Chapter 3

Economic Incentives

Economic incentive-based instruments, such as taxes or subsidies, are used by policymakers to create prices for the externalities (i.e., economic damages) that farming produces. These policy instruments effectively alter prices in existing markets or create new markets so that producers have incentives to control pollution at socially desirable levels. In this chapter, we detail a variety of economic incentive-based instruments that may be used for nonpoint pollution control and evaluate these instruments according to several criteria related to instrument design, implementation, and the incentives created.

Introduction and Overview

Agricultural nonpoint-source pollution occurs at greater levels than are socially optimal because markets fail to accurately relay the social costs of pollution to producers. Economic incentive-based instruments, such as taxes or subsidies, are used by policymakers to create prices for the externalities (i.e., economic damages) that are produced. These policy instruments effectively alter prices in existing markets or create new markets so that producers have incentives to control pollution at socially desirable levels.

Economists have suggested a variety of incentive-based instruments to control nonpoint-source pollution. However, no general comparison of instruments exists. In this chapter, we provide a detailed discussion of a variety of economic incentive-based instruments that may be used for nonpoint pollution control. Specifically, we show that:

- Incentives must be designed to transmit the goals of policymakers. Producers respond differently to various incentives, depending on the *base* to which the incentive is applied (e.g., the incentive base of a fertilizer tax is fertilizer) and the complexity of the instrument.
- Design-based incentives are generally superior to performance-based incentives.
- Second-best, input- and technology-based incentives are most conducive to policy.
- Coordination of existing programs and improved targeting of incentives are needed for further improvements to water quality.
- Properly designed market-based systems may be effective alternatives to existing programs to control nonpoint pollution.

This chapter begins with a general overview of incentives. Next, we review the two main classes of incentive bases: (1) performance-based incentives (i.e., incentives based on runoff, measured ambient concentrations, or damages), and (2) design-based incentives (i.e., incentives based on inputs and technology). (Table 3-1 lists the economic incentives that are covered in this chapter and provides examples of actual application of each.) Within each class, we consider a variety of specific incentive bases and how each has been applied at the Federal level, evaluating each instrument according to (1) the incentives it provides, (2) its relative complexity, (3) informational requirements of a resource management agency in designing the instrument and of producers in using the instrument to evaluate their decisions, (4) flexibility of the instrument to changing economic and environmental conditions, and (5) potential administration and enforcement costs. In addition, we discuss how policy design issues relate to policies that have been implemented at the Federal level (noting that major State policies are similar). Finally, we review two alternative types of incentives—compliance mechanisms and market mechanisms—and discuss practical experience with these pollution control methods.[1]

[1] We limit our focus to nonpoint policies. However, point sources of pollution will influence damages as well. Point source and nonpoint-source pollution control policies should therefore be conjunctive (see Shortle and Abler, 1997; Shortle and others, 1998a).

Table 3-1—Types of incentives and examples from Federal programs

Incentives	Federal program applications
Performance-based:	
Runoff	None in existence
Ambient	None in existence
Design-based:	
Expected runoff	None in existence
Variable inputs	None in existence
Technology (i.e., fixed inputs, production techniques, etc.)	USDA Conservation Compliance, Swampbuster, ACP, WQIP, EQIP
Acreage at the extensive margin	Conservation Reserve Program

Characteristics of Economic Incentives

Policymakers can use economic incentives to create prices for nonpoint pollution externalities so that producers will control pollution at more socially desirable levels. Incentives may alter prices in existing markets (e.g., a nitrogen tax increases the price of nitrogen) or they can create new markets that did not previously exist (e.g., a market for expected runoff levels is created by either taxing expected runoff levels and forcing producers to "buy" expected runoff from society, or by issuing permits for expected runoff levels to producers and allowing them to sell permits among themselves). Profit-maximizing producers are then forced to consider the social cost of pollution when making management decisions. Management choices are then more consistent with society's environmental objectives.

Economic incentives are generally classified as either a tax or a subsidy.[2] In the case of nonpoint pollution, taxes make it more expensive for producers to pollute by increasing the cost of pollution-causing activities. Alternatively, subsidies make it less expensive for producers to not pollute by decreasing the cost of pollution-mitigating activities. The effect of each can be the same, depending on how they are applied.

The major benefit of economic incentive-based policies is that producers can choose whatever strategy is most profitable for them. In addition, producers' strategies can change as relative prices for inputs and outputs change, or as new technologies become available. Pollution abatement costs will generally be lower with incentives than with command and control policies because producers may be able to utilize site-specific attributes (which a resource management agency may have limited information about) to their advantage in reducing control costs. In addition, innovators may have an incentive to develop and market new approaches that help producers reduce pollution control costs.

[2] The permit price in a market for pollution permits essentially operates as a tax.

Two Types of Taxes and Subsidies

For simplicity, we focus on constant, per-unit incentives (e.g., a sales tax) and lump sum incentives. For a tax, total payments equal the (constant) *per-unit tax* rate multiplied by the tax base. The relationship between total subsidy receipts and the subsidy base is slightly different. A subsidy can be used to provide the same outcome as a tax with the same per-unit rate. However, subsidy payments are often determined relative to a benchmark level. For example, a subsidy applied to fertilizer use might be based on a reduction in use from a specific level. The greater the reduction in total fertilizer use, the greater the subsidy. No subsidy would be provided if there were no reduction in fertilizer use.

A *lump sum* instrument is a fixed tax or subsidy that can be used to influence discrete choices or to determine the distributional outcomes of policies. With respect to discrete choices, lump sum instruments can be made contingent on particular actions. For example, a producer can be paid a lump sum amount if he/she adopts a particular tillage practice, and paid nothing if adoption does not occur. Alternatively, lump sum instruments that are not contingent on particular actions are not applied to a base and therefore do not influence marginal incentives.

Subsidies Versus Taxes for Pollution Control

Taxes and subsidies can be designed to have the same effect on producers' production and pollution control decisions. However, taxes and subsidies will have different impacts on farm profits and on a resource management agency's budget. Taxes will generally reduce farm profits and increase agency budgets, while subsidies will have the opposite effect. However, it is possible to use taxes without reducing farm profits by pro-

A subsidy implicitly supports the view that polluters are not responsible for pollution. Instead, polluters are given the "right" to pollute and society must pay polluters for cleaner water. An alternative view is that society holds the "rights" to cleaner water and that polluters should pay for pollution control (i.e., the "polluter pays" principle). This alternative view is supported by taxes and regulatory policies, and has shaped many point-source programs. For example, point-source control policies under the Clean Water Act hold polluters responsible for treatment costs.

viding producers with a lump sum refund of expected tax payments.

Subsidies require specification of a benchmark level from which they will be determined (e.g., with a point-source emissions subsidy, firms receive a larger subsidy the further emissions are reduced below the benchmark). The specification of this benchmark may create perverse incentives. For example, suppose abatement of point-source pollution is to be subsidized. Establishing a firm-specific pollution abatement benchmark at current discharge levels would penalize firms that have already undertaken pollution abatement. For example, a firm that has been able to reduce emissions to 4 tons on its own would have a 4-ton benchmark while a firm that has not reduced emissions and produces 8 tons would have an 8-ton benchmark. For a given pollution level, the firm with the 8-ton benchmark will receive a larger subsidy and be rewarded for not attempting to reduce pollution on its own. Therefore, establishing a benchmark at current discharge levels would create an immediate, perverse incentive for a firm to produce as large a discharge as possible in order to elevate its benchmark (Baumol and Oates, 1988). Finally, when subsidies are used, society (as opposed to the polluter) must pay for pollution control.

Performance-Based Incentives

Performance-based incentives are taxes or subsidies pursuant to a firm's production and pollution control decisions. Two outcomes of producers' decisions are the most logical targets of incentives for reducing nonpoint water pollution: runoff from a field and ambient water quality conditions.

In theory, a tax or subsidy can be based on how much runoff leaves a site so that the external cost of pollution is considered by producers when they make their production and pollution control decisions.[3] This is akin to an effluent tax on factory discharge. Unfortunately, runoff cannot be monitored at reasonable cost given current monitoring technologies. Only with advances in monitoring technologies will runoff-based instruments become viable policy tools for controlling nonpoint pollution.

Even if runoff was observable, its suitability as an incentive basis would be limited by the natural variability of runoff and other nonpoint processes. Optimally, incentives provide producers with information about the impacts of their choices on expected damages from pollution, and assign them responsibility accordingly. However, a single runoff-based incentive rate can only provide information about how individual choices are expected to impact runoff and not damages. This is because a runoff-based incentive induces producers to consider the impacts of their choices on mean runoff levels, and choices made to achieve a particular mean runoff level do not correspond to a unique level of expected damages. Instead, these choices could have a variety of unintended impacts to damages, due to random events. Thus, runoff-based incentives will not generally provide producers with enough information to accurately consider the external costs of each of their decisions. Similar results occur when trying to achieve an ambient water quality goal at least cost.

