7 Regulatory Strategies

If the state wants to control, say, the pollution of a river, it may approach the issue in a number of ways. It may decide to regulate the pollution directly by means of a government department or agency; it may rely on the polluting firms to self-regulate (perhaps under state oversight); or it might delegate the control function to third parties such as public interest groups and the commercial partners of the polluters.[1] Under these different arrangements, a number of instruments may be deployed. Thus, if the state regulates directly, the dumping of noxious substances may be made unlawful or, alternatively, the state may give rewards (e.g. tax deductions) to those existing polluters who reduce the levels of their discharges. Manufacturers might be compelled to tell the public how much pollution is caused in making each product or rights might be allocated so as to allow the victims of pollution to recover damages from polluters. In relation to many risks it may be appropriate to regulate by means of a mixture of instruments and to apply these through a variety of bodies—be these governmental, self-regulatory, corporate, commercial, or public interest group.[2]

This chapter looks at the main instruments that the state can use to regulate directly. Chapter 8 will then examine self-regulation and other modes of delegating the regulatory function to bodies beyond the state. In doing so, it will look at the case for relying on audits of corporate risk management systems (sometimes called 'meta-regulation') and the challenge of finding optimal mixes of regulatory instruments and institutions.

It is clear that choosing the right strategy for regulating matters. A regulatory system will be difficult to justify—no matter how well it seems to be performing—if critics can argue that a different strategy would more effectively achieve relevant objectives. How, though, can we map out the array of different regulatory techniques? A starting point, when focusing on direct state regulation, is to consider the basic capacities or resources that governments possess and which can be used to influence industrial, economic, or social activity. These have been described as follows:[3]

[1] See N. Gunningham and P. Grabosky, *Smart Regulation* (Oxford, 1998).

[2] Ibid., esp. ch. 6.

[3] See C.C. Hood, *The Tools of Government* (London, 1983), 5; T.C. Daintith, 'The Techniques of Government' in J. Jowell and D. Oliver (eds), *The Changing Constitution* (3rd edn, Oxford, 1994). Lawrence Lessig offers an alternative breakdown of 'modalities of regulation' into: law, markets, norms and architecture—see L. Lessig, *Code and Other Laws of Cyberspace* (New York, 1999).

To command—where legal authority and the command of law is used to pursue policy objectives.

To deploy wealth—where contracts, grants, loans, subsidies, or other incentives are used to influence conduct.

To harness markets—where governments channel competitive forces to particular ends (for example, by using franchise auctions to achieve benefits for consumers).

To inform—where information is deployed strategically (e.g. so as to empower consumers).

To act directly—where the state takes physical action itself (e.g. to contain a hazard or nuisance).

To confer protected rights—where rights and liability rules are structured and allocated so as to create desired incentives and constraints (e.g. rights to clean water are created in order to deter polluters).

A number of basic regulatory strategies are built on the use of the above capacities or resources and can be distinguished from each other as follows.[4]

Command and Control

The essence of command and control (C & C) regulation is the exercise of influence by imposing standards backed by criminal sanctions.[5] Thus, the Health and Safety Executive may bring criminal prosecutions against occupiers who breach health and safety regulations. The force of law is used to prohibit certain forms of conduct, to demand some positive actions, or to lay down conditions for entry into a sector.

Regulators who operate C & C techniques are sometimes equipped with rule-making powers (as is often the case in the USA). In the UK, however, it is common for regulatory standards to be set by government departments through primary or secondary legislation and then enforced by regulatory bureaucracies. C & C thus involves the setting of standards within a rule, it often entails some kind of licensing process to screen entry to an activity, and may set out to control not merely the quality of a service or the manner of

[4] On regulatory strategies in general use, see S. Breyer, *Regulation and Its Reform* (Cambridge, MA, 1982), esp. ch. 8; A. Ogus, *Regulation: Legal Form and Economic Theory* (Oxford, 1994), esp. pts. III and IV; Gunningham and Grabosky, *Smart Regulation*, ch. 2.

[5] On command and control and alternatives, see R. Baldwin, 'Regulation: After Command and Control' in K. Hawkins (ed.), *The Human Face of Law* (Oxford, 1997); N. Keohane, R. Revesz, and R. Stavins, 'The Choice of Regulatory Instruments in Environmental Policy' (1998) 22 *Harvard Environmental Law Review* 313–67.

production but also the allocation of resources, products, or commodities and the prices charged to consumers[6] or the profits made by enterprises.

The strengths of C & C regulation (as compared to techniques based, say, on the use of economic incentives such as taxes or subsidies) are that the force of law can be used to impose fixed standards with immediacy and to prohibit activity not conforming to such standards. In political terms, the regulator or government is seen to be acting forcefully and to be taking a clear stand: by designating some forms of behaviour as unacceptable; by excluding dangerous parties from relevant areas; by protecting the public; and establishing penalties for those engaging in offensive conduct. Some forms of behaviour can thus be outlawed completely and the ill-qualified can be stopped from practising activities likely to produce harms. The public, as a result, can be assured that the might of the law is being used both practically and symbolically in their aid.

C & C regulation is not, however, problem-free and, during the 1980s in particular, a number of North American socio-legal scholars and economists alleged a series of weaknesses.[7] Such concerns were echoed by many politicians on both sides of the Atlantic—particularly those predisposed to doubt the value of governmental rather than market-based modes of influence.

CAPTURE

A first worry was that in C & C regulation the relationships between the regulators and the regulated might tend to become too close and lead to capture—the pursuit of the regulated enterprises' interests, rather than those of the public at large.[8] A number of versions of capture theory have been put forward.[9] 'Life-cycle' accounts suggest that agencies progress through various stages until, lonely, frightened, and old, they become the protectors of the regulated industry, rather than of the public interest;[10] 'interest-group' explanations stress the extent to which regulators can be influenced by the claims and political influence of different groups; and 'private-interest' or

[6] For more detailed discussion of price control mechanisms see Chapter 17 below.

[7] See e.g. Breyer, *Regulation and Its Reform*; R.B. Stewart, 'Regulation and the Crisis of Legalisation in the United States' in T. Daintith (ed.), *Law as an Instrument of Economic Policy* (Berlin, 1998); id., 'The Discontents of Legalism: Interest Group Relations in Administrative Regulation' (1985) *Wisconsin Law Review* 685; E. Bardach and R. Kagan, *Going by the Book: The Problem of Regulatory Unreasonableness* (Philadelphia, 1982); Gunningham and Grabosky, *Smart Regulation*, 41–7.

[8] See C. Hood, *Explaining Economic Policy Reversals* (Buckingham, 1994), 21.

[9] For a review of these, see B. Mitnick, *The Political Economy of Regulation* (New York, 1980); also see P.J. Quirk, *Industry Influence in Federal Regulatory Agencies* (Princeton, 1981); G. Wilson, 'Social Regulation and Explanations of Regulatory Failure' (1984) 32 *Political Studies* 203.

[10] See M.H. Bernstein, *Regulating Business by Independent Commission* (New York, 1955).

economic analyses see regulation as a commodity liable to fall under (or to be established under) the sway of the economically powerful.[11]

The proximity of regulator to regulatee relationships that is associated with C & C techniques might be thought to be particularly conducive to capture in so far as agencies, when drawing up and enforcing rules, must rely to some extent on the cooperation of the regulated firms. Thus, the argument runs, regulators require a good deal of information in order to carry out their functions—say to fix appropriate standards on issues such as acceptable pollution levels or price increases. The primary, and best, source of such information will often be industry. The regulator, accordingly, requires some assistance from the regulated firms in order to make C & C regulation work. This gives the regulated firms a degree of leverage over regulatory procedures and objectives, a leverage that, over time, produces capture.

In response to allegations that C & C regulation is particularly prone to capture, it should be noted that many versions of capture theory would attribute capture to factors that operate in a manner unaffected by the particular regulatory technique employed. They might point, for instance, to broad political, institutional, or economic considerations.

LEGALISM

A second major concern with C & C regulation has been its alleged propensity to produce unnecessarily complex and inflexible rules, and indeed, a proliferation of rules that leads to over-regulation, legalism, delay intrusion on managerial freedoms, and the strangling of competition and enterprise.[12] Eugene Bardach and Robert Kagan have expressed concern at the extent to which US regulators have tended to over-regulate with over-inclusive rules (rules that apply to an unnecessarily wide array of instances or actions) and have given a number of reasons why such problems tend to occur. First, rule-makers find it very difficult to design precisely targeted rules (the informational demands are severe) and the tendency is to avoid such design and drafting difficulties by writing over-inclusive rules. Second, for political reasons, regulators tend to respond to particular problems or tragedies with general, or 'across-the board', rules and solutions. This gives the appearance of 'doing something about that sort of thing'. Third, pressures to reduce discretions in favour of the 'rule of law' (so as to make regulatory actions rule-governed) may come from politicians, those regulated, or consumers, and

[11] See R. Posner, 'Theories of Economic Regulation' (1974) 5 *Bell Journal of Economics* 335; G. Stigler, 'The Theory of Economic Regulation' (1971) 2 *Bell Journal of Economics* 3.

