. In the same way that all shareholders benefit from the actions of those who undertake to discipline management, all insurance companies benefit from those who undertake to discipline health care providers. This generates a classic free rider problem, which may lead insurance providers to undersupervise health care providers.

2. *Adverse selection.* As if the moral hazard problems were not enough, health insurance companies are also subject to serious adverse selection problems. As a result, certain types of insurance products will not be sold on private markets, even if there is substantial demand for them. To see why, consider the following example. Suppose a company starts out naively, offering a medical insurance package that anyone can purchase at any time of their life. In order to calculate the premium level, they take the total cost of health care for their clients over the year, and divide that by the number of clients. This gives the average cost per client, which they then top up to cover overhead. They set this as the premium (this is known as “community rating”).

The problem with this arrangement is that even though it is difficult to know exactly when you will need to be hospitalized, roughly two-thirds of all health care spending occurs in the last five years of life. Since people over the age of 60 are much more likely to be in the last five years of their life than people under the age of 30, this means that the package being offered by the insurance company is a much better deal for those of more advanced years. As a result, if the premium were \$5000 per year, the only people who would buy it would be those who think that their medical expenses will be that high, (e.g. people with a history of medical problems, women who intend to get pregnant, men or women over the age of 60, etc.) No young healthy person in their right mind would be willing to buy insurance from this company. If another company came along offering health insurance limited to people aged 20-40, they could charge a fraction of the premium, and would be able to steal away all of the best customers from the first firm.

Because of this adverse selection problem, a certain class of persons always tend to become "uninsurable." Their anticipated health care expenses are so high that no one else will be

willing to join a "pool" that they belong to. For example, in the United States, everyone over the age of 65 is effectively uninsurable. Notice that this is another example of collectively self-defeating behaviour. When I am 32 years old, I naturally want to have health insurance that will pay for my medical care when I am retired. However, at the moment, I would be crazy to join an insurance scheme that covers individuals over the age of 65. I am better off signing up in a pool that covers only people in their 30s. Of course, when everybody in their 30s does the same thing, every company that does insure people over the age of 65 will go bankrupt. As a result, I may find that as I get older, there are fewer and fewer companies around that are willing to sell me health insurance.

3. Transaction costs. Finally, medical care is subject to enormous transaction costs. When patients are covered by different insurance companies, it means that the hospital or doctor has to keep track of every product or service that is used on every patient, and bill the appropriate insurer. For example, hospitals in the United States provide itemized bills that include such items as bandages, meals, and even orderly services, such as being pushed down the hall in a wheelchair. (On an anecdotal note, after a brief stay in an American hospital, I noticed on the bill I received that I had been charged US\$8 for a Tylenol pill. Since the retail price of a single Tylenol pill is less than 10 cents, transaction costs presumably made up the balance of the expense. This fact might motivate patients to bring their own Tylenol the next time they check into the hospital, except for the fact that they don't have to pay, the insurance company does!)

Health care experts in the United States estimate that between 20 and 25 percent of health care spending in a private insurance system is absorbed into administrative costs. Public health care systems have considerably lower transaction costs. In 1987 the United States spent \$95 per capita on health insurance overhead, whereas Canada spent only \$18 (Richards 1997, 122). The amount of money being wasted here is significant – between 2.5 and 3 per cent of the U.S. GDP. To dramatize this, consider the fact that the United States spent a whopping \$36 per capita on health research, versus only \$13 per capita in Canada. Nevertheless, Americans managed to spend twice as much maintaining the private sector health bureaucracy as they did on research.

The Canadian health care system attempts to resolve these problems by having the government monopolize the market for primary health insurance. This is referred to as a "single

payer" system, and it is different from many European systems, where the government provides health care services directly. In Canada, only insurance is provided by government. Most actual health care is provided by the private sector (either by non-profit hospitals, or by for-profit physicians, either self-employed, or organized into partnerships or corporations). This government intervention in the insurance sector is very effective. It completely eliminates the adverse selection problem, and does an enormous amount to reduce transaction costs. Since there is a single payer, detailed billing records need not be kept. And by retaining the private delivery of health care services, the Canadian system creates a much more decentralized market for health care provision. Most doctors in Canada are essentially small business owners. Clinics are run by small groups of doctors, who bill the government insurance plan for all health care services that they provide. All of their overhead (secretaries, supplies, etc.) count as expenses against these earnings. This gives them an incentive to innovate, and to economize on delivery costs.

In the United States, by contrast, an attempt has been made to get around the failure of the insurance market by *vertically integrating* health insurance companies with health care providers. Medicine in the United States is increasingly dominated by HMOs (health management organizations). Instead of "signing up" with an insurance company, people join an HMO. For a flat fee, the HMO commits itself to providing primary health care services to that individual (this system of payment is referred to as "capitation," i.e. "by the head" instead of "by the service."). When the person becomes ill, he or she is cared for by a physician who is (usually) a salaried employee of the HMO.