Ambient-based incentives are based on the ambient pollution levels in the water resources affected by farming's activities. These incentives are (seemingly) advantageous for two reasons. First, economic theory suggests instrument bases (ambient pollution levels) should be close to the externality (damages from pollution). Second, ambient pollution can be monitored without the resource management agency having to observe the actions of each producer. However, these advantages quickly disappear when informational requirements and other complexities associated with policy design are taken into account (table 3.2).

[3] Incentives can also be applied to each farm's pollution loading to the stream, which runs off from fields influenced by transport characteristics. The results are similar. The only difference is that a loadings incentive requires the producer to determine some transport impacts, while the runoff incentive places this burden on the regulator.

Table 3-2—Evaluation of performance-based incentives

Criteria	Runoff-based	Ambient-based		
		Efficient, CE(r), CE(x)	Cost-effective: CE(a)	Second-best (uniform, limited information)
Incentives provided	N/A	Instrument exists only under very restrictive assumptions.	Poor Not efficient. Exists only when producers are all risk-neutral and when producers and the resource management agency share identical expectations. Additional instruments required for optimal entry/exit.	Poor Not cost-effective. Additional instruments required for optimal entry/exit.
Overall complexity	N/A	N/A	Medium-High A producer must be able to evaluate how he/she and others influence the incentive base.	Medium-High A producer must be able to evaluate how he/she and others influence the incentive base.
Information required by producers	N/A	N/A	High Each producer needs information about production and runoff characteristics of all producers and pollution transport.	High Each producer needs information about production and runoff characteristics of all producers, and pollution transport.
Flexibility	N/A	N/A	High Producers can respond to changing market conditions. Agency has to set one rate for each producer.	High Producers can respond to changing market conditions. Agency only has to set one uniform rate.
Administration/ enforcement costs	Currently prohibitive	N/A	High High information costs. Potentially high monitoring costs in some cases.	Medium-High Medium to high information costs. Potentially high monitoring costs.

N/A = not applicable because instrument is impractical. These rankings are subjective, based only on theoretical properties as opposed to empirical evidence. A more reliable table would be based on empirical results that compare each type of policy according to a consistent modeling framework that is representative of the nonpoint problem.

Incentives Provided by the Instruments

Ambient-based incentives can be designed to achieve an efficient or cost-effective (CE) outcome only under highly restrictive conditions (Horan and others, 1998a,b). For example, a CE outcome designed to achieve a mean ambient pollution target—a CE(a) outcome—can be achieved only when producers are risk-neutral and producers and the resource management agency have the same expectations about the nonpoint process. The ambient tax/subsidy rate that leads to the CE(a) outcome is uniformly applied across producers and equals the social cost of a marginal increase in mean ambient pollution levels. Such a tax/subsidy rate transmits the policy goal of the policymakers to the producers. A CE(a) outcome is possible in this case because the goals of producers would coincide with those of policymakers (i.e., to control mean ambient pollution levels at least cost). When the expectations of producers and the resource management agency differ, a cost-effective solution cannot be achieved because the goals of the producers will differ among themselves and from the goals of the resource management agency (see appendix 3A).

Risk-averse producers will not like the additional risk (due to the natural, weather-related uncertainty associated with ambient pollution levels) that ambient-based incentives create. Instead, risk-averse producers will prefer design-based instruments that can produce the same social outcome and have the same expected impacts on profitability. Moreover, ambient-based instruments cannot produce the CE(a) outcome when used alone. This is because producers' production and pollution control choices have uncertain impacts on ambient pollution levels, creating risk that cannot be adequately controlled with an ambient-based incentive alone (Horan and others, 1998b).

When some producers are risk-averse and/or when ambient-based incentives cannot be designed to accurately transmit the resource management agency's goals, then ambient-based incentives can only be second-best (i.e., achieve policy goals at least cost given risk aversion and heterogeneous expectations about the nonpoint process). Potentially high transaction costs may necessitate that second-best incentives be applied at uniform rates across producers.

Ambient pollution levels depend on the mix of sites in production in the region. If a suboptimal mix of sites is in production, then each producer will face the wrong incentives for input use and technology choices (since these incentives depend on ambient pollution levels, which depend on the mix of sites in production), and equilibrium ambient pollution levels will be suboptimal relative to CE(a) or second-best levels.

By themselves, ambient-based incentives do not provide incentives for optimal entry and exit.[4] Additional lump sum instruments, however, will induce optimal entry and exit into production in the region (Horan and others, 1998a). The lump sum incentives would take the form of a tax applied to producers who produce on extramarginal sites (if they do not produce on this land, they pay no tax) or a subsidy given producers who voluntarily retire extramarginal acreage.[5] It is not necessary to apply lump sum taxes or subsidies to producers who produce on marginal and inframarginal land unless their decision to produce is influenced by the magnitude of the tax. A lump sum refund of their expected tax would reduce their expected tax burden to zero without compromising cost effectiveness.

Relative Complexity of the Instruments

Ambient-based instruments are complex from a producer's perspective because producers must be able to evaluate how their actions and the actions of others affect the incentive base (since the incentive base depends on group performance). Given the large number of nonpoint polluters that may exist within a region, such instruments are likely to be too complex for producers to make accurate evaluations. In that case, producers will receive incorrect incentives from ambient-based instruments.

Informational Requirements

Ambient-based instruments place a large informational burden on producers. To attain a CE(a) or second-best outcome, each producer would have to have information about the actions of other producers and how these actions affect ambient pollution levels or expected damages. Given the large number of nonpoint polluters that may exist within a region, producers are not likely to acquire such information.

[4] *Entry and exit* refer to the process of production sites being entered into or removed from production. Optimal entry and exit occurs when production occurs at positive levels on the marginal and inframarginal sites, but ceases on extramarginal sites.

[5] For example, the Conservation Reserve Program (CRP) uses subsidies to induce producers to retire environmentally sensitive land from production.

The resource management agency also has significant informational requirements. For any performance-based or design-based instrument, the agency must have information about producers and the pollution process so that it can evaluate the impact of the policy instrument on ambient water quality (more information is better, although policies can be designed with less than perfect information). With ambient-based instruments, the agency has the additional burden of having to understand how producers evaluate the impacts of their decisions on water quality. In other words, the agency must understand each producer's belief structure about the nonpoint process. This added requirement is likely to limit the ability of the resource management agency to construct CE(a) or second-best ambient-based incentives.

Flexibility Provided by the Instrument

Producers have flexibility in their production and pollution control decisions under ambient-based incentives in that they may utilize any private knowledge they may have to further reduce costs, or they may alter their decisions as economic and environmental conditions change. The resource management agency may have more flexibility than with some design-based instruments because there is only a single incentive rate to adjust as underlying economic and environmental relationships change. In contrast, several types of rates must be altered as underlying relationships change when incentives are applied to several inputs.

Administration and Enforcement Costs

Information costs associated with setting ambient-based instruments at appropriate levels may be significant. Monitoring costs depend on how easy it is to monitor ambient pollution levels or damages. Monitoring may be relatively easy in some cases (small reservoir or lake) but relatively difficult in others (ground water or major river with many tributaries).

Application of Performance-Based Incentives

Performance-based incentives have not generally been applied in the United States. One possible exception is a tax being used in Florida to reduce phosphorus discharges to the Everglades. The Everglades Forever Act calls for a uniform, per-acre tax on all cropland in the Everglades Agricultural Area. The tax was implemented in 1994 at a rate of $24.89 per acre per year, and will increase every 4 years to a maximum of $35.00 per acre by 2006 unless phosphorus is reduced 25 percent basinwide (State of Florida, 1995). Reductions in phosphorus are determined through monitoring of runoff water that collects in drainage ditches. This type of tax is based on acres of cropland—a design base; however, its application depends on phosphorus levels—a performance base. The tax creates the incentive to adopt best-management practices, and also for producers to apply pressure on recalcitrant neighbors. The number of producers is not so large that free-riding is much of a problem.

This tool is flexible in that producers are not restricted in how they manage their operations to meet the phosphorus goal. However, the basis upon which the tax is placed—acres of cropland—is not necessarily consistent with the goal of the tax, phosphorus reduction. A more efficient approach (and potentially practical, given the small number of polluters) would be to tax phosphorus loads directly.

Design-Based Incentives

Design-based incentives are based on a producer's variable input use and production technology.[6] Producers have no uncertainty about design-based incentives when making decisions, and each producer's decisions may be observed by a resource management agency (although not always easily). However, input use and technology are further removed from damages than with performance-based instruments. Design-based incentives can be based on expected runoff (which is estimated based on inputs and technology) or on inputs and technology directly. After evaluating each subclass, we discuss practical applications of design-based incentives.

Expected Runoff-Based Incentives

Expected runoff levels from cropland may be estimated (before runoff actually occurs) with a simulation model that incorporates all production and pollution control decisions. The incentive base (expected runoff) is therefore design-based because it depends explicitly on inputs and technology (table 3-3).[7]

[6] The inputs and technology targeted by policy may include aspects of pollution control that are unrelated to production.