[12] See Stewart, 'Regulation and the Crisis of Legalisation'; Bardach and Kagan, *Going by the Book*; G. Teubner, *Juridification of Social Spheres* (Berlin, 1987); R. Harris and S. Milkis, *The Politics of Regulatory Change* (New York, 1989).

these pressures may induce the excessive production of rules. Fourth, regulators often wish to respond to a mischief before public concern dies down—while the memory of the disaster is still fresh. Working to the resultant short time-scales tends to produce rules that are broad-brush, rather than precisely targeted. Finally, there is what is dubbed the 'regulatory ratchet',[13] whereby regulatory rules tend to grow rather then recede because revisions of regulations are infrequent; work on new rules tends to drive out attention to old ones; and failure to carry out pruning leads the thickets of rules to grow ever more dense.[14]

In the context of British telecommunications, it has been argued that detailed, prescriptive rules can be a barrier to entry, can inhibit competition, and can discriminate between incumbent licensed operators and new entrants. When it was regulator, OFTEL urged a movement away from control by means of detailed rules contained in the licences of those given privileged access, towards 'open state' regulation that is based on general authorizations and which gives a stronger role to general competition and consumer protection laws, backed up by detailed guidance only where necessary.[15]

STANDARD-SETTING

Setting appropriate standards has been argued to pose major difficulties for regulators because the informational demands are so severe.[16] Thus, anti-competitive effects must be addressed; the appropriate *type* of standard must be selected—be this an output standard specifying a level of performance or an input standard calling for a particular design or specification of operation or machinery—and the level of exposure to judicial review may be high.[17] Setting the appropriate level of performance is, moreover, technically difficult and liable to be contentious. To give a simple instance, employing the example of pollution again, even if it is assumed that the regulator knows the beneficial values of particular levels of cleanliness in a river, and is clear on social objectives, setting the optimal level of allowable pollution (the level that minimizes the sum of abatement and damage costs) would require data on the differing abatement costs of all of the various polluters on the riverbank. The efficient level of pollution will, indeed, be specific to each enterprise, yet

[13] Bardach and Kagan, *Going by the Book*, ch. 7.

[14] On responses to these problems, see R. Baldwin, *Rules and Government* (Oxford, 1995), 183–5 and below, Chapter 14.

[15] See OFTEL, *Second Submission to the Culture, Media and Sport Select Committee: Beyond the Telephone, the Television and the PC—Regulation of the Electronic Communications Industry* (London, March 1998). Ofcom took over OFTEL's functions in December 2003.

[16] See Chapter 14 below for a general discussion of standard-setting, also Breyer, *Regulation and Its Reform*, 109–19; Ogus, *Regulation*, ch. 8.

[17] See Ogus, *Regulation*, ch. 8.

the regulator has usually to produce a generalized across-the-board rule. The result will be a broadly inefficient regime, with some enterprises finding it hugely expensive to meet the standard and others able to go better than standard at very little cost but given no incentive to do so.[18] A further worry, moreover, is that commands, and the standards they mandate, may prove unresponsive to changes in technologies, risks and other regulatory challenges so that, even if appropriate today, they may not deliver the right solutions tomorrow.

ENFORCEMENT

A final major difficulty said to be particularly associated with C & C regimes is that of enforcement. The complex rules attending such regimes have to be brought to bear on the ground by bodies of officials or inspectors, but enforcement is expensive, the techniques used give rise to contention, and the effects of enforcement are said to be uncertain.[19] On the latter point, for instance, the rules used in C & C systems may be too narrow or too broad in scope. They may, accordingly, fail to cover conduct that should be controlled, or else may constrain activity that should be unrestricted. In addition, there may be problems of 'creative compliance'—the practice of avoiding the intention of the law without breaking the terms of the law.[20] Command methods may also lack force when court sanctioning is weak and the rules, as a result, fail to pose a credible deterrent. Where, moreover, there is credibility of deterrence, there may be problems because adversarial industry to regulator relationships develop and this produces poor information flows to the regulator and a climate of defiance and resistance that produces poor compliance.[21]

Regulators employing C & C techniques thus face substantial difficulties of rule use. Not only must the rules employed be capable of enforcement and be accessible to regulated firms or individuals, but the appropriate types and levels of standards must be fixed, problems of scope (or inclusiveness) must be overcome, and issues of creative compliance dealt with. Such problems, moreover, must often be faced in political environments that are unlikely to produce the resources necessary for effective enforcement and are hostile to rules that impose compliance costs on industry or interfere with managers. In the light of such difficulties, some commentators have advocated a move

[18] See C. Sunstein, 'Paradoxes of the Regulatory State' (1990) 57 *University of Chicago Law Review* 407–41.

[19] For further discussion of enforcement, see Chapter 11 below and Baldwin, *Rules and Government*, ch. 6.

[20] See below, p. 232.

[21] Gunningham and Grabosky, *Smart Regulation*, 45; J. Braithwaite and T. Makkai, 'Trust and Compliance' (1994) 4 *Policy and Society* 1.

away from command-based strategies towards alternative, 'constitutive', 'less restrictive', or 'incentive-based' styles of control.[22] On the governmental stage also, numerous administrations and international bodies came to favour this kind of shift at around the end of the second millennium.[23] The strategies now to be described may be seen as the main alternatives to the C & C style of regulation—they are state-initiated but, as will be seen, may delegate regulatory functions in varying degrees to non-state actors.

Incentive-based Regimes

Regulating by means of economic incentives might be thought to offer an escape from highly restrictive, rule-bound, C & C regimes.[24] According to the incentives approach, the potential mischief causer, say a polluter, can be induced to behave in accordance with the public interest by the state or a regulator imposing negative or positive taxes or deploying grants and subsidies from the public purse. Thus, not only can taxes be used to penalize polluters, but rewards can be given for reductions in pollution, or financial assistance can be given to those who build pollution-reducing mechanisms into their production or operational processes. An example of such an incentive strategy at the broadest level was the differential tax on leaded and unleaded petrol that was introduced into Britain in 1987.[25]

The posited advantages of such schemes are numerous. They are, for instance, said to involve relatively low levels of regulatory discretion (as compared to C & C systems) because financial punishments or rewards operate in a mechanical manner once the regime is established. These low

[22] See Stewart, 'Regulation and the Crisis of Legalisation'; Breyer, *Regulation and Its Reform*. On incentive-based regulation see Ogus, *Regulation*, ch. 11. For a European view of the limits of command law, see G. Teubner, *After Legal Instrumentalism? Strategic Models of Post-Regulatory Law*, EUI Working Paper No. 100 (Florence, 1984).

[23] See, e.g., the UK Better Regulation Task Force, *Imaginative Thinking for Better Regulation* (London, 2003); OECD, Report on Regulatory Reform (Paris, 1997), vol. 2, pp. 193–202.

[24] See Ogus, *Regulation*, ch. 11; Daintith, 'Techniques of Government'; R. Breyer and R.B. Stewart, 'The Discontents of Legalism: Interest Group Relations in Administrative Regulation' (1985) *Wisconsin Law Review* 685. On the limitations of incentive-based regimes, see J. Braithwaite, 'The Limits of Economism in Controlling Harmful Corporate Conduct' (1982) 16 *Law and Society Review* 481.

[25] See generally: A. Ogus, 'Corrective Taxation as a Regulatory Instrument' in C. McCrudden (ed.), *Regulation and Deregulation* (Oxford, 1999); A. Ogus, 'Corrective Taxation and Financial Impositions as Regulatory Instruments' (1998) 61 *MLR* 767; S. Rose-Ackerman, 'Efficient Changes: A Critique' (1973) 6 *Canadian Journal of Economics* 572; W.J. Baumol, 'On Taxation and the Control of Externalities' (1972) 62 *American Economic Review* 307; P. Burrows, 'Pricing versus Regulation for Environmental Pollution' in A.J. Culyer (ed.), *Economic Policies and Social Goals* (London, 1974); D. Driesen, 'Alternatives to Regulation? Market Mechanisms and the Environment' in R. Baldwin, M. Cave, and M. Lodge (eds), *The Oxford Handbook of Regulation* (Oxford, 2010).

levels of discretion and structured modes of application reduce the dangers of regulatory capture in so far as regulators are not involved in constant negotiations, close relations, and information exchanges with regulatees as in the usual C & C scheme.

They are also said to leave managers free to manage. It is up to the regulated firm, not the bureaucrat or regulator, to balance the costs of polluting against those of abatement in a particular context and to devise means of reducing the mischief most efficiently. Managers are, accordingly, able to be more flexible concerning their modes of production than most C & C regimes allow.

Incentive-based regimes are, additionally, claimed to be cheaper to administer than commands[26] and to involve relatively light burdens of information collection and compliance costs. They, moreover, are said by proponents to encourage individual regulated firms to reduce harmful conduct as much as possible (to give an 'incentive to zero'), not merely down to the level that is demanded by the standard stipulated in a C & C regime—a standard liable, in any event, to be fairly lax because C & C regulators tend, for political reasons, to have to set a general standard soft enough to be met by poorer performers in the industry without causing financial crises or unacceptable unemployment.