The central advantage of the HMO system is that it eliminates moral hazard. In effect, the internal diversification and risk-pooling provided by the size of the corporation acts as a substitute for health care insurance. Further cost control is achieved by making it in the health care provider's financial interest to provide as little care as possible (since they receive the "capitated" fee for the patient regardless of whether they provide any care). Some people worry that this reverses the incentives too dramatically, making it tempting for doctors to compromise patient care in order to improve returns. In any case, the point here is simply that the arrangement eliminates moral hazard. Transaction costs are also reduced, because a hospital owned by a particular HMO will only treat patients who are clients of that same firm. The weakness of the HMO system is that it does little to handle the adverse selection problem, because HMOs are still free to reject anyone who asks to sign up. As a result, Americans still

rely upon government to provide health insurance to a significant segment of the population. In particular, everyone over the age of 65 receives government health insurance, since they are essentially uninsurable in the private sector.

When compared to the American HMO model, the central weakness of the Canadian single payer system is that it does not entirely eliminate moral hazard problems. Cost inflation is significantly mitigated by the fact that, when there is only one insurance provider, it is generally in their interest to investigate and determine how much various services are actually worth, and when they are needed. Thus provincial governments in Canada negotiate with health care providers to determine the fee that will be charged for each service. By eliminating the free rider problem that arises among private health insurance firms, a government monopoly is able to exercise much more effective discipline over health care providers. However, the Canadian system relies largely on a "fee-for-service" structure – doctors are paid for every procedure or test they perform. This gives them an incentive to overtreat patients. As a result, various Canadian provinces have been experimenting with paying physicians a salary -- for example, Quebec's CLSCs (*Centres locaux de services communautaires*) pay their doctors at a flat rate. Various attempts have also been made to introduce capitation-style arrangements with private clinics. While these hold out the promise of better cost control, they have the downside of limiting patient choice among physicians – one of the most keenly felt weaknesses of the HMO system.

Whatever the specific merits of the Canadian and American systems, they both illustrate the same set of principles. The problem starts with market failure in the health insurance sector. The standard way to resolve this is to take the problematic transactions "off the market," substituting bureaucratic administration for private exchange relations. In the United States, a corporate bureaucracy is used instead of a government bureaucracy. Because HMOs are still in the process of consolidating their control of the health care sector in the United States, any evaluation of the two systems is somewhat premature. For the moment, however, it is worth noting that the United States has a lot of catching up to do. Canada outperforms the United States on all of the standard indices used to measure health outcomes (e.g. life expectancy, infant mortality, etc.) It does so using a significantly smaller fraction of its GDP (9 per cent versus 15 per cent). Looking at the long run, it is worth noting that the Canadian system is also far more decentralized than the American. In the United States, there are single HMOs with millions of

clients. Thus the American system, despite the fact that it is technically "private," relies much more heavily on vertical management structures than the Canadian. It is therefore in a much worse position to take advantage of market incentives.

6.6 Limitations of government

Socialized medicine provides an fairly clear-cut example of how the state can administer a particular sector of the economy more efficiently than the private sector. It is important to recall that markets produce efficient outcomes *only* when they are properly structured – so that the only way for individuals and firms to maximize their private gains is through price competition. Health care is a field where these conditions are almost entirely absent, largely due to information asymmetries. Furthermore, the state is able to correct many of these problems through a fairly clean and indirect intervention in the insurance sector. As a result, the case for public health care is very easy to make. But this is not always the case. In order for the state to produce efficiency gains, it is not enough for there to be market failure in the private sector. It must also be the case that the state is able to perform more efficiently. But as we saw in the previous chapter, bureaucratic organization is subject to diminishing returns – generally speaking, the larger a "vertically integrated" organization gets, the less efficiently it performs. Just as there are diminishing returns to management in corporations, government intervention is also subject to diminishing returns.

Assessing the relative merits of public and private sector provision of particular goods and services is fraught with difficulty. On the one hand, many advocates of the "free market" have an entirely irrational bias against the public sector, which is aided and abetted by a number of popular misunderstandings of how markets work. On the other hand, people with little exposure to economic reasoning have a marked tendency to overestimate the ability of governments to intervene effectively in the economy. Sometimes people think that the state should be engaged in a particular activity simply because they wrongly believe that the only way to guarantee a socially useful outcome is to have "someone in charge." Economist Joseph Stiglitz argues that this "control mentality" is based on two fallacies: "it overestimates the powers of direct control and it underestimates the powers of indirect control" (1989, 34).

Governments have no choice but to use bureaucratic means to achieve most of their objectives, and the bureaucratic organizational form is subject to its own peculiar inefficiencies.