[7] There may be legal problems with basing permits on the resource management agency's expectations about runoff instead of actual runoff, especially given the limited ability of modelers to accurately predict runoff from input use.

Table 3-3—Evaluation of expected runoff-based incentives

Criteria	Efficient, CE(a), or CE(x)	Cost-effective: CE(r)	Second-best (imperfect information, uniform)
Incentives provided	Instrument exists only under very restrictive assumptions.	Good Cost-effective but not efficient. Additional instruments required to ensure cost-effective entry/exit.	Fair Cost-effective but not efficient. Additional instruments targeted at entry/exit may increase efficiency.
Overall complexity	N/A	Medium-High Instrument may be site-specific or uniform. Producers must evaluate how their production and pollution control decisions influence the instrument base.	Medium-High Instrument may be site-specific or uniform. Producers must evaluate how their production and pollution control decisions influence the instrument base.
Information required by producers	N/A	Medium Producers need information about their own runoff process. However, this information can be provided by the resource management agency	Medium Producers need information about their own runoff process. However, this information can be provided by the resource management agency.
Flexibility	N/A	High Producers are able to respond to changing market conditions. Agency has to set only one rate for each producer	High Producers are able to respond to changing market conditions. Agency has to set only one rate for each producer.
Administration and enforcement costs	N/A	High A simulation model must be developed to determine expected runoff levels for each acre in production. All input and technology decisions must be monitored.	Medium-High Use of limited information may reduce costs. A simulation model must be developed to determine expected runoff levels for each acre in production. All input and technology decisions must be monitored.

N/A = not applicable because instrument not practical. These rankings are subjective, based only on theoretical properties as opposed to empirical evidence. A more reliable table would be based on empirical results that compare each type of policy according to a consistent modeling framework that is representative of the nonpoint problem.

Incentives Provided by the Instrument

Expected runoff-based tax/subsidy rates can be designed to achieve an efficient outcome, (i.e., to control expected damages), CE(a) outcome (i.e., to control expected ambient pollution levels), or CE(x) outcome (i.e., to control input use and technology) only under highly restrictive conditions (Shortle and others, 1998b).[8] This is because expected runoff-based instruments provide producers with incentives to control mean runoff levels from their field, and these incentives generally differ from the goal of policymakers who wish to achieve an efficient, CE(a), or CE(x) outcome (Shortle, 1990; Horan 1998). Expected runoff-based instruments can be designed to achieve a CE(r) outcome (to control runoff at least cost) because the goals of producers then coincide with those of policymakers. The optimal incentive rate in this case would be site-specific, equal to the social value of a marginal increase in mean runoff from the site.

An expected runoff-based instrument will be effective only if producers understand how their production and pollution control decisions influence expected runoff. This information may be provided to producers by the resource management agency in the form of a tax or subsidy schedule based on input and technology choices, or the agency may provide producers with access to the runoff simulation models. Differing expectations about the runoff process are not important here as they were with ambient-based instruments because the incentive is based on the resource management agency's expectations. There would be no benefit to producers from using their own expectations.

Political or legal reasons or transaction costs may prevent a resource management agency from implementing site-specific incentives. Instead, a single incentive rate may be applied uniformly to each site. No matter what policy goals are chosen, a uniform instrument provides incentives for producers to reduce expected runoff levels at least cost. Therefore, the instrument is a cost-effective method of achieving a set of mean expected runoff levels, even if the mean levels achieved do not correspond to the policy goals (i.e., a uniform incentive always leads to a CE(r) outcome). A cost-effective uniform incentive rate equals the average of the expected marginal social costs created by runoff from each site, plus (in the case that policy goals are not to control expected runoff) an additional term (a risk premium) to account for the risk associated with controlling expected runoff as opposed to the policy goal (Shortle and others, 1998b). A uniform expected runoff incentive is not likely to reduce administration costs significantly because the resource management agency would have to construct a model of each site to determine compliance and all inputs and technologies would have to be monitored for use in the model.

If expected runoff incentive rates are set at levels to attain the CE(r) outcome, then the mix of production sites may not be cost-effective because of suboptimal entry and exit (see appendix 3A). The cost-effective mix of sites may be obtained by providing lump sum incentives to producers who produce on marginal or extramarginal sites. It is not necessary to provide lump sum subsidies to producers on inframarginal sites unless, in the case of expected runoff taxes, their decision to produce is influenced by the magnitude of the tax. However, a lump sum refund of these producers' taxes would reduce their tax burden to zero without compromising efficiency.

Second-best policies may be designed when producers retain private information. The resource management agency may have imperfect information about production practices, land productivity, and other site-specific characteristics that affect runoff or economic returns. Producers may be reluctant to truthfully provide any private information to the resource management agency for fear that this information might be used against them in the design of environmental policy. While it may be possible to develop a cost-effective incentive scheme that induces producers to truthfully report their private information, it is implausible due to large informational requirements and related monitoring and enforcement costs (see Shortle and Abler (1994) for the case of such an input-based incentive scheme).

Alternatively, it is possible to design incentives to attain a second-best benchmark that allows producers to retain their private information.[9] In the absence of administration and enforcement costs, policy designed with limit-

[8] Specifically, an efficient rate exists when either (1) the producer makes only a single decision that influences runoff or (2) the covariance between marginal damages and marginal runoff levels is zero for each input (Shortle and others, 1998b).

[9] Policies designed under imperfect information cannot be designed to attain a specific outcome. With limited information, the resource management agency can design policy based only on how it expects producers to react. Therefore, policy would have to be designed to attain an expected outcome.

ed site-specific information will generally be less efficient than policy designed under perfect information. However, given the large costs of obtaining site-specific information, policy designed where producers retain their private information may actually be optimal.

The efficiency of a second-best incentive can be increased if additional instruments are used for entry and exit. Lump sum incentives for achieving optimal entry and exit would take the form of a tax applied to producers producing on extramarginal sites (if they do not produce on this land, they pay no tax) or a subsidy applied to producers who voluntarily retire extramarginal sites (e.g., USDA's Conservation Reserve Program). It is not necessary to provide lump sum taxes or subsidies to producers who produce on marginal and inframarginal sites unless, in the case of an expected runoff tax, their decision to produce is influenced by the magnitude of the tax. However, a lump sum refund of their expected tax would reduce their expected tax burden to zero without compromising optimality.

Overall Complexity of the Instrument

An expected runoff incentive is administratively complex because input use and technology must be monitored for each site in order to determine expected runoff levels (using a simulation model). In addition, the resource management agency would have to develop a model to simulate runoff from each agricultural production site.

Informational Requirements

Each producer must understand how production and pollution control decisions affect runoff if the instrument is to be effective. Information on the relationship between runoff and production and pollution control decisions may be provided to each producer by the resource management agency. To attain a cost-effective outcome, the resource management agency requires perfect information about production and runoff characteristics. Less information is required in designing policies to achieve second-best outcomes. However, efficiency is increased as more information is used to design policy.

Flexibility Provided by Instrument

An expected runoff-based incentive is fairly flexible. Producers have flexibility in that they may utilize any private knowledge they may have to further reduce costs, or they may alter production decisions as economic and environmental conditions change. The resource management agency may have more flexibility than with some design-based instruments because there is only a single instrument base (expected runoff levels) for which incentive rates must be altered as underlying economic and environmental relationships change. When incentives are applied to several inputs, several types of rates must be altered as underlying relationships change.

Administration and Enforcement Costs

Monitoring costs are high for expected runoff-based instruments because the use of each input and technology must be monitored to determine (through the use of a simulation model) expected runoff. Also, providing producers with information about runoff relationships for each production site (by providing access to simulation models) would likely be expensive. Information and administration costs would be higher with site-specific instruments than with uniform or second-best instruments designed using less-than-perfect information.

Input- and Technology-Based Incentives

The second subclass of design-based incentives is based more directly on inputs and technology (Shortle and Abler, 1994). A summary of input- and technology-based instruments (not including expected runoff-based instruments), according to evaluative criteria, is presented in table 3-4.

Incentives Provided by the Instruments

Input- and technology-based incentives can be designed to achieve an efficient or any type of cost-effective outcome (i.e., a CE(a), CE(r), or CE(x) outcome; see table 2-1). The reason is that input and technology choices, while not always equivalent to specific policy goals, are the means by which a resource management agency can achieve its goals. For example, if a resource management agency had absolute control over agricultural production in a region and wanted to achieve an efficient outcome, it would do so by specifying input use and technologies for the region.