The advantages of incentive regimes can, however, be exaggerated and a number of cautionary points should be borne in mind.[27] Such systems often have to be put into effect by means of highly complex systems of rules (the field of taxation, for instance, is not one renowned for simplicity).[28] Many of the problems associated with C & C regulation might thus be replicated in putting such systems into effect on the ground. Inspection and enforcement mechanisms might, moreover, have to be employed to prevent regulatees evading their liabilities (e.g. to taxes). The system might, thus, come to resemble C & C regulation and the distinction between incentives and penalty mechanisms might be less than first appeared.[29] As for overall costs, it cannot be assumed that these will be lower under a taxation as opposed to a command regime. In the former, the task of determining optimal abatement may be thrust on to regulated parties in circumstances where this involves higher costs than those that would have been incurred by a public regulator.[30]

Proponents of incentive systems tend to assume that those regulated operate, on the whole, in an economically rational manner. In practice,

[26] Baumol, 'On Taxation and the Control of Externalities', 307.

[27] For evaluation, see Ogus, *Regulation*, 250–6 and Breyer, *Regulation and Its Reform*, 278–80.

[28] See R.S. Markovits, 'Antitrust: Alternatives to Delegalisation' in G. Teubner (ed.), *Juridification of Social Spheres* (Berlin, 1987).

[29] See Bardach and Kagan, *Going by the Book*, chs. 8, 9, and 10; J. Braithwaite, 'The Limits of Economism in Controlling Harmful Corporate Conduct' (1982) 16 *Law and Society Review* 481.

[30] See Ogus, 'Corrective Taxation and Financial Impositions', 776–7.

however, many problems (e.g. hazards in the workplace) are the product of irrational, accidental, or negligent behaviour.[31] Incentive mechanisms may, accordingly, influence responsible parties more effectively than irresponsible, careless, or ill-informed individuals or firms—yet it is the latter group who are most in need of regulation. Regulatory lag may also prove a significant problem with incentive regimes because they operate indirectly. Thus, within a firm the effects of tax incentives may have to be transmitted from finance directors through operations managers to floor staff and this, even if successful, may take some time—the fish in the river may long be dead. Incentives may thus prove to be poor regulatory tools where periodic crises occur in the sectors involved, where such sectors are subject to rapid economic change, or where preventative measures need to be taken, rather than harmful effects penalized.

A core difficulty with incentive regimes may be predicting the effect on the ground of a given incentive. To continue the river pollution example, it will be very difficult to predict how much a certain level of taxation will clean up the river—or whether certain thresholds will be passed (e.g. pollution will rise to levels that the fish cannot survive). The effect on each firm sited on the river will differ and (assuming firm rationality) will depend, *inter alia*, on the profit derived within each production process from each unit of pollution. Fixing incentive levels in order to achieve acceptable outcomes may thus make informational demands at least as severe as those encountered within C & C regimes. It might be responded to this point that, in practice, the tax authorities could adopt a trial-and-error approach so that, in the river example, tax levels could be modified in response to water cleanliness readings and the desired purity could be arrived at over time.[32] The difficulty with such a solution, however, is that a process of incremental adjustment might work to control small risks but would not be acceptable if the regulated risks were potentially catastrophic. Thus, if the river contained the last breeding stock of a rare fish, it would be difficult to justify operating on a trial-and-error basis and running the risk of under-deterrence and fish kill through sub-optimal taxation.

The mechanical application of incentives may also bring disadvantages. Within C & C systems, enforcement can be used flexibly in an effort to achieve desired results and to limit the imposition of restrictions on particular firms or individuals where unduly onerous effects would result. In so far as incentive regimes operate mechanically, such tailoring to individual circumstances will not be possible. If a flexible and discretionary approach is adopted in relation to incentives (and there is no reason why this cannot be the case),

[31] See Braithwaite, 'Limits of Economism'.

[32] W. Baumol and W. Oates, 'On Taxation and the Control of Externalites' (1972) 62 *American Economic Review* 307.

another supposed difference from, and advantage over, C & C regulation falls away.

Presentationally and politically, a move from C & C towards incentive regimes may prove popular with firms regulated (especially where subsidies are offered), but public concern may arise on the grounds that socially harmful activity is not being stigmatized or condemned and that a licence is being given for undesirable behaviour.[33] Subsidies may be objected to as making payments from the public purse to those engaged in offensive conduct and negative incentives or taxes may be criticized not only for their failure to designate certain acts as unacceptable but also for taking away from industry the very resources that might have been committed to measures aimed at avoiding the undesirable consequences of their actions (e.g. to filtration systems).

As far as democratic accountability and access to the regulatory process are concerned, similar consultative and other procedures to those used in command and control regulation may be used. If it is hard to predict the effects of given incentives on the ground, however, it may be difficult to produce the results that such democratic inputs favour, and this can be seen as a weakening of accountability and access. Other accountability concerns may relate to the distributional effects of taxes (the question of who eventually pays[34]) and the degree to which complex tax rules tend to offer well-resourced regulated firms better access to rule-making processes than can be enjoyed by individual harm sufferers and small public interest groups.[35]

Market-harnessing Controls

COMPETITION LAWS

A direct method of regulating by channelling market forces is to influence competition within an area. Competition laws can thus be used instead of, or in conjunction with, regulation in order to sustain such levels of competition as will ensure that the market provides adequate services to consumers and the public.[36]

[33] See Ogus, *Regulation*, 225; also W. Beckerman, *Small is Stupid: Blowing the Whistle on the Greens* (London, 1995).

[34] See Ogus, 'Corrective Taxation and Financial Impositions', 775—who notes the regressive effects of some taxes and the difficult issue of how some corporations' tax costs are subsequently distributed between shareholders, employees, and consumers.

[35] Ibid., 786.

[36] On competition law generally, see R. Whish, *Competition Law* (6th edn. Oxford, 2008). On regulation versus competition, see below, Chapter 23.

Such laws can also be used to control market behaviour so as to prevent anti-competitive or unfair practices such as 'predatory pricing' by dominant operators (setting prices for one's products below cost in order to drive competitors from the market)[37] or effecting cross-subsidies from monopolistic to competitive sectors.

The telecommunications industry provides an example of competition law being used instead of classical C & C regulation. Thus, in contrast to the UK's use of a sectoral agency (OFTEL, later superseded by Ofcom) with sector-specific rules, the New Zealand government, on privatizing in the late 1980s, relied on general competition laws, applied in the courts, as a mechanism for influencing the telecommunications industry.

The broad advantages of reliance on competition laws are that they can be applied across the board to different sectors, the need for industry-specific regulation is avoided, and barriers to entry may be lower than in regimes incorporating large numbers of highly prescriptive rules. Consistent principles can also be developed across sectors and there are economies of scale in applying rules broadly.

Competition laws produce lower levels of intrusion into firms' internal decisions than are involved in C & C regimes, and flexibility in the industry tends to be greater under competition law regimes than in cases where behaviour is structured by an overseeing agency. Finally, enforcement involves relatively light burdens on the public purse because it depends on private actions in courts, rather than interventions by publicly funded regulatory agencies. Experience in New Zealand telecommunications suggests, however, that a number of drawbacks can be encountered when heavy reliance is placed on competition laws.[38] The broad principles established in competition laws may, for instance, not provide solutions to operational, technical, or commercial problems. Such issues are left to the parties to resolve in the courts and more effective solutions might, under certain conditions, be produced by a specialist overseeing agency. An agency, moreover, might develop and apply a greater level of expertise than the parties or the courts in dealing with such issues as the economics of interconnections. Guidelines established by a regulatory agency can reduce uncertainties and transaction costs for operators more efficiently than competition laws or the courts.

The courts system may, furthermore, be slow to develop guidelines on central industrial issues. Thus, following difficulties concerning the application of general competition rules to a dispute over interconnection by a new

[37] See J. Vickers, 'The Economics of Predatory Prices' (1985) 6 *Fiscal Studies* 24.

[38] For reviews of New Zealand experience, see: New Zealand Commerce Commission, *Telecommunications Industry Inquiry Report* (Wellington, June 1992); C. Blanchard, 'Telecommunications Regulation in New Zealand: How Effective is "Light-Handed" Regulation?' (1994) 18 *Telecommunications Policy* 154–64.

entrant (issues fought from New Zealand to the Privy Council in 1994[39]), the New Zealand government considered whether a new mix of institutions and rules would be appropriate.[40] One difficulty encountered in relying on judicially developed principles on such issues as interconnection is that rulings only emerge as cases happen to arise. Principles, accordingly, may develop sporadically, slowly, and may leave key issues untouched. Developing such principles, moreover, may involve asking the courts to stand in the shoes of business people and to make business decisions.[41] Evidential problems may also compound such reliance on the courts, thus, competition law may have a limited role in dealing with entry barriers where it is difficult to show these have been established on purpose by a dominant undertaking.