The reason that the state is able to intervene effectively in some instances is that the inefficiencies generated by bureaucracies are generally *different* from those generated by improperly structured markets. This means that in some circumstance, the inefficiencies generated by the public sector will be less than those generated by the private sector. But assessing this can be very difficult to do. This is because many of the benefits provided by the state, as we have seen, take the form of externalities (either public goods, or the elimination of public bads). As a result, it is inappropriate to think that simple monetary measures will provide any insight into the performance of the public sector. It is better to keep in mind the specific weaknesses that impair government performance, in order to determine how likely they are to arise in a particular sector. What follows is a brief summary of these problems:

1. Principal-agent problems. Governments are subject to the same principal-agent problems as firms. In a perfect world, governments would simply tell their managers what to do, and managers would carry out their orders to the best of their ability. In the real world, effective control is only possible if the principal has an enormous amount of information, and the agent has very few opportunities for strategic behaviour. Neither of these tend to be that case with the state. While the scale of state makes it difficult for it to assemble and manage information, the fact that it is democratically governed means that it is generally much less autocratic than a comparably-sized firm. (For example, concern over privacy and civil liberties means that various branches of government are legally prohibited from sharing information – something that would never happen in a corporation). This means that public sector employees have ample opportunity to engage in strategic behaviour.

The state also faces peculiar difficulties when it comes to providing its managers with appropriate performance incentives. In this area, the public sector generally has more limited means at its disposal than the private. Most obviously, governments are unable to offer their managers an "ownership stake." Furthermore, financial compensation is often capped by limited popular tolerance of large salaries in the public sector. As a result, while government employees in Canada are, on average, paid more than their private sector counterparts, the government is unable to offer its top managers salaries that are competitive with the private sector (Richards

1997, 35).^{*} The result can be a form of adverse selection between the private and public sector. Finally, it is important to note that any moral incentives that managers in the public sector have had – the ideal of disinterested public service – appears to have suffered under more than three decades of public hostility toward "government bureaucrats."

2. *Taxation.* Taxation is a very blunt instrument, and tends to generate deadweight losses. The deadweight losses occur partly because taxes increase transaction costs, and so prevent certain otherwise beneficial exchanges from occurring. Deadweight losses also arise when individuals take evasive action in order to avoid paying. This is not always a problem with Pigovian taxes (where this evasive action is precisely what generates the efficiency gain). But it is a problem with broad-based income or consumption taxes, which *may* have the effect of discouraging socially beneficial activities – such as working. Pigovian taxes tend to be tricky, on the other hand, because it is very hard to know just how high to set them. Often the information that would be required is simply not available. Similarly, it is very hard to "recapture" positive externalities by imposing taxes on those who benefit from them. As a result, the tax system functions in a very "rough-and-ready" way. Most improvements in efficiency generated by the tax system will only show up when one examines the benefits that individuals receive over their entire lifetime, or one looks at very large groups of individuals in the short term.

3. *Rent-seeking.* One of the key institutional characteristics of the state is that it has the power to redistribute wealth. While this unique power is what allows the state to carry out a number of economic activities more efficiently than the private sector, it also has the consequence of attracting what are called **rent-seekers**. In the classic case of beneficial state intervention, two individuals are in a situation where they can benefit from a redistribution of goods between them.

* Data on the "compensation premium" that low-level public-sector workers enjoy remains controversial. The underlying problem, however, is the importation into the public sector of an employer-employee bargaining model that was originally developed for use in the private sector, viz. labour unions. In the private sector, union demands are moderated by the fact that they risk forcing their employer into bankruptcy. Thus competition imposes discipline upon both corporations and unions. Since the government cannot go bankrupt, the threat of "back-to-work" legislation must be used as a substitute to moderate the demands of public sector unions. This can lead to economically irrational outcomes, since the use of such legislation is generally governed by the political situation of the ruling party, rather than by the fiscal situation of the state.

However, due to some underlying collective action problem (such as a transaction cost), they are unable to effect this transfer through voluntary exchange. In this case, the state can step in, reassign ownership of these goods, and improve the welfare of both parties. But this power can be misused. A corrupt government official may choose to redistribute goods in a way that benefits one party and harms another. Similarly, an individual may convince an unsuspecting government official that an intervention would be efficiency-promoting when it in fact benefits that individual at the expense of others.

People who petition the government to effect transfers that are not in the "general interest," i.e. efficiency-promoting, are referred to as rent-seekers. These rent-seekers are a familiar feature of the political landscape. For example, companies can try to get competitive, or they can try to secure an advantage over their rivals through special tax breaks, or by convincing officials to impose tariffs on foreign competitors. People who suffer from a particular misfortune may form a lobby group to seek compensation instead of seeking to avoid such misfortunes in the future. All of this can be very wasteful, particularly because even when these groups do not succeed in getting an unproductive transfer, they waste resources in trying to secure one.

All of these problems that afflict the public sector are quite serious. However, the idea that because government has these problems, we would be better off without it, is as ridiculous as the idea that because corporations have similar problems, we would be better off without them, or that because markets sometimes fail, we would be better off without them. In all cases, the solution is to strike a balance. Some goods will be produced by simple contract and exchange between private individuals (with minimal constraint by social norms – individuals must respect each others' rights). Some will only be produced by corporations (medium normative constraint – trust, reciprocity within the corporate group, along with principal-agent relations and some use of sanctions). Some goods will only be produced by the state (maximum normative constraint – state is able to guarantee compliance among all actors). All three are simply different kinds of institutional mechanisms designed for delivering the goods.

Key words

direct taxation

entitlement

natural monopoly

network externalities

Pigovian taxes

public bads

public goods

regulation

rent-seeking

welfare state