Instruments must target all inputs and technology choices to attain an efficient, CE(a), or CE(r) outcome. The cost-effective incentive rate would be site-specific,

Table 3-4—An evaluation of input- and technology-based incentives

Evaluative criteria	Efficient or cost-effective: CE(a), CE(r), or CE(x)	Second-best (i.e., uniform, limited set of inputs, imperfect information)
Incentives provided	Good	Fair
	Additional instruments are needed to ensure optimal entry/exit.	Not efficient. Additional instruments required for optimal entry/exit.
Relative complexity	Medium	Low
	Efficiently or cost-effectively designed instrument is site-specific and applied to each input and technology choice. Producers can easily evaluate instruments.	Incentives applied only to a few input and technology choices, and may be uniformly applied to all producers. Producers can easily evaluate instruments.
Information required by producers	Low	Low
	Producers need information about only their own production processes.	Producers need information about only their own production processes.
Flexibility	Medium	Medium-High
	Producers are able to respond to changing market conditions. Incentives for each production and pollution control decision. Resource management agency must set multiple rates for each producer.	Producers are able to respond to changing market conditions. Incentives for only some production and pollution control decisions. Resource management agency must set multiple rates for each producer.
Administration and enforcement costs	Medium-High	Low-Medium
	Site-specific incentive applied to each production and pollution control choice requires an extensive amount of monitoring.	Costs are reduced the more uniformly the incentives are administered, the fewer inputs are targeted, and the less site-specific information the resource management agency pursues.

Note: These rankings are subjective, based only on theoretical properties as opposed to empirical evidence. A more reliable table would be based on empirical results that compare each type of policy according to a consistent modeling framework that is representative of the nonpoint problem.

equaling the expected social cost of a marginal increase in the use of the input (Shortle and others, 1998a; Shortle and Abler, 1997). Note that the social cost of a marginal increase in the use of an input is negative for those inputs that decrease pollution (e.g., a nitrogen inhibitor). The use of such inputs should be subsidized.

The use of per-unit, input-based incentives alone will not create the incentives necessary to induce producers to adopt the efficient technology (e.g., placing the appropriate taxes on variable inputs may not induce a switch from conventional tillage to conservation tillage).[10] If a suboptimal technology is used, then input use may also be suboptimal since all production decisions are interdependent. Therefore, the optimality of input taxes/subsidies is conditional on the technology chosen. Additional instruments, targeted at technology, are required to attain the efficient, CE(a), or CE(r) outcome.

Lump sum incentives that are contingent on technology choices can produce optimal adoption. For example, a lump sum tax can be applied to producers who adopt a suboptimal technology, or a lump sum subsidy can be applied to producers who adopt the optimal technology. If there are adjustment costs to technology adoption, a cost-sharing approach can also be used to induce adoption.

[10] This is because the choice of production technology has a nonmarginal impact on damages, but the linear instruments only account for the marginal impacts of each producer's choices.

Producers may have available to them a variety of crop production and pollution control technologies and will likely be operating with a suboptimal technology prior to the implementation of nonpoint pollution control policies. The cost of switching to an alternative technology may be significant. Nowak (1987) identifies 15 constraints to adoption (see box, p. 50), most having to do with the costs of obtaining information, management and capital constraints, and perceptions about risk. These constraints explain the frequent use of suboptimal crop management strategies.

Additional instruments may be necessary to ensure optimal entry and exit. The use of input and lump sum technology taxes/subsidies may not result in efficient or cost-effective entry and exit into the region. It may therefore be necessary to apply a lump sum tax/subsidy to producers producing on extramarginal sites to ensure optimal entry and exit. Otherwise, there will be an inefficient mix of sites in production, resulting in too much pollution for the region. The optimal lump sum tax would be applied to producers who produce on extramarginal sites and would ensure that they do not earn after-tax profits on these sites. Alternatively, a lump sum subsidy could be given to producers who retire extramarginal sites. The optimal value would ensure that these producers are better off when they do not produce on extramarginal sites. Lump sum subsidies to producers on marginal and inframarginal sites are unnecessary unless, in the case of input and technology taxes, their decision to produce is influenced by the magnitude of the other taxes. However, a lump sum refund of these producers' taxes would reduce their tax burden to zero without further compromising efficiency.

The resource management agency may have imperfect information about production practices, land productivity, and other site-specific characteristics that affect runoff or economic returns, and producers may be reluctant to truthfully reveal any private information. The agency may therefore have to design a second-best benchmark that allows producers to retain their private information.[11] In the absence of administration and enforcement costs, policy designed with limited site-specific information will generally be less efficient than policy designed under perfect information. However, given the large costs of obtaining site-specific information, policy designed when producers retain their private information may actually be optimal.

[11] See footnote 10.

Political or legal reasons or costs may limit the ability of a resource management agency to implement site-specific incentives for each input that contributes to pollution. Instead, incentives may be applied uniformly across sites and applied to only a few inputs, reducing administration costs. The choice of inputs to target could be based on ease of observation or measurement. Some management practices, such as the rate at which chemicals are applied, are very difficult to observe without intensive and obtrusive monitoring.

An optimal uniform incentive rate equals the average of the expected marginal social costs created by the input use at each site, plus adjustments to account for the average marginal impacts of input substitution on expected social costs and profit levels (Shortle and others, 1998a). The adjustments are needed because placing incentives on the most easily observed inputs can lead to substitution distortions and undesirable changes in the input mix (Eiswerth, 1993; Stephenson, Kerns, and Shabman, 1996). For example, a tax on herbicides would reduce herbicide use, but may increase mechanical cultivation and soil erosion, which in turn has undesirable impacts on water quality. The resource management agency would have to carefully consider the management alternatives to the undesirable practices, and have in place economic incentives or other measures to counter any undesirable characteristics of the alternatives.

The efficiency of second-best, input-based incentives can be increased if additional instruments are used for technology adoption and entry/exit. Specifically, lump sum technology taxes/subsidies could be administered to all producers to ensure optimal technology adoption, and lump sum taxes/subsidies could be administered to producers on extramarginal sites to ensure proper entry and exit (e.g., the CRP). The efficiency gain from using these lump sum instruments diminishes as the uniformity of the lump sum taxes/subsidies grows. Lump sum tax refunds could be provided to producers on marginal and inframarginal sites, reducing their tax burden to zero without further compromising efficiency.

Relative Complexity of the Instrument

Input-based instruments are relatively simple because they are applied as an excise tax/subsidy on variable inputs. Technology-based instruments, since they are lump sum, are also relatively simple. However, the site-specific nature of efficient or cost-effective instruments increases their administrative complexity.

Second-best instruments are designed to be more simple. Other things equal, uniform instruments will be administratively less complex than site-specific instruments, and instruments applied to only a few inputs will be less complex to administer than instruments applied to all inputs. Finally, instruments designed with limited information will be less complex from an administrative perspective.

Informational Requirements

The resource management agency must have perfect information about production and runoff functions for any efficient or cost-effective solution that attempts to control nonpoint pollution. However, second-best policies may be designed with only limited information about site-specific characteristics. Producers have no special informational requirements with input- and technology-based incentives.

Flexibility Provided by Instrument

Producers have flexibility in their production and pollution control decisions under input- and technology-based incentives in that they may utilize any private knowledge they may have to further reduce costs, or they may alter their decisions as economic and environmental conditions change. A resource management agency would have less flexibility with these instruments since a number of incentive rates would have to be adjusted as underlying environmental and economic relationships change.

Administration and Enforcement Costs

Administration, monitoring, and enforcement costs are relatively high for all efficient or cost-effective input- and technology-based instruments due to their site-specific nature and the necessity of monitoring each input and technology used. Second-best instruments are less costly to apply because they do not have to be site-specific, nor does every input and technology choice have to be monitored for each producer. Information costs may also be reduced with second-best policies.

Application of Design-Based Incentives

Studies of actual or proposed economic incentive-based policies for reducing agricultural nonpoint-source pollution are limited. Only a few States have used input-based incentives, and their impact on agricultural nonpoint pollution problems has not been determined. Economists must therefore rely on simulative modeling techniques to gauge how these instruments might perform. Technology subsidies (cost-sharing and incentive payments) and land retirement (extensive margin) subsidies (CRP) are the only tools that have been extensively used for reducing agricultural nonpoint-source pollution.

Input-Based Incentives

The empirical literature on input-based incentives consists primarily of different incentive policy simulations (e.g., Abrahams and Shortle, 1997; Babcock et al., 1997; Helfand and House, 1995; Larson, Helfand, and House, 1996; Tsai and Shortle, 1998; Weinberg and Wilen, 1997). These studies all contend that incentives can be targeted at a limited number of inputs (such as irrigation water or chemical use) and still achieve environmental goals with cost effectiveness. However, the choice of base is important. Cost effectiveness is increased if incentive bases are highly correlated with policy goals (Russell, 1986), and if the incentives encourage producers to reduce sufficiently the use of pollution-causing inputs while not using more of other pollution-causing inputs or less of pollution-mitigating inputs. For example, Helfand and House (1995) and Larson, Helfand, and House (1996) explore alternative tax policies to limit aggregate expected nitrogen runoff levels from lettuce production in the Salinas Valley, California. They find that taxing irrigation water is more cost-effective than taxing nitrogen fertilizer inputs, and almost as cost-effective as regulating both inputs optimally. Water had a higher correlation with runoff, and producers were more likely to use less water than less nitrogen when faced with a given incentive. Peters, McDowell, and House (1997) also found that tax rates on nitrogen fertilizer must be high to reduce expected nitrogen loss due to an inelastic demand for fertilizer.