To point to some of the problems to be anticipated in using competition laws is not, of course, to say that such laws cannot play a very useful role in combination with other mechanisms of influence, such as C & C regulation in the classical style. Competition laws can thus substitute for excessively prescriptive C & C regulation on some issues and the latter can be used to impose structures and final solutions for industries in circumstances where competition law would be slow to provide answers on these fronts.

FRANCHISING

Franchising is a system of control that can be employed in naturally monopolistic sectors to replace competition *in* the market with competition *for* the market. It has been employed notably in the British independent television, radio, and rail industries. The underlying idea is that if applicants for franchises make competitive bids for an exclusive (or at least protected) right to serve a market for a given period and under conditions, they will bid on assumptions of efficient operation and, as a result, consumers will benefit—they will be served by operators who are not under immediate competitive pressure but who will behave in many ways as if they are. A fuller discussion of franchising is offered in Chapter 9.

REGULATION BY CONTRACT

Government departments or agencies can use the state's wealth and spending power to achieve desired objectives by specifying these in the contracts it agrees with enterprises. It can be stipulated, for example, that parties contracting to supply goods or services shall pay their own employees a

[39] See *Clear Communication* v. *New Zealand Telecommunications Corp.* [1994] 6 TCLR 138 (1995) 1 NZLR 385 (PC).

[40] See Baldwin, Scott, and Hood, *A Reader on Regulation*, ch. 1.

[41] See New Zealand Commerce Commission, *Telecommunications Regulation*, 83.

minimum wage.[42] The regulatory aspects of the contract may be incidental to the main purpose, which may be commercial, but the effect is to impose a regulatory standard across all firms contracting with the government. There is no need for a command base. A form of contracting out—Compulsory Competitive Tendering (CCT)—of local authority services has been used by government as a means of reducing service costs, and it brings with it local authority regulation of those who provide services under contractual terms. In some sectors, similarly, dependence on public funding has been used as a basis for encouraging both the development of self-regulation and the imposition of 'consensual forms of regulation'.[43]

TRADABLE PERMITS

A further technique that seeks to harness markets is the use of tradable permits to engage in an activity that has been deemed to require control (e.g. discharging pollutants into a water course).[44] Like franchising, the strategy can be used to control both entry into the market and subsequent behaviour within the market. Examples of the use and advocacy of tradable permits are to be found. Thus, since 1991 the US Environment Protection Agency (EPA) has sought to control sulphur dioxide emissions by allocating tradable emission permits to coal-burning electric power plants[45] and the EU launched its Emissions Trading Scheme in January 2005. By 2007, the Stern Review had advocated the broad use of trading mechanisms to combat climate change.[46]

Emissions trading is discussed in greater detail in Chapter 10 below, but a sketch of issues will be given here. In typical regimes, the public agency issues a given number of permits and each of these allows a specified course of behaviour (e.g. a polluting discharge of a fixed amount). Following the initial allocation, permits may be traded and this allows, say, a generating company to switch to cleaner fuels and sell its excess allowances to other firms. The initial distribution of permits may be carried out by auction or according to

[42] See T.C. Daintith, 'Regulation by Contract: The New Prerogative' (1979) *Current Legal Problems* 41. On governing through contracts, see I. Harden, *The Contracting State* (Buckingham, 1992) and N. Lewis and J. Goh, *The Private World of Government* (Sheffield, 1998).

[43] See Baldwin, Scott, and Hood, *A Reader on Regulation*, ch. 1; M. Cave, R. Dodsworth, and D. Thompson, 'Regulatory Reform in Higher Education in the UK: Incentives for Efficiency and Product Quality' in M. Bishop, J. Kay, and C. Mayer (eds), *The Regulatory Challenge* (Oxford, 1995).

[44] For a review of market-based instruments versus other regulatory tools, see: N. Keohane, R. Revesz, and R. Stavins, 'The Choice of Regulatory Instruments in Environmental Policy' (1998) 22 *Harvard Environmental Law Review* 313–67.

[45] D. Ellerman, R. Schmalensee, E. Bailey, P. Joskow, and J.-P. Montero, *Markets for Clean Air* (Cambridge, 2000).

[46] N. Stern, *The Economics of Climate Change* (Cambridge, 2007).

public interest criteria. The incentives within such systems are provided by the market in permits.

Advantages claimed for the strategy are, first, that permits can be allocated to those who will generate most wealth per unit of pollution. This is because those willing to pay most for the permits will be those who derive the most profit from polluting—in this sense, it can be argued, (at least on a set of not uncontentious assumptions), that the pollution is being put to the use that society values most. Second, the incentive to reduce harmful behaviour can, as in taxation regimes, operate down to zero, since the process of abatement will release permits for resale until the point where no harm is being done at all. Third, managers, again, are less restricted than in C & C regulation because they are free to decide whether and how to reduce harmful conduct in order to release permits. Fourth, regulatory discretions (and dangers of capture) are kept low because markets rather than bureaucrats are imposing restraints, and, finally, regulatory costs are low since, once established, the market in permits runs on its own accord.

The problems to be anticipated in relation to schemes with marketable permits are, however, numerous. Enforcement still has to be carried out to prevent non-permit holders from creating harms and to stop permit holders from exceeding the terms of their permits. Inspectorates, accordingly, require funding. Regulatory lag may also be a problem. If, for example, permits are used to control river pollution, it may be difficult to adjust pollution levels rapidly so as to cope with sudden drops in the river's capacity to absorb pollution (as might occur in a heatwave or drought). The difficulty is that permits are already issued, they are in the marketplace and bearing a given entitlement. (A response to the difficulty might be to give permits a floating entitlement that is adjustable by the regulator. This would give flexibility but might prejudice the operation of the market and would impose severe informational demands on the regulator.)

Permits, moreover, do not provide the resources needed to compensate the victims of harmful conduct and, politically, permits may create difficulties with electorates, since they may be seen as 'licences to pollute'. The system, in addition, demands that there be a healthy market in permits—which calls for such factors as a large number of potential buyers possessed of adequate information. If the market is deficient (perhaps because of uncertainties or lack of information), the value of permits may be low and the incentives to desist from harmful conduct may be weak. A further problem is that markets in permits may allow hoarding and the creation of barriers to enter into certain markets. This will be more likely where conditions favour collusion between certain large firms. The effects may be generally anti-competitive and may be unfair to less well-resourced firms. As for the areas where markets in permits can be used, some harms or pollutants may have to be prohibited absolutely and, accordingly, the tradable permit system will

be inappropriate. Finally, it should be cautioned that democratic accountability and influence may be low once the system is up and running, since the market (and its degree of genuine competitiveness) will govern the price to be placed on pollution. Where markets are imperfect, it is also likely that information flowing into the public domain is below optimal levels.

Disclosure Regulation

Structuring the disclosure of information provides a mode of regulation that is not heavily interventionist. It does not regulate the production process, the level of output allowed, prices charged, or the allocation of products. Disclosure rules usually prohibit the supply of false or misleading information and may also require mandatory disclosure—perhaps obliging suppliers to provide information to consumers on price, composition, quantity, or quality (familiar demands in the food and drinks sectors).[47] Disclosure regulation may also involve the supply of information to the public directly by a scrutinizing regulator or governmental official. Thus, in October 1997, the then Agriculture Minister, Jack Cunningham, first put into action a policy of 'naming and shaming' food manufacturers who failed to comply with regulations on safety, product quality, and authenticity. Following a departmental survey, the Minister named sixteen pork and bacon brands as guilty of failing to declare the added water content of their products. These included suppliers of Tesco and J. Sainsbury.[48] More recently, the Environment Agency published details of the ten firms who had been fined the highest sums following its prosecutions in 2007.[49] 'Naming and shaming' is not, however, the only reason for state disclosure. Governments may also disclose information for exhortatory reasons (e.g. health campaigns), in order to raise standards by drawing attention to best practices in a field or to rank service providers in 'league tables' (as with schools).[50]

Disclosure regulation allows the consumers of products and services (or even voters more generally) to make decisions on the acceptability of the processes employed in producing those products or services. To rely on consumer or

[47] In the food sector there is a pressure group devoted to disclosure—the Food Labelling Agenda (FLAG). See generally K. Yeung, 'Government by Publicity Management: Sunlight or Spin?' (2005) *Public Law* 360–83. On voluntary disclosure by producers for 'ethical branding', marketing, and other reasons, see Yeung, loc. cit.

[48] See *Financial Times*, 29 Oct. 1997: '"Naming and Shaming" over Pork Product Labels'.

[49] Environment Agency, *Spotlight on Business* (Bristol, 2008), 22. The Health and Safety Executive also publishes details of convictions secured.

[50] The Major government instituted state-sponsored school league tables in the 1990s—see Yeung, 'Government by Publicity Management'.

voter preferences in this manner does, however, restrict the potential of disclosure as a regulatory instrument.

The main problems to be anticipated are, first, that users of the information disclosed, be they consumers or other citizens, may make mistakes; they may fail to use the information properly; fail to understand the implications of the data given; mis-assess risks; neglect to collect the full range of relevant information; lack the resources and expertise to research issues fully; and so may come to harm. Second, information users may not respond in anticipated ways to the flow of information. Considerations of economics rather than policy, politics, or social concern may shape their decisions. Thus, consumers, when purchasing products, may choose according to price, rather than other factors. They may, for instance, buy cheap products without responding to information suggesting that dangers are involved in consumption or that production of the goods involves a host of socially undesirable consequences (e.g. discharges of polluting effluents).

Third, the costs of producing the information may be excessive, as may the costs of processing it. Thus, if information disclosure rules were employed instead of C & C regulation in relation to food safety, a visit to the supermarket would involve a very lengthy process of scrutinizing labels. It might, in many circumstances, be far more efficient for consumers to rely on the expertise and protection of public regulators and inspectorates, rather than depend on their own individual assessments of risks.

Fourth, the risks associated with some products or activities may be so great that policymakers may feel that it is inappropriate merely to inform affected parties about these matters and C & C methods may be deemed necessary.[51] Fifth, where information regulation is employed there is always a danger that the information will be inaccurate and unjustifiable claims made. Policing of the quality of information will, accordingly, be necessary. This increases the costs of information-based regulatory regimes. Finally, standards may have to be applied to various items of information so that affected parties may make appropriate use of any data given. In the absence of such standards, information may be offered in a manner that does not assist, for example, consumers. Thus 'may cause cancer' is a phrase that discloses little concerning the size of any risk of cancer generated by using the product.

Given the above limitations of disclosure regulation, the case for the strategy is liable to be strongest where: the hazard involved is not potentially catastrophic or the difference between high- and low-quality products or processes is not likely to give rise to grave consequences; the relevant information can be processed at a reasonable cost; risks can be assessed accurately by affected parties; consumers of the products at issue, or other affected

[51] See I. Ramsey, *Consumer Protection* (London, 1989).

parties, can be relied upon to give proper consideration to the information given; and the accuracy and utility of information can be monitored and ensured through enforcement at acceptable cost. It can also be argued that, even where information strategies cannot be used as free-standing replacements for traditional command methods—as in environmental protection—the two approaches can be used as complementary instruments.[52] There is evidence that, at least where clear standards and credible penalty systems are found, public disclosures can create additional and strong incentives for compliance in such areas as pollution control.[53]

Direct Action and Design Solutions

Governments can use their resources to achieve desired results by taking direct action. Rather than set and enforce standards on, say, dust extraction levels in factories, central governments or local authorities can build properly ventilated premises and lease these to private manufacturers. Public ownership of infrastructure can, moreover, be combined with the franchising out of operations (leasing for fixed periods subject to conditions on use and renewal would produce similar results). Long-term investments can, by such methods, be rendered amenable to planning by government, and the replacement of unsatisfactory operators can be facilitated. Thus, in London the bus transport network is publicly owned, but routes are put out to competitive tendering or franchising.[54]

An advantage of direct action is that public money can be used to ensure furtherance of democratically established objectives in circumstances where firms, particularly small ones, might not invest in the required measures. A degree of subsidization may, by such means, be effected and public resources can be used to assist firms to reduce harms rather than to fund C & C enforcement regimes or to apply penalties that take money away from the enterprises that are asked to spend on avoiding undesired consequences.

Such subsidization, however, may give rise to distributional issues—concerning the fairness of access to subsidized premises, for instance—and subsidies may produce undesirable distortions of competition. An equally difficult problem may be that the public funding of a certain aspect of a production process may encourage firms to build operations around the

[52] See Gunningham and Grabosky, *Smart Regulation*, ch. 6 (by N. Gunninghan and D. Sinclair).

[53] See J. Foulon, P. Lanoie, and B. Laplante, 'Incentives for Pollution Control: Regulation or Information?' (2002) 44 *Journal of Environmental Economics and Management* 169–87.

[54] See S. Glaister, D. Kennedy, and T. Travers, *London Bus Tendering* (London, 1995) and S. Glaister, *Deregulation and Privatisation: British Experience* (Washington, DC, 1998).

funded element. As a result, innovation may not be driven by the market and the enterprises' responsiveness to markets and potential new technologies or processes may be blunted. Thus if the well-ventilated manufacturing premises are publicly owned and there are no other controls on dust levels in the air, there is little incentive for the private sector to devise new, more efficient ways to control dust. The manufacturers of dust extraction systems, for example, would be potentially selling their new designs to the procurement departments of public bodies, rather than to private firms. The incentive to innovate would, accordingly, be far weaker than under a regime of taxing dust exposures—which would lead companies to press extraction manufacturers for ever better ways of reducing dust and tax liabilities.

Finally, it should be noted that the 'direct action' approach tends to assume, perhaps unrealistically, that where the state provides a solution, this will remove the targeted mischief unproblematically. The reality, however, may be that the state may fall down on its ongoing obligations just as badly as the private sector. It cannot be assumed, for instance, that, in the above example, the state's dust extraction systems will be perfectly maintained and effective over time. Public bodies' failures to renew filters and maintain machinery may be as pronounced as those of private firms.

A different way that the state can use its resources to eliminate problems is through the use of design solutions. Thus, rather than regulate the mischief, the state can organize affairs so that the mischief cannot arise—or opportunities for the mischief to eventuate are minimized. It can 'design out' problems in a variety of ways. These include constructing the physical environment in a certain manner—as where parking is controlled by concrete bollards or road accidents are reduced by a road architecture that makes speeding impossible. The law can also be used for such design purposes—as where statutes set up markets in a configuration that ensures healthy competition and consumer satisfaction.[55] The labels of 'techno-regulation', 'architecture-based', and 'code' approaches are attached to such design strategies and some commentators make a case for dealing with some of the most daunting regulatory challenges in this way. Thus Lessig suggests that it is possible to regulate cyberspace through control of the software code that shapes the structure of cyberspace and dictates access to and participation in that space.[56] This could be done, he argues, by mandating software designers to build certain elements into software code in pursuit of public regulatory objectives.[57] The degree to which regulatory actors can escape such

[55] See R. Brownsword, 'Code, Control and Choice: Why East is East and West is West' (2005) 25 *Legal Studies* 1–21; D. Garland, *The Culture of Control* (Oxford, 2001); B. Morgan and K. Yeung, *An Introduction to Law and Regulation* (Cambridge, 2007), 102–5.

[56] See L. Lessig, *Code and Other Laws of Cyberspace* (New York, 1999).

[57] L. Lessig 'The Law of the Horse: What Cyberlaw Might Teach' (1999) 113 *Harvard Law Review* 501, 514–22.

architectural/code controls may, however, prove a point of contention—as may the extent to which such controls need to be combined with other types of regulatory instrument. These matters are explored in some detail in the extensive scholarly debates that relate to cyberspace and its control.[58]

NUDGE STRATEGIES

A regulatory strategy that purports to offer a user-friendly and a low-intervention alternative to more draconian controls is 'nudging'. This approach is highly influential in many government circles following the publication of Thaler and Sunstein's 2008 book *Nudge*.[59] Nudging involves structuring the architecture of decisions (so-called 'choice architectures') so that it is easier for consumers or others (such as regulatees) to act in ways that are beneficial. Studies in the fields of decision-making[60] and behavioural economics suggest that people tend to make poor choices for a number of reasons that Thaler and Sunstein identify (they process information in shorthand ways that are biased by immediate concerns and experiences, they tend to be too optimistic, and so on).[61] Nudging makes it easier to make the sensible decision but, according to a philosophy entitled 'libertarian paternalism', it purports to leave the target person or firm free to choose to take the non-sensible course of action. An example of nudging is establishing a presumption that all citizens consent to be organ donors unless they register their unwillingness to donate (which, Thaler and Sunstein stress, they should be able to do easily).

The nudging approach thus allows for decisions to be manipulated by public authorities, provided that it leaves decision-makers free to choose to behave as they, rather than the public authorities, see fit. In its ideal form, therefore, the approach cleverly combines an element of paternalism with the preserving of freedom of choice. It also offers the hope of using small changes in choice architectures to achieve considerable changes in outcomes.

Critics, however, might have two central worries about nudging. The first is that it is difficult, in real-life situations, to draw the line between manipulations that do not threaten freedoms of choice and those that do. Thaler and

[58] See A. Murray, *The Regulation of Cyberspace: Control in the Online Environment* (London, 2006); A. Murray and C. Scott, 'Controlling the New Media: Hybrid Responses to New Forms of Power' (2002) 65 *Modern Law Review* 491–516.

[59] R. Thaler and C. Sunstein, *Nudge: Improving Decisions about Health, Wealth and Happiness* (New Haven, 2008). Warning: the thesis of *Nudge* is expressed amidst a mass of personal anecdotes that some readers may find frustrating and exhausting. See also R.H. Thaler, C.R. Sunstein, and J.P. Balz, 'Choice Architecture' (2010). Available at SSRN: http://ssrn.com/abstract=1583509 (last accessed 7 December 2010).

[60] See, e.g., H. Simon, *Administrative Behaviour* (1947; 4th edn, New York, 1997).

[61] Thaler and Sunstein, *Nudge*, ch.1.