The uniformity of incentives across sites is also an issue. Helfand and House (1995) determined the use of uniform input taxes within a region to be almost as cost-effective as site-specific taxes. This result is not supported by others, however. Babcock and others (1997), Russell (1986), and Tsai and Shortle (1998) find that targeting incentives to specific sites may significantly outperform uniform approaches due to local geographic and hydrologic conditions. These studies, however, did not consider the additional administrative and information costs associated with improved targeting.

Finally, empirical research suggests that input-based incentives are likely to have only indirect effects on technology choices (or other types of discrete choices such as crop choice and rotation) (Hopkins, Schnitkey, and Tweeten, 1996; Taylor, Adams, and Miller, 1992). If a set of input taxes induces an inefficient set of discrete choices, then input use is likely to remain inefficient as well. For example, inefficient input use can be expected if an input tax policy induces farmers to adopt an inefficient crop rotation. This is because the (efficient) tax rates will fail to provide farmers with appropriate incentives under the production relationships that correspond to an inefficient rotation. The result may be inefficiently high pollution levels (Hopkins, Schnitkey, and Tweeten, 1996; Taylor, Adams, and Miller, 1992).

Technology Adoption Subsidies

USDA and most States have long offered farmers incentive payments for the adoption of conservation practices. Historically, payments were based on the installation cost of primarily structural practices, such as terraces. More recently, the advent of programs such as the Water Quality Incentive Program (WQIP) and Environmental Quality Incentive Program (EQIP) have made payments available for nonstructural management practices, such as conservation tillage. These payments are designed to offset any private losses a farmer may incur by adopting the practice, any increased risk (in terms of uncertain yields) over the first several years of implementation, and any other short-term adoption constraints (see box, "Constraints to Adoption of Alternative Management Practices").

The incentive payments offered by USDA are technology-intensive in that they focus on management practices. Efficiency will be increased if technology-based incentives are used in conjunction with input-based incentives. In order for the short-term subsidy to elicit a change in technology, it must equal the present value of the stream of expected net losses from adopting the practice, if the practice reduces profits. If the practice increases profits, then the subsidy's value is simply that amount necessary to overcome adoption constraints.

Even though incentive payments have been an important tool for many programs, their effectiveness may be limited. USDA financial assistance programs indicate that practice profitability, rather than short-term subsidies, is the most important factor for long-term adoption. The Rural Clean Water Program of the 1980's demonstrated that cost-shared practices had to be attractive on their own merits (EPA, 1990). In a study of soil conservation decisions in Virginia, Norris and Batie (1987) found that farm financial factors, as opposed to cost-sharing, were the most important influences on the use of conservation practices. This suggests that either subsidy levels were not high enough or that subsidies were not offered long enough to be effective.

WQIP incentives may also have been inadequate for encouraging many farmers to adopt practices less damaging to water quality. A 1994 Sustainable Agriculture Coalition study found that WQIP incentive payments were too low in some regions to secure the adoption of recommended practices, including waste management systems, conservation cover, conservation tillage, critical area planting, filter strips, pasture and hayland management, pasture and hayland planting, planned grazing systems, stripcropping, nutrient management, pest management, and recordkeeping (Higgins, 1995).

An Economic Research Service (ERS) study (Cooper and Keim, 1996) used the results of farmer surveys from the Eastern Iowa-Illinois Basin, Albemarle-Pamlico, Georgia-Florida, and Upper Snake Area Study projects (joint ERS-Natural Resources Conservation Service (NRCS)-U.S. Geological Survey projects to study relationships between production practices and water quality) to model the probability of adopting a preferred farming practice as a function of WQIP incentive payments. The practices studied included split fertilizer applications, integrated pest management, legume crediting, manure crediting, and soil moisture testing. Results suggested that adoption rates of 12 to 20 percent could be achieved with no payment, indicating that some practices were profitable on their own merit in some regions. However, the adoption rate would not increase beyond 30 percent with the actual WQIP payments of $10/acre. A substantial payment increase would be required to encourage 50-percent adoption for any of the prac-

Constraints to Adoption of Alternative Management Practices

Nowak (1991) identified 15 constraints to adoption:

1. Basic information about the practice is lacking. Producers do not have adequate information to assess the economic and agronomic properties of a practice, and how the practice might meet overall goals (e.g., profitability or stewardship). A producer will not blindly adopt a new practice without adequate information.

2. Cost of obtaining information is too high. Information is not costless, and the cost or difficulty of obtaining site-specific information may be prohibitive to the producer.

3. Complexity of the proposed production system is too great. There is an inverse relationship between the complexity of a practice and adoption rate.

4. Practice is too expensive. If adoption costs are high in terms of capital outlays and reduced margins, then producers will not be in an economic position to adopt the practice, even if water quality protection is an important goal.

5. Labor requirements are excessive. If a practice requires more labor than the farm manager feels is available, then the practice cannot be adopted.

6. Planning horizon is too short. Some producers may have a short planning horizon because of planned retirement or other factors. If the time associated with recouping initial investments, learning costs, or depreciation of new equipment is beyond the operator's planning horizon, then the practice will not be adopted.

7. Supporting infrastructure is lacking. Producers rely on a network of providers of support and services, such as chemical dealers, implement dealers, extension agents, and other producers. An innovative practice may not be part of the traditional support network's knowledge base. A producer could not adopt such a practice without an adequate support network in place.

8. Producer lacks adequate managerial skill. Many of the new production systems rely on increased management skills, particularly IPM, nutrient management, and precision farming. Producers who do not have the necessary management skills will not adopt such practices.

9. Producer has little or no control over adoption decision. In some cases, a producer cannot make a decision to adopt an alternative practice or production system without the input and approval of partners, landlord, or lender. If these other parties are not convinced of the merits of a proposed change, then the practice cannot be adopted.

10. Information about the practice is inconsistent and conflicting. A producer may hear different messages about the impact of a practice on farm profitability, input needs, and water quality. A producer will be reluctant to adopt a practice until the information about it becomes more consistent.

11. Available information is irrelevant. The information available about the performance of a practice may be based on performance in another county or even another State. A producer may be unwilling to adopt a new practice until information about the practice under local conditions is developed, especially if the new practice entails some investments or changes that are essentially irreversible.

12. Current production goals and new technology conflict. A new technology may not fit into existing production systems or policy settings. For instance, participating in the commodity programs may restrict the ability of a producer to incorporate rotations into his or her operation. A producer may be unwilling to adapt his current operation to fit a new practice.

13. The practice is inappropriate for the physical setting. A practice that was developed for one particular setting, such as flat fertile fields in the Midwest, may cause yield losses, reductions in net returns, or even environmental damage when applied in another setting. A producer will be unwilling to adopt a practice that is inappropriate for his or her setting.

14. Practice increases risk. A new practice may increase the variability of returns. An increase in the risk of a negative outcome may be unacceptable to producers who are risk-averse.

15. Belief in traditional practices outweighs new technology. Some producers are unwilling to abandon practices that are "tried and true," and are therefore perceived as being less risky.

tices. Thus, WQIP payments may be insufficient for adopting and maintaining practices beyond the 3 years that incentives are provided.

The ERS results are supported by a Cornbelt survey (Kraft, Lant, and Gillman, 1996) in which only 17.5 percent of farmers indicated they would be interested in enrolling in WQIP. An additional 27.8 percent stated they might be interested. The average payment requested by those expressing some interest in the program was almost $76 per acre, much greater than the WQIP maximum of $25. Only 18.8 percent were willing to accept $25 per acre or less.

Practice subsidies have also been found to increase the adoption of alternative management practices. Ervin and Ervin (1982) found that government cost-sharing was a significant variable for explaining soil conservation efforts in one Missouri county. Similarly, Nielsen, Miranowski, and Morehart (1989) studied aggregate soil conservation investments and found that cost-shares were a significant variable when conservation tillage was included as an investment. It is important to note that soil-conserving practices produce water quality benefits only as an indirect effect. These practices are designed primarily to enhance long-term soil productivity, which is of immediate economic concern to farmers.

Entry and Exit Subsidies: Land Retirement

The USDA Conservation Reserve Program (CRP) uses subsidies to retire cropland especially prone to producing environmental problems. In exchange for retiring highly erodible or other environmentally sensitive cropland for 10-15 years, CRP participants are provided with an annual per-acre rent and half the cost of establishing a permanent land cover (usually grass or trees). Payments are provided for as long as the land is kept out of production. These subsidies ensure a degree of extramarginal efficiency (i.e., that entry/exit issues are considered to some degree).

CRP eligibility has been based on soil erosion (first 9 signups) and potential environmental benefits (signups 10 and up). With the 10th signup, the cost effectiveness of CRP outlays was increased by using an environmental benefits index (EBI) to target funds to more environmentally sensitive areas. The EBI measures the potential contribution of enrollment bids to conservation and environmental program goals. The seven coequal EBI components are surface-water quality improvement, groundwater quality improvements, preservation of soil productivity, assistance to farmers most affected by conservation compliance, encouragement of tree planting, enrollment in Hydrologic Unit Area Projects of USDA's Water Quality Program, and enrollment in established conservation priority areas. Enrollment bids with a higher EBI to rental payment ratio were accepted ahead of bids with lower ratios. Thus, to some degree, the EBI ensures that land with characteristics most related to environmental quality is enrolled first.

The CRP has converted a total of 36.4 million acres of cropland to conservation uses since 1985, about 8 percent of U.S. cropland. Net social benefits of the CRP are estimated at $4.2-$9 billion (Hrubovcak, LeBlanc, and Eakin, 1995).

Compliance Mechanisms

Instead of offering farmers a payment to adopt alternative practices, existing program benefits can be withheld unless the change is made. So-called compliance mechanisms tie receipt of benefits from unrelated programs to some level of environmental performance. Examples include USDA's Conservation Compliance program to reduce soil erosion and the Swampbuster program to discourage the drainage of wetlands (USDA, ERS, 1994). As applied to agricultural nonpoint-source pollution, program benefits could be withheld if a conservation or water quality plan containing the appropriate technologies is not developed and implemented. Producers would have an incentive to develop the plan as long as the expected program benefits outweighed the costs of implementing the plan.

The effectiveness of compliance mechanisms for controlling agricultural nonpoint-source pollution is limited by the extent to which those receiving program benefits are contributing to water quality problems. In addition, the effectiveness of a compliance approach varies with economic conditions. Generally, program benefits decrease when crop prices are high. It is precisely during these times that agriculture's pressures on the environment are greatest and the incentive effects of compliance are at their lowest. Budgetary reasons may also force the reduction of program benefits, reducing the incentive effect of compliance mechanisms.[12]

[12] The Federal Agriculture Improvement and Reform (FAIR) Act of 1996 reduces commodity support programs through 2003.

The compliance approach's cost effectiveness depends on how the policy is designed. If the policy requires that particular practices be adopted, then cost effectiveness would be poor if it is not possible to choose the practices optimally. If compliance is based on performance, then producers have an incentive to find the least-cost approach to meeting the performance requirements. However, compliance cannot generally allocate pollution control among farms in a least-cost way because program incentives are unlikely to be distributed in a way that reflects contributions to water quality damages (farms with high damages receiving more program benefits). The administration and enforcement costs for compliance may be high. Individual water quality plans must be developed, and farm-level monitoring and enforcement carried out.

The Food Security Act of 1985 enacted conservation compliance provisions for the purpose of reducing soil erosion. The provisions require producers of program crops who farm highly erodible land (HEL) to implement a soil conservation plan. Reducing soil erosion has implications for water quality. Violation of the plan would result in the loss of price support, loan rate, disaster relief, CRP, and FmHA benefits.

The 1996 NRCS Status Review (USDA, NRCS, 1996) determined that only 3 percent of the nearly 2.7 million fields required to have a conservation compliance plan were not in compliance. USDA estimates that nearly 95 percent have an approved conservation system in place. An additional 3.8 percent are following an approved conservation plan with a variance granted on the basis of hardship, climate, or determination of minimal effect. These results indicate that farmers had sufficient incentives to develop and adopt alternative conservation practices.

Evaluations of conservation compliance report minimal or moderate increases in crop production costs and significant reductions in soil erosion (Thompson and others, 1989; Dicks, 1986), although regional assessments show significant variation in costs and benefits. Two studies conclude that conservation compliance is a win-win situation with increased farm income and reduced soil loss (Osborn and Setia, 1988; Prato and Wu, 1991). However, others show reductions in soil loss are achieved only with decreases in net farm income (Hickman, Rowell, and Williams, 1989; Nelson and Seitz, 1979; Lee, Lacewell, and Richardson, 1991; Richardson et al., 1989; Hoag and Holloway, 1991; Young, Walker, and Kanjo, 1991). The majority of HEL can apparently be brought into compliance without a significant economic burden. A national survey of producers subject to compliance found that 73 percent expected compliance would not decrease their earnings (Esseks and Kraft, 1993).

Conservation compliance has resulted in significant reductions in soil erosion. Annual soil losses on HEL cropland have been reduced by nearly 900 million tons (USDA, NRCS, 1996). Average soil erosion rates on over 50 million HEL acres have been reduced to "T," or the rate at which soil can erode without harming the long-term productivity of the soil. If conservation plans were fully applied on all HEL acreage, the average soil erosion rate would drop from 16.8 tons per acre per year to 5.8 tons (USDA, NRCS, 1996).

Finally, conservation compliance has been calculated to result in a large social dividend, primarily due to offsite benefits. An evaluation using 1994 HEL data indicates the national benefit/cost ratio for compliance is greater than 2 to 1 (although the ratios vary widely across regions) (USDA, ERS, 1994). In other words, the monetary benefits associated with air/water quality and productivity outweigh the costs to government and producers by at least 2 to 1. Average annual water quality benefits from conservation compliance were estimated to be about $13.80 per acre (USDA, ERS, 1994). However, these findings do not necessarily indicate that existing compliance programs are cost-effective nonpoint pollution-control mechanisms.

Market Mechanisms

The creation of markets for pollution allowances is an innovative approach to reducing pollution from sources with different marginal costs of control. For point sources of pollution, a simple market works as follows. Each source is provided with a permit defining the level of emissions it may discharge, where aggregate allowable emissions for the watershed are determined based on some policy goal.[13] A market is then created

[13] Permits may be allocated to polluters in a number of ways. They may be auctioned or sold to polluting firms by the government, or distributed free of charge on any basis that is deemed fair. The implicit assumption when firms must pay for permits is that they do not hold the right to pollute. When permits are provided free of charge, initial property rights reside with polluters. The initial allocation does not affect the final outcome, only the distribution of wealth.

by letting firms redistribute emissions levels among themselves by buying or selling "allowances," which are essentially authorizations to increase emissions. For example, if firm A purchases an allowance from firm B, then firm A can increase its emissions by the amount specified by the allowance and firm B must decrease its emissions by the same level.

Firms with initial emission levels greater than their initial permit holdings will have to either purchase more allowances or reduce emissions, depending on the relative cost of each method. Firms with higher marginal costs of emissions reduction will purchase allowances from firms with a lower marginal cost of emissions reduction. This sort of trading scheme makes it beneficial for firms with lower pollution control costs to reduce emissions by more than firms with higher control costs, reducing pollution control costs for the watershed as a whole. Point-source allowance markets have been used for a number of years with varying degrees of success. Most successful has been the market for SO_2 emissions allowances, which has significantly reduced firms' compliance costs for meeting air quality regulations (USGAO, 1997).

Permit Markets Involving Nonpoint Sources

A market could be designed to include nonpoint sources. In such a program, point sources would have the option of purchasing allowances from nonpoint sources to meet their emissions reductions requirements. Trading between point sources and nonpoint sources is possible when the pollutants are common to both point and nonpoint sources (e.g., nitrogen and phosphorus), or when the effects of pollutants on expected damages can be used to determine appropriate trading ratios between different types of pollutants. Costs of reducing agricultural nonpoint-source loads in a watershed may be less than reducing point-source loads, especially where point-source discharges are already being constrained by the National Pollution Discharge Elimination System (NPDES) permits of the Clean Water Act.

Point/nonpoint trading is most feasible when both point and nonpoint sources contribute significantly to total pollutant loads (Bartfeld, 1993). If the nonpoint source contributions are very large in relation to the point-source contributions, then the point sources will be unable to purchase enough nonpoint-source allowances to make much difference in water quality. On the other hand, if point sources are very large in relation to the nonpoint sources, savings from trading may not justify the administrative expense of a trading program.

However, point/nonpoint trading is not suitable for all types of water bodies (Bartfeld, 1993). Trading is most suitable for water bodies with long pollutant residence times, such as lakes and estuaries. In water bodies with short pollutant residence times, water quality impacts of nonpoint-source pollution vary with flow levels. During wet periods when nonpoint-source discharges are greatest, stream flow is also higher, and the impacts of nonpoint-source pollutants on stream water quality are lessened through dilution. On the other hand, streams will experience little nonpoint-source discharge during dry periods when flow is low. It is during these periods that point-source discharge impacts on water quality are most severe. Trading will do little to protect water quality during these low-flow conditions.

Efficiency of a trading program is increased if nonpoint sources can trade with other nonpoint sources. Trading between nonpoint sources will occur, however, only if there is an enforceable cap on runoff (or expected runoff). Otherwise, producers would have no incentive to purchase pollution allowances. As with all pollution control policies, trading will be effective only if policy goals represent an improvement over current situations.

Choice of Permit Base for Nonpoint Sources

As with other incentives, the characteristics of nonpoint pollution make it difficult to establish effective markets for nonpoint pollution allowances. Allowances for nonpoint emissions cannot be directly traded because these emissions cannot be measured (Letson, Crutchfield, and Malik, 1993). Even if emissions permits were allocated to nonpoint sources, there would be no way of knowing whether a source was in compliance.