Sunstein give examples of easy cases in which consumers are confronted with helpful information on products but, in other circumstances, the rigging of the decision architecture may make 'non-sensible' choices (as seen by the nudger) quite difficult to take. Thaler and Sunstein, as noted, would respond that opt-outs must be easy but such assurances count for little if there is no reliable way to identify and protect the easy opt-out. They contend that it would be 'ridiculous' to have an inflexible rule on when opt-out costs are too high: 'the precise question of degree is not important. Let us simply say that we want these costs to be small'.[62] Critics, however, are liable to say that this is too easy a response. 'Ease of opt-out', they would stress, is a contentious issue that lies at the heart of nudging and the proponents' answer evidences the dangerousness of nudge: it treats the centrally important issue of opt-out feasibility as a small and relatively uncontentious matter. This approach, the objectors would say, sows the seeds of an illiberal system of control.

The second main worry is related and is that the processes of nudging are value-laden yet low in transparency. It might, thus, be contended that whether a nudged-for outcome is 'good' or 'beneficial' is not always obvious. The evaluation of an outcome's merits may reflect the nudger's conception of the good rather than the nudgee's or it may, simply, be an outcome whose merits are debatable and contested. Nudging, the objection goes, is not a device that is easily confinable to the pursuit of uncontentious benefits. In response, Thaler and Sunstein suggest that nudges are inevitable (all decisions are structured) and so they might as well be made benignly.[63] Nudge-sceptics would, however, say that this response misses a key point. Some manipulations of decisions, and control systems, are more open than others. If a government issues a law that prohibits citizens from smoking in public places, this is a mode of control that is open, discussed, and implemented after representative processes have been followed. If nudging is used, the process used to effect a nudge may be far more hidden from view—the nudge may flow from an administrator's decision on how to design a public building: a decision not subjected to advanced disclosure or debate. The danger of nudging is that, under the banner of neutrality, control regimes become less overt, less accountable, and more paternalistic.

Such concerns about the accountability and openness of nudging are not necessarily assuaged by Thaler and Sunstein's comments about the occasions *when* nudging will have the most potential for good. The authors suggest that nudging will be most useful where the nudgers or 'choice architects' have high levels of expertise and the nudgees face difficult decisions on which they have poor feedback and few opportunities for learning.[64] Sceptics would immediately voice worries that systems in which 'experts'

[62] Thaler and Sunstein, *Nudge*, 249.
[63] Thaler and Sunstein, *Nudge*, 235–7.
[64] Thaler and Sunstein, *Nudge*, 247.

manipulate the choices of less well-informed parties are exactly those scenarios which there are the greatest dangers that regulatees' and citizens' preferences will be overridden in the name of expert judgements of a spuriously neutral nature. There is, however, a Thaler and Sunstein response. They say that nudgers will be best able to make good guesses about what is best for the nudgees: 'when they have much more expertise at their disposal, and when the differences in individuals' tastes and preferences are either not very big (nearly everyone prefers chocolate ice cream to licorice) or when differences in tastes and preferences can be easily detected'.[65]

Whether this response will placate critics is doubtful. In the first place, it puts considerable faith in experts to identify those circumstances where divergencies of preference are small and to withstand the temptation to impose their own vision of the good. Second, it places nudging in the realm of expert judgement rather than that of open discussion. Third, the discussion of preferences fails to come to grips with the challenges posed by both distributions and strengths of preferences. Ice cream preferences are noted by Thaler and Sunstein, but their discussion only highlights the problem of allowing expert nudgers (or advocates of nudging) to judge preferences. If those parties who prefer liquorice ice cream consider that they are a group worthy of consideration, and if their preferences are very strong, they may object vehemently to nudges that favour chocolate ice cream and they may argue that a movement away from un-nudged choice of ice cream is an example of high-handed expertise at its most undemocratic. They might add that nudges that favour some groups within society rather than others are highly political in nature and should not be swept under the nudge carpet. If we are to have controls over choices that matter to us, and which affect social justice, they might say, let us do so after a proper process of open deliberation.

A third concern relates to the applicability of 'nudging' to the behaviour of corporations. Applying a nudge strategy to corporations presupposes much about the rationality and risk management capacity of such enterprises. Where potential harms may emerge from the cumulative actions of numbers of decision-makers, the nudging of particular decision-makers may not suffice to control the harm's emergence.This suggests that nudging has limited potential, especially in those industries where production chains are complex and extended.

A final worry about nudging is that, whether it is applied to individuals or corporations, its effectiveness may depend not only on the organizational capacity and rationality of the regulatees but also their dispositions. If regulatees are ill-disposed to comply with regulations or are committed to either

[65] Ibid.

creative compliance or an errant course of action, they are unlikely to respond well to nudges.[66] Nudging, accordingly, would not be a satisfactory way to regulate the movements of highly ill-disposed persons who present security risks. It might be added that, for similar reasons, nudging will often prove unsuitable as a means of controlling potentially catastrophic risks.

To conclude, in its ideal form and location, nudging offers a useful means of regulating social and corporate behaviour. This is not, however, a simple tool that carries with it no dangers. The most interesting issue with regard to nudging is not whether it can, in some situations, prove useful and uncontentious. Providing information on tobacco dangers is likely to meet general approval. The more acute questions are whether governments can identify and contain the potential for illiberality that this tool carries and whether its use can be targeted adequately at those areas where it will operate effectively and acceptably.

Rights and Liabilities

In the case of the factory that pollutes the river, the state might decide not to tax pollution or impose standards in a C & C regime, but to allocate rights (for example, to the enjoyment of clean water) so as to encourage socially desirable behaviour.[67] Thus, the argument goes, the prospective polluter will be deterred from such activity by his or her potential liability to pay damages when sued by the holder of the right to clean water (say, the angling club or the riparian owner downstream). The deterrent effect will be provided by the expected cost of polluting—which is the quantum of expected damages multiplied by the probability of those damages having to be paid out. In economic terms, the efficient level of deterrence is that which will ensure that the factory owner will spend money on avoiding pollution up to the point where the cost of avoidance exceeds the value of the harm caused by the pollution. (Beyond that point it is efficient to let the pollution occur and compensate the 'victims', rather than spend on abatement.)

If society desires this efficient level of deterrence, difficulties are encountered because the precise deterrent effects of liability rules are difficult to predict. Rights and mirroring liabilities may, moreover, fail to deter efficiently

[66] For an argument that the evidence is weak that nudging can be used effectively to increase population health, see T. Marteau, D. Ogilvie, M. Roland, M. Suhacke, and 'Judging Nudging' (2011) *BMJ* 342. Available at: http://www .bmj.com/content/342/bmj.d228.full (last accessed 26 January 2011).

[67] See generally Breyer, *Regulation and Its Reform*, 174–7; G. Calabresi and A. Melamed, 'Property Rules, Liability Rules and Inalienability: One View of the Cathedral' (1972) *Harvard Law Review* 1089.

for a number of reasons. Many undesirable events, for example, are the results of accidents, random events, and irrational behaviour. Deterrence, for this reason, does not operate in a mechanical and frictionless manner.[68] A further difficulty is that sub-optimal deterrence may occur where the wealth of the potential harm-causer is insufficient to allow them to fear a level of loss that correlates to the efficient level of deterrence. Thus a small oil tanker operator whose firm is worth $10 million cannot be adequately deterred, and induced to take appropriate precautions, by a potential liability of $60 million (which sum reflects the harm caused by potential spillage). That operator can only fear a potential loss of up to $10 million. This 'shallow pockets' issue would require a response beyond bare liability rules—and compulsory insurance to cover possible losses of $60 million or more might be appropriate.

Under-deterrence may also occur in liability regimes because enforcement costs for individuals may prove discouraging and lead many parties not to proceed to enforce their rights. Coordinating between victims may not always prove feasible, or it may involve high transaction costs. Evidential difficulties may reduce to a low level the probability of proving that the harm involved was caused by the actions of the defendant polluter. (If there is only a 50 per cent chance of proving causation, this halves deterrence. Uncertainties in the legal rules creating rights and liabilities will have a similar effect.) Many victims in the pool of victims may lack the resolve to proceed against the harm-causer and, to the extent that claims are not pursued, deterrent effects will be sub-optimal.

In reflection of such factors, the harm-causer will be likely to be able to settle out of court for negotiated sums that are lower than those that would create efficient levels of deterrence. Courts, of course, might attempt to correct for levels of deterrence that are too low—for example by granting damages that do not merely compensate for harms done but also include a punitive element that makes up for the under-deterrence liable to arise for the reasons cited. The courts will, however, face considerable informational hurdles if taking this course. The judiciary would find it extremely difficult to amass all relevant information about the array of potential actions for damages likely to follow, say, a pollution incident. If such actions are brought separately and serially, the court will not know at a given time in the process how many claims are to be aggregated in calculating total deterrence, nor will it be able to assess the gravity of claims to be brought at a future date.