Nonpoint permits provide producers with incentives to reduce pollution. Therefore, as we have shown throughout this chapter, permits can be applied to a number of bases. In this section, we consider two types of permit markets. The first market is defined by point-source polluters trading emissions allowances for allowances based on expected runoff by nonpoint polluters. The second market is defined by point-source polluters trading emissions allowances for allowances based on input use by nonpoint sources. In both cases, allocative efficiency is increased by allowing trades to occur among like sources.

No matter which base is chosen, nonpoint allowances will not generally be traded for point-source allowances one-for-one due to the different allowance bases, the random nature of nonpoint pollution, and the heterogeneous nature of nonpoint-source contributions to pollution. Instead, a trading ratio must be established to define how many nonpoint allowances must be purchased by a point source to equal one unit of emissions allowances, and vice versa.

Permit Market Based on Expected Runoff

A market based on allowances for expected runoff creates the same incentives as taxes/subsidies applied to expected runoff. Under such a system, an efficient, CE(a), or CE(x) outcome will be attainable only under very restrictive conditions (Shortle and others, 1998b). However, a CE(r) outcome is possible in which allowances are traded at a uniform rate. Optimally, agricultural producers would be allowed to trade allowances among themselves, and also with point-source polluters. Expected runoff allowances cannot be traded one-for-one with point-source emissions allowances, however. A uniform trading ratio equal to the price of an emissions allowance relative to the price of an expected runoff allowance defines the number of emissions allowances that must be traded for one unit of expected runoff. As a result of the uniform trading ratio, high social-cost nonpoint polluters will use more inputs than is efficient while low social-cost nonpoint polluters will use fewer inputs than is efficient. Similarly, high social-cost point-source polluters will emit more than is efficient while low social-cost point-source firms will emit less than is efficient.

There are several problems with basing an allowance market on expected runoff. First, monitoring and enforcement costs will be high because the simulation models used to determine compliance require that the technology used and the use of each input be monitored. Second, producers must know how their production decisions affect runoff if the market is to be effective. Government intervention to help ensure that the necessary information is available to producers would likely be expensive. Finally, legal problems may be created if permits are based on the resource management agency's expectations about runoff as opposed to actual runoff, especially given the limited ability of modelers to accurately predict runoff from input use and management practices.

Permit Market Based on Input Use

Shortle and Abler (1997) suggest trading point-source emissions for nonpoint variable production inputs. The efficient trading ratio is defined to be the marginal rate of substitution of emissions for input use such that expected damages and pre-permit profits are held constant (Shortle and Abler, 1997). With *n* production sites and *m* inputs that influence pollution, *n* x *m* markets (trading ratios) are required to achieve efficiency. Obviously, the transaction costs of such a market system would be considerable (Shortle and others, 1998b).

A second-best allocation could be obtained by allowing trades to occur at uniform rates and by limiting the number of inputs to be traded. The resulting outcome is the same as would occur when uniform input taxes are applied to the same limited set of inputs. The second-best input allowance market economizes on transaction costs associated with monitoring and enforcement of permits for the unrestricted inputs and would reduce the incentives for noncompliance by reducing arbitrage opportunities. Little can be said qualitatively about the second-best prices relative to efficient prices derived by Shortle and Abler (1997). Whether the input allowance prices in the restricted set are higher or lower than their efficient counterparts depends not only on the effects of the input on environmental quality, but also on substitution relationships with other restricted and unrestricted factors.

Uniformity of prices across polluters reduces the cost-effectiveness of pollution control because it eliminates potential gains from different treatment of polluters according to their relative impacts on ambient conditions. The inefficiencies that occur from uniform input prices when differential prices are optimal are analogous to the inefficiencies that can occur when uniform emissions charges are used in place of an optimally differentiated structure (Baumol and Oates, 1988). High control-cost or low social-cost polluters will end up devoting too many resources to pollution control while low control-cost or high social-cost polluters will devote too few resources to pollution control. However, if the differences in the economic gains are small before transaction costs are considered, then even small savings in transaction costs may be justified. If the differences in the gains are large, then the transaction cost savings must be comparably large.

The determination of which inputs are likely to be the best prospects for regulation will depend on the nature of any resulting substitution effects, correlation with environmental quality, and enforcement and monitoring costs. Finally, monitoring and enforcement would be easier for a second-best input market than a market based on expected runoff. Consequently, the costs associated with these activities will probably be less under a market for inputs.

Empirical Evidence

Point/nonpoint trading programs have been set up to restore water quality in several U.S. water bodies, notably Dillon and Cherry Creek Reservoirs in Colorado, and Tar-Pamlico Basin in North Carolina (Hoag and Hughes-Popp, 1997). These existing programs are designed such that point-source polluters purchase emissions allowances from nonpoint polluters. The amount of allowances purchased depends on the amount of expected runoff to be reduced by nonpoint polluters, and the trading ratio. Under existing programs, expected runoff reductions from nonpoint sources in the basin occur through installation of best-management practices (BMP's) and the development of nutrient management plans. For example, the ratio at which nonpoint expected runoff allowances can be converted to point-source emissions allowances is 2:1 for the Dillon Reservoir, and 3:1 for cropland and 2:1 for livestock for Tar-Pamlico. However, it should be noted that permits were not issued to nonpoint sources.

In several existing programs, the expected cost of reducing nonpoint-source loadings was estimated to be lower than the cost of (further) reducing point-source loadings (table 3-5), suggesting that trades may be beneficial for both parties. However, no trades have occurred (Hoag and Hughes-Popp, 1997). One significant factor may be program design. Because nonpoint sources are not regulated, any trades are not enforceable. Instead, if nonpoint-source reductions failed to meet water quality goals, then point sources would be held responsible for meeting the goal through increased point-source controls. Also, agricultural producers may not have wished to participate for fear of being labeled as polluters and becoming regulated in the future.

Table 3-5—Estimated marginal phosphorus abatement costs for point and nonpoint sources

	Abatement cost	
Location	Point source	Nonpoint source
	$/pound	
Dillon Reservoir, CO	860-7,861	119
Upper Wicomico River, MD	16-88	0-12
Honey Creek, OH	0-10	0-34
Boone Reservoir, TN	2-84	0-305

The range of estimates in each case reflects varying stringency of controls or differences among sources (for example, agricultural versus urban sources).

Source: Malik, Larson, and Ribaudo, 1992.

The Tar-Pamlico program provides good examples of several other problems facing existing point/nonpoint trading programs. The largest point-source polluters in this area formed an association and traded as a group (to reduce transaction costs) at a pre-determined price. Members of the association could purchase nitrogen reduction allowances by contributing to the North Carolina Agricultural Cost Share Program at a fixed price of $56/kg (this price has recently been reduced to $29/kg). The State would then handle the task of getting agricultural producers to participate in the program and deciding how much reduction alternative farming practices would achieve. However, the fixed price was based on average control costs, thus reducing the potential benefits that would have been obtained through margin pricing (Hoag and Hughes-Popp, 1997). Also, the program's requirement of a 2:1 trading ratio may have increased the cost of a trade to levels that have been unattractive to point sources. Initial loading reduction goals for the program were met by the point sources through changes in the production process at a cost of less than $56/kg. Finally, the program is hampered by a lack of generally applicable models or data linking land use practices to water quality effects (Hoag and Hughes-Popp, 1997).

No markets currently exist for trading allowances based on nonpoint inputs. However, literature on second-best input taxation offers some insights into the efficiency loss resulting from the use of uniform prices and trading ratios applied to only a few inputs (see the discussion of input-based incentives in this chapter under the heading, "Applications of Design-Based Incentives").

Summary

Economic incentives have many desirable characteristics. They rely on market systems to achieve desired outcomes, they allow producers to respond to changes in economic conditions, and (for a given policy objective) they allocate costs of control efficiently among producers by allowing producers to use their own specialized knowledge about their operations.[14] This chapter has focused primarily on the two main classes of incentives: performance-based and design-based. The choice of base is important in determining (1) the types of incentives provided to producers, (2) the degree of flexibility producers retain in their production and pollution control decisions, (3) the complexity of policy design, (4) the informational requirements of both producers and the resource management agency, and (5) the administration and enforcement costs of the policy.

Instruments perform best when the incentives provided by the instrument coincide with the goals of the resource management agency. For example, an ambient-based instrument can be designed to achieve a mean ambient goal at least cost (when producers have appropriate expectations about the nonpoint process). However, an ambient-based instrument cannot be designed to achieve an efficient outcome because the incentives provided by the instrument (i.e., to control expected ambient pollution levels) differ from goals of policymakers (i.e., to control expected damages). Likewise, a cost-effective expected runoff-based instrument exists when the objective of policymakers is to achieve a mean runoff goal. However, an expected runoff-based instrument cannot be used to achieve an efficient outcome or to achieve an ambient water quality goal at least cost due to differences in policy goals and incentives provided by the instrument. As another example, suppose nitrogen runoff is a problem in a particular watershed. In this case, incentives applied to fertilizer use and irrigation are likely to be more effective than incentives applied to technology choices that are less correlated with water quality or incentives designed to retire land from production.

Performance-based instruments can be inferior to design-based instruments on several grounds. First, runoff-based instruments are not presently feasible because runoff cannot currently be monitored at reasonable cost with current monitoring technology. Second, optimal ambient-based instruments exist only when producers and the resource management agency share the same expectations about the nonpoint process.

Third, the informational requirements for both the resource management agency and producers are increased with ambient-based instruments relative to design-based instruments. For example, producers must be able to evaluate how their actions and the actions of others influence the incentive base for ambient-based instruments to be effective. Moreover, producers have to make predictions about the actions of other polluters before they can predict how their own actions will influence the incentive base. Similarly, the resource management agency must understand how producers will evaluate the incentives. Thus, the agency is required to know what information is available to each producer and how each producer will evaluate that information. Neither producers nor a resource management agency are likely to be able to obtain and process such large amounts of information, which are not required with design-based instruments.

Finally, ambient-based instruments will be less effective if producers are risk averse. In this case, efficiency can be increased if these performance-based instruments are combined with design-based instruments.

Of the two types of design-based instruments described (i.e., instruments based on expected runoff and instruments based directly on input use and technology adoption), second-best input- and technology-based incentives are most conducive to meeting specified policy goals. Ideally, instruments should be applied to all inputs and technologies used and be site-specific. However, empirical evidence suggests only a slight welfare loss from using uniform policies applied to only a few key inputs and technologies. The degree of uniformity, the inputs and technologies targeted, and the amount of site-specific information utilized in policy design that provides the best level of control at lowest welfare and administration cost is conditional on local setting, availability of information, and the skill of the resource management agency. Input and technology incentives may be constructed to perform relatively well in promoting least-cost control when the tax or subsidy is closely correlated to pollution control performance (Russell, 1986). For example, if fertilizer application rates are closely correlated with nutrient loadings to a stream because

[14] Economic incentive policies also create incentives for research into more efficient technologies. This is discussed in chapter 7.

of local geographic and hydrologic conditions, then a tax on fertilizer application will achieve a level of control almost as efficiently as a tax on nutrient loadings (Russell, 1986).

In contrast, expected runoff-based instruments are likely to be more costly to administer than other design-based instruments because the resource management agency has to monitor input use and technology choices for each production site and develop a model to predict runoff from all sites.

Regardless of the choice of instrument base, economic efficiency is increased when additional instruments are used to limit the scale of production in the region. Otherwise, the mix of production sites will be suboptimal, resulting in too much pollution. Optimal policies would ensure that an optimal mix of land remains in production. However, determining the optimal mix involves a comparison of each site's private net returns to the site's contribution to external social costs, an impractical process when there are a large number of agricultural production sites. Instead, second-best principles can be used to limit the costs of such policies. As with the CRP, the resource management agency may develop alternative criteria on which to limit production, such as identifying extramarginal land on the basis of resource characteristics. For example, land consisting of poor soils, steep slopes, or sandy soils overlying ground water used for drinking water, or land that is close to reservoirs might be identified as extramarginal in the sense that the management practices necessary to reduce the risk of water quality damages to acceptable levels would be prohibitive. Such cropland could be retired through a number of mechanisms, including lump sum taxes, subsidies, regulation, or long-term easements.

Coordination of existing programs and improved targeting of incentives will lead to further water quality improvements. Design-based subsidies are being used by USDA and States to promote the adoption of management practices believed to protect water quality. One drawback is that these subsidies are not designed to affect the long-term profitability of a practice. As a result, evidence suggests that they have not successfully promoted the long-term adoption of practices believed necessary to meet water quality goals. A subsidy-based policy could be strengthened by offering long-term subsidies that increase net returns. Another drawback is the technology-based focus of these incentives. While input use may be altered as an indirect effect of adopting alternative practices or technologies, programs will be more successful if incentives are applied directly to input use when this use is highly correlated to water quality impairment.

A final drawback of a subsidy-based policy is that it encourages increases in the scale of production (i.e., production on extramarginal acreage), resulting in more pollution. A separate policy instrument may be required to decrease the scale of production and increase relative efficiency. A lump sum payment or subsidy to retire marginal cropland could achieve this control. (A lump sum tax could also achieve this goal, but such a tax carries the same political baggage as a design tax.) Such a payment is similar to the current CRP, which retires marginal cropland in order to achieve environmental benefits. Coordinating a CRP-like program with long-term incentive programs targeted at both technologies and input use could provide more cost-effective control of nonpoint-source pollution in sensitive watersheds than current programs.

Properly designed market-based systems may be effective alternatives to existing incentive programs. Market-based systems would reduce overall pollution control costs by combining point-source and nonpoint-source policies and allowing markets to allocate pollution control costs more efficiently. The two types of market-based systems that seem to offer the greatest potential are those based on expected runoff and those based on input use. Which type of system performs better is an empirical issue. However, the principles from second-best design incentives may be used in the construction of markets for polluting inputs. A market based on a limited number of inputs may minimize administration costs and still achieve significant pollution control if the inputs are highly correlated with water quality impairments.

The current institutional setting makes point/nonpoint trading difficult and does not favor the establishment of nonpoint/nonpoint trading. A necessary component of a trading program is that the activity the permits are based on (emissions or inputs) can be regulated. Regulations, in the form of emissions permits authorized under the Clean Water Act, exist for point sources. However, nonpoint sources are currently exempt from any regulations. Binding constraints must be imposed on the permitted activities through an enforceable permit system if the market is to operate effectively.

Appendix 3A—
Illustration of Some Results

Proposition 1. An ambient-based incentive can be designed to achieve the cost-effective solution based on a mean ambient target only when producers and the resource management agency share the same expectations about the nonpoint process.

Proof. Denote a producer's site-specific joint distribution function defined over all random variables as h_i (v, W) where v is an (nx1) vector with ith element v_i. In general, a producer's site-specific joint distribution, h_i (v, W), differs from the resource management agency's, denoted by g (v, W).

Denote the site-specific ambient tax rate by t_i. Assuming producers to be risk-neutral, each producer will choose input use to maximize expected after-tax profit, restricted on the choice of technology:

$$V_i(A_i) = \underset{x_{ij}}{Max}\{\pi_i(x_i, A_i) - t_i E_i\{a\}\}$$

where E_i is the mean operator corresponding to h_i (v, W). The first-order necessary condition for an interior solution is

$$\frac{\partial \pi_i}{\partial x_{ij}} - t_i E_i\{\frac{\partial a}{\partial r_i}\frac{\partial r_i}{\partial x_{ij}}\} = 0 \quad \forall i, j \qquad (3A\text{-}1)$$

Comparison of (3A-1) with (2B-4) implies the following condition must hold in the optimal solution:

$$t_i E_i\{\frac{\partial a}{\partial r_i}\frac{\partial r_i}{\partial x_{ij}}\} = \lambda * E\{\frac{\partial a*}{\partial r_i}\frac{\partial r_i^*}{\partial x_{ij}}\} \quad \forall i, j \qquad (3A\text{-}2)$$

where the superscript (*) denotes that these variables are set at their optimal levels in the cost-effective solution. Further manipulation of (3A-2) yields the condition

$$t_i = \frac{\lambda * E\{\frac{\partial a*}{\partial r_i}\frac{\partial r_i^*}{\partial x_{ij}}\}}{E_i\{\frac{\partial a}{\partial r_i}\frac{\partial r_i}{\partial x_{ij}}\}} \quad \forall i, j \qquad (3A\text{-}3)$$

In general, equation (3A-3) is overdetermined with m equations and one unknown. An optimal tax rate exists only when either (1) producers have a single production choice that influences runoff, or (2) h_i (ν, W) = g(ν, W) $\forall i,j$.

Proposition 2. A cost-effective expected runoff incentive tax will result in too few sites in production.

Proof. The optimal tax rate is λ_i^*, where λ_i^* is defined as the value of λ_i in the solution to equations (2B-8) and (2B-9). When faced with an optimal expected runoff tax, the after-tax profits associated with production on the ith site are

$$\pi_i(x_i, A_i) - \lambda_i^* E\{r_i\}$$

In a competitive market, production will occur on a site as long as after-tax profits are positive, i.e., as long as

$$\pi_i(x_i, A_i) - \lambda_i^* E\{r_i\} \geq 0$$

The marginal site, n, is the site for which after-tax profits vanish, i.e.,

$$\pi_n(x_n, A_n) - \lambda_n^* E\{r_n\} = 0 \qquad (3A\text{-}4)$$

In general, $n \neq n^*$ where n^* is the solution to (2B-8) and (2B-9) unless (2B-9) is satisfied. Assuming constraint (2B-3) is binding, condition (2B-9) requires that $\pi_{n*} = 0$, which generally differs from (3A-4), which implies π_n (x_n, A_n) = $\lambda_n^* E\{r_n\} > 0$. Therefore, the number of production sites will be too small. An additional instrument is needed to ensure optimal entry and exit. A lump sum refund of the total tax bill would be sufficient to satisfy (2B-9).