One final problem is that insurance may limit the deterrent effect of liability rules and may generally make deterrence very difficult to assess. Under certain conditions, insurance may spread risks very widely and undermine deterrence. On the other hand, very high or even excessive levels of

[68] See D. Harris, M. Maclean, H. Genn, and S. Lloyd-Bostock, *Compensation and Support for Illness and Injury* (Oxford, 1984), 328 and on the deficiencies of liability rules in providing compensation see ch. 12.

deterrence (and for firms financial difficulties) may be caused if insurance is subject to restrictions, withdrawals, and crises, so that effective cover at affordable prices is not available. Thus, in the tort sector, what has been described as a crisis was experienced in the mid-1980s in the United States and Canada[69] and it has been the unpredictability of the liability insurance market that has urged a number of North American commentators to look to regulatory devices as alternatives to the tort system.[70]

Public Compensation/Social Insurance Schemes

Economic incentives to avoid undesirable behaviour can be created not merely by systems of taxation and subsidy but also by schemes of compensation or insurance that link premiums paid to performance records. One field in which a good deal of research into insurance-based incentives has been conducted is that of the working environment.[71] A review conducted in 1994[72] pointed to a number of insurance-based schemes dealing with workplace safety and health around the world. National schemes were encountered in several EU countries, the USA, Canada, Japan, and New Zealand, with strategies under development in Denmark, Poland, and elsewhere. These were all no-fault liability schemes and essentially compensatory, though some also provided means of funding improvements in conditions—as in the French, Swedish, and Albertan systems.

In the typical scheme, workers surrender their rights to sue employers for damages relating to health and safety failings, and, in return, are entitled to statutory compensation, often amounting to full payment of lost earnings plus costs. The employer's premiums depend on their organization's past claims experience.[73]

[69] See V. Finch, 'Personal Accountability and Corporate Control: The Role of Directors and Officers Insurance' (1994) 57 *MLR* 880, 915.

[70] See, e.g., G. Priest, 'The Current Insurance Crisis in Modern Tort Law' (1987) 96 *Yale LJ* 521; R.B.Stewart, 'Crisis in Tort Law? The Institutional Perspective' (1987) 54 *University of Chicago Law Review* 184; M. Trebilcock, 'The Social Insurance–Deterrence Dilemma of Modern North American Tort Law: A Canadian Perspective on the Liability Insurance Crisis?' (1987) 24 *San Diego Law Review* 929.

[71] See the work of the Eurofound: the European Foundation for the Improvement of Living and Working Conditions, a European Community institution, reported in: *Catalogue of Economic Incentive Systems for the Improvement of the Working Environment* (Dublin, 1994) (hereafter 'Eurofound Catalogue') and S. Bailey (ed.), *Economic Incentives to Improve the Working Environment* (Dublin, 1994).

[72] Eurofound Catalogue.

[73] See S. Bailey, 'Economic Incentives for Employers to Improve the Management of Workplace Risk'—paper to W.G. Hart Legal Workshop, 4 July 1995.

A central issue attending such schemes is whether state-administered or private insurance mechanisms should be employed. In relation to private provision, doubts exist concerning the extent to which private insurance companies can be relied upon to provide incentives to improve working conditions. The primary concern of private insurers is not to reduce hazards, but to generate profits for shareholders. Such insurers might not be prepared to spend money to isolate poor-risk, dangerous employers beyond profit-maximizing levels. It is true that competition in the insurance market will to some extent drive insurance companies to spend money on discriminating between risks, but there are limits to competitive pressures and, in any event, there is a tension between the basic function of insurance (to spread risks) and risk discrimination (isolating poor risks). This tension also imposes limits on the willingness of private insurers to identify poor risks and to apply localized economic incentives.

In such conditions, the tendency will be to confine risk discrimination to those sectors in which statistical guidance on the quantum of risks is readily available and affordable. Thus, in motor insurance, with a wealth of accidents, and, as a result, useful data available at reasonable cost, discrimination might be high, whereas in relation to workplace safety—where accidents are infrequent but often serious—weak statistics might be expected to lead to low levels of risk discrimination and the linking of cover and premiums to very broadly defined categories of risk.

For such reasons, Eurofound, the European Foundation for the Improvement of Living and Working Conditions, has proposed a publicly administered scheme linking premiums not to statistics on accident records—which were said 'not to make any sense' for firms with under 100 employees[74]—but to factors that could be measured properly such as the conditions of the working environment, the state of the factory's machinery, and so on. Such schemes, said Eurofound, would encourage the accurate reporting of accidents, whereas reliance on past accident records might be expected to encourage firms to massage their statistical returns—for example, by placing pressure on employees not to report accidents (e.g. by offering bonuses to accident-free teams of workers, and creating peer pressures not to report). Insurance-based schemes might also be combined with the use of incentives to improve conditions by allowing premium reductions to companies taking harm-reducing measures (e.g. moving to the use of low-emission materials or low-noise machines).

The further advantages pointed to by proponents of insurance-based schemes[75] are that they make employers conscious of the costs of their actions. Employers considering increasing pressures on workers to take risks

[74] Eurofound Catalogue, 19. [75] Ibid., 24–5.

so as to escalate production levels will be aware that the potential extra profits derived from improved production will have to be weighed against the potential increases in insurance premiums that will follow an inspection by the insurance fund. Prevention will thus be given a higher priority by firms than would be the case under C & C regulation because harms will impinge more directly on their profits. Insurance-based schemes are said to offer incentives and financial motivations to *all* employers, in contrast with C & C strategies, which are so expensive to enforce that they are patchily and poorly applied on the ground.

A further strength claimed for incentive schemes is that they can achieve incentives to go better than fixed standards—indeed, incentives to zero can be instituted. This contrasts with C & C systems, which offer incentives to comply with designated standards but not to perform to higher standards. Employers, it is also said, will respond to the emergence of new hazards under incentive schemes without the need for new legislation.

To balance such sanguinity, however, some caveats do have to be entered. Compensation for workers may produce some undesirable incentives. Thus, if compensation is seen as generous or an easy option, this may encourage some individuals to accept injuries, dangers, or disabilities in return for cash. To work properly, moreover, such a scheme would have to involve the periodic inspection and rating of all employers and their premises. The resource implications are huge. Thus, inspection as envisaged would not be possible in the UK using the present staffing and resources of the Health and Safety Executive, whose current scheme of inspection involves, in the case of medium-sized firms, several years between visits. It might, indeed, be argued that the important difference between the proposed insurance scheme and the existing C & C system lies in the assumptions that are made concerning resources: that with a commensurate increase in resources, C & C could achieve as much.

The differences between an insurance-based scheme and C & C regulation may, thus, be liable to overstatement. In the former, inspectors would check compliance with rules designed to limit risks and would penalize non-compliance by imposing an adjusted premium. In C & C regimes, fines or administrative orders take the place of premiums as sanctioning devices. The insurance-based scheme, it could be contended, is merely a C & C regime with a variation in the sanction. Fines, after all, might be described as disincentives.

Conclusions: Choosing Regulatory Methods

In deciding whether to regulate or to leave matters to the market it is wise, as noted in the last chapter, to be realistic about the levels of performance that

can be expected of regulatory regimes. To compare a friction-free vision of regulation with the imperfect operation of the market is to bias any analysis in favour of regulation. Similarly, in comparing different regulatory strategies, an effort must be made to take into account all the respective difficulties that will be encountered in their implementation. Thus, to compare C & C, with all its enforcement difficulties, to a series of 'less-restrictive' devices that are assumed to be enforceable in a problem-free manner is not to offer a balanced perspective.[76]

Enforcement, as has been noted, is not a difficulty confined to C & C regimes.[77] Nor, moreover, should the *positive* aspects of enforcement be ignored when reviewing C & C regulation. Enforcement procedures can be seen as the lifeblood of many regulatory systems. In Britain, for instance, enforcement practices tend to be more flexible, more administrative, and less prosecutorial than those encountered in the USA, where the most committed critics of C & C are to be found. C & C operates on the ground in a less restrictive and legalistic fashion on this side of the Atlantic, and it is the enforcement practices adopted that ameliorate many of the difficulties encountered in C & C regimes.[78] The objections to C & C, it could be said, often relate to a style of applying C & C regulation—one that is not the norm, say, in Britain.

The difference between C & C and other regimes may, indeed, be one prone to exaggeration since, as noted, many or most schemes require implementation through rules—be these command- or incentive-based. Proponents of C & C have to cope with difficulties of fixing the appropriate level of precision and inclusiveness in rules, of using rule formulations that cope with potential creative compliers, and of incorporating the right kinds of standards.[79] 'Alternative' regulatory methods often need rules, however, on matters such as: *when* incentives will apply; the *conditions* under which franchises will be held or marketable permits transferred; the *kind of information* to be disclosed; the *use* of publicly provided premises; the *extent and form* of liabilities; or the *nature* of premium variations in a social insurance system. Just as enforcement difficulties cannot be assumed away when moving to alternative or 'less restrictive' regulatory methods, neither, it should be repeated, can those problems that attend rule-making processes.[80]

[76] For an argument viewing C & C as a 'last resort', see Breyer, *Regulation and Its Reform*, ch. 9.

[77] See Ogus, *Regulation*, 250–6; Breyer, *Regulation and Its Reform*, 278–80; R. Smith, 'The Feasibility of an Injury Tax Approach to Occupational Safety' (1974) 38 *Law and Cont. Prob.* 730; P. Burrows, *The Economic Theory of Pollution Control* (Oxford, 1979), 33–5.

[78] See D. Vogel, *National Styles of Regulation: Environmental Policy in Great Britain and the United States* (Ithaca, NY, 1986).

[79] See Chapter 14 below and generally Baldwin, *Rules and Government*.

[80] See Markovits, 'Antitrust'.

It should also be cautioned that an historical association between certain regulatory methods and certain styles of implementation—for example, between C & C and the use of highly restrictive rules—should not be taken as a demonstration of inevitable or exclusive linkage. In North America in the 1980s, an enthusiasm for alternative methods of regulation was to a degree fuelled by concerns that C & C methods had led to a 'crisis of legalisation'.[81] Other possible causes of over-proliferation and complexity in rules can, however, be pointed to. Relevant factors may have been: the particular demands made of regulators by North American judges when seeking to control the rationality, fairness, and accessibility of rules and rule-making processes; the existence of certain conditions leading to litigiousness; the operation of certain statutory rule-making procedures; or the political contexts within which particular regulatory institutions operated.[82] Given the potential relevance of such factors, it is difficult to conclude with confidence that a move from C & C to alternative strategies constitutes even a start in combating excessive legalization. There may be a temptation when considering 'alternative' regulatory methods, to isolate their least attractive features and designate these as C & C intrusions—that, however, is, again, to rig the debate.

It should be remembered, at this point, that in most regulatory contexts combinations of regulatory methods tend to be employed. Thus, potential polluters may face some C & C regulations, but also may be subject to licensing or franchising conditions or sets of incentives operating though taxation and subsidy rules. They may have to supply information of various kinds, they are likely to be enmeshed in a network of liability rules, and may be able to avail themselves of publicly provided assets or services. In relation to a given regulatory issue it is, accordingly, necessary to look for the particular mixture of regulatory strategies that will best meet desired objectives—procedural and substantive.[83] It may, indeed, be necessary to consider mixes of regulatory strategies that may go beyond state-instituted regimes. The next chapter, accordingly, examines the potential of self-regulatory, corporate, and third-party controls.

Finally, it should be stressed that regulatory strategies will often have to change over time, either because they are under-performing or in order to meet the new challenges that are posed by such matters as new risks and risk creators or freshly imposed objectives. Such responsiveness will require that regulators are able to assess their own performance (a matter returned to in Chapter 12; see also Table 7.1) but also that they are able to institute the orders of change

[81] See Stewart, 'Regulation and the Crisis of Legalisation', 108–9.

[82] See, e.g., Bardach and Kagan, *Going by the Book*, and R.A. Kagan, 'Should Europe Worry about Adversarial Legalism?' (1997) 17 *OJLS* 165.

[83] See Gunningham and Grabosky, *Smart Regulation*, 14–19, ch. 6.

that are required for optimal regulation.[84] These may be 'first-order' changes such as adjustments of emissions standards. They may be more dramatic 'second-order' shifts in the types of control instrument used—say, from command and control rules to tax incentives, or they may be transformational 'third-order' changes that involve wholesale revisions of the regulatory landscape. These might involve, for instance, re-nationalizations or radical restructurings of industrial sectors or across-the-board replacements of state-operated regulation with market-driven trading regimes. Choices between these orders of change are at least as important as choices of control instruments and mixes of these. It is essential, accordingly, for regulators to operate systems that allow them to recognize the circumstances in which first- or second-order adjustments are insufficient or even counter-productive and when transformational shifts of strategy are required. What counts most is getting the broadest strategy right. As Russ Ackoff, the management thinker, said: 'The more efficient you are at doing the wrong thing, the wronger you become. It is much better to do the right thing wronger than the wrong thing righter. If you do the right thing wrong and correct it, you get better.'[85]

[84] On different orders of change, see J. Black, 'What is Regulatory Innovation?' in J. Black, M. Lodge, and M. Thatcher, *Regulatory Innovation* (Cheltenham, 2005), 8–11.

[85] S. Stern, 'A Fond Farewell to a Brilliant Thinker' *Financial Times*, 10 November 2009.

Table 7.1. Regulatory strategies: posited strengths and weaknesses

Strategy	Example	Strengths	Weaknesses
1. Command & Control	Health and Safety at Work	Force of law.	Intervenes in management.
		Fixed standards set minimum acceptable levels of behaviour.	Prone to capture.
		Screens entry.	Complex rules tend to multiply.
		Prohibits unacceptable behaviour immediately.	Inflexible.
		Seen as highly protective of public.	Informational requirements severe.
		Use of penalties indicates forceful stance by authorities.	Expensive to administer.
			Setting standards is difficult and costly.
			Anti-competitive effects.
			Incentive is to meet the standard, not go better.
			Enforcement costly.
			Compliance costs high.
			Inhibits desirable behaviour.
2. Incentives	Differential tax on leaded and unleaded petrol	Low regulator discretion.	Rules are required.
		Low-cost application.	Poor response to problems arising from irrational or careless behaviour.
		Low intervention in management.	
		Incentive to reduce harm to zero, not just to standard.	Predicting outcome from given incentive difficult.
		Economic pressure to behave acceptably.	Mechanical, so inflexible.
			Regulatory lag.
			Politically contentious as rewards wrongdoer and fails to prohibit offence.
3. Market-harnessing controls		Responses to market driven by firms, not bureaucrats.	No expert agency to solve technical or commercial problems in the industry.
(*a*) Competition laws	Airline industry	Can be applied across industries.	
		Economies of scale in use of general rules.	Uncertainties and transaction costs.
		Low level of intervention.	Courts slow to generate guidance.
		Flexibility for firms.	Principles develop sporadically.
(*b*) Franchising	Rail, television, radio	Enforcement is low cost to public.	Evidential difficulties.
		Low level of restriction.	Need to specify service.

		Respects managerial freedoms. Allows competition for market as substitute for competition in the market. Managers rather than bureaucrats respond to market preferences.	Tension of specification and responsiveness/innovation. Uncertainties impose costs on consumers. Requires competition for franchise but may be few bidders. Need to enforce terms of franchise.
(*c*) Contracting	Local authority refuse services	Combines control with service provision. Sanctioning by economic incentive or non-renewal. Easier to operate than licensing system.	Potential confusion of regulatory and service roles. Poor transparency and accountability. Judicial control weak.
(*d*) Tradable permits	Sulphur dioxide emissions (USA)	Pollution by greatest wealth producer. Incentive to reduce harm to zero. Managerial freedom considerable. Regulatory discretion low. Regulatory costs low.	Enforcement may require inspectorate. Regulatory lag, lack of rapid response in crisis. No compensation for victims. Requires healthy market for permits. Barriers to entry may be created. Some harms need to be prohibited absolutely.
4. Disclosure	Mandatory disclosure in food/drink sector	Low intervention. Allows consumer to decide issues. Lower danger of capture. Useful in low-risk sectors.	Information users may make mistakes. Economic incentives (e.g. price) may prevail over information (on, e.g., risk). Cost of producing information may be high. Risks may be so severe as to call for prohibition. Policing of information quality and fraud may be required. Information may be in form undermining its utility.
5. Direct action and design solutions			
(*a*) Direct interventions	State-supplied work premises	Can separate infrastructure provision from operation. Assures acceptable level of provision. Useful where small firms in poor position to behave responsibly.	Fairness of subsidies may be contentious. Funding costly. Public sector involvement contentious. Innovations may not be market driven.

(*continued*)

Table 7.1. Continued

Strategy	Example	Strengths	Weaknesses
		Allows state to plan long-term investments.	
(b) 'Nudge' strategies	Consent to organ donation is assumed unless positive opt-out is exercised	Low cost, combines influence with residual freedom of choice.	Freedoms may be undermined if opt-out is less than easy. Transparency and accountability of nudging may be low. May not work well where decision processes are complex. May impact poorly on regulated parties who are committed to errant conduct.
6. Rights and liabilities laws	Rules of tort law; right to, e.g., light or clean water	Self-help. Low intervention. Low cost to state.	May not prevent undesired events that result from accidents and irrational behaviour. Individuals may not enforce due to costs. Evidential difficulties and legal uncertainties reduce enforcement. Victims may lack resolve and information to proceed, so deterrence sub-optimal. Difficult for courts to deter efficiently. Insurance may temper deterrent effects.
7. Public compensation / social insurance	Workplace safety schemes (USA, Canada, Japan, New Zealand)	Insurers provide economic incentives. Low intervention in management. Low danger of capture. Encourages accurate reporting of incidents. Makes employers aware of costs of activities. Good coverage, applied to all employers. No need to legislate for each individual harm.	Incidence levels may be too low to allow risk discrimination. Tension of loss-spreading and incentive to behave responsibly. Inspection and scrutiny of performance expensive. May operate in very similar manner to command and control mechanism.