UNIVERSITY OF LONDON THESIS

Degree PhD Year 2006 Name of Author SALMONS, R. M.

Trading mechanisms for the implementation of environmental policy

Roger Salmons
University College London

Submitted for PhD in Economics,
September 2005

Abstract

Relative performance standards are used in a wide range of environmental policy areas, and there is increasing interest on the part of practitioners in the potential for using market-based mechanisms to implement these standards. This thesis analyses the properties of a generic mechanism – *performance-based credit trading* – that can be used to implement any regulatory target that can be expressed in terms of a linear aggregate performance rule. The general formulation of the rule is very flexible and it can accommodate a variety of different forms of regulatory target, including relative performance standards for proportions, weighted averages, and rates.

The environmental effectiveness and cost-efficiency of performance-based credit trading are analysed, and it is demonstrated that if all markets are perfectly competitive, then the mechanism will ensure that the regulatory target is met at the lowest possible economic cost. Comparative statics analyses are undertaken to assess the relationship between the performance rule parameters and the price of *performance credits* created under the mechanism, and it is shown that for particular types of performance standard it is possible that the price may fall as the standard becomes more stringent. Two specific policy applications – industrial energy efficiency and packaging recovery – are used to explore impacts of the mechanism on the prices and quantities of the market commodities included in the performance rule; to show how market information can be used to estimate the cost of the regulatory intervention; and to demonstrate how *performance-adjustment factors* can be used to manipulate the distributional impacts of the mechanism. Finally, the implications of relaxing the assumption of perfect competition are considered, and it is demonstrated that both the environmental effectiveness, and the cost-efficiency of the mechanism may be sensitive to the assignment of *obligations* and *initial property rights*.

Contents

List of Tables

List of Figures

To Helen

Acknowledgements

I would like to thank my supervisor, Dr Joe Swierzbinski, for all his advice and guidance during the production of this thesis. I would also like to acknowledge Professor David Pearce, who inspired me during my M.Sc. course, and encouraged me to embark on a career in research.

Above all I would like to thank my wife, Helen, for her patience and encouragement. Without her love and support, this would not have been possible.

Preface

The past twenty-five years has witnessed a significant evolution in the approach to environmental policy design. In particular, there is now a general recognition of the advantages that economic instruments can provide in terms of reducing the costs of maintaining and improving environmental quality. During this time, the academic literature has focused primarily on two instruments – emissions taxes / charges, and tradable emission permit schemes with absolute limits (so called "cap and trade" schemes); analysing their respective properties, comparing their relative merits, and investigating various policy applications. In contrast, little consideration has been given to the use of economic instruments to implement relative performance standards. To the extent that they have been considered, it has been to show that they cannot induce the socially optimal outcome, unlike emission taxes and tradable emission permits. This thesis attempts to redress the balance (at least to some degree), by investigating the properties of a generic market-based mechanism – performance-based credit trading – that can be used to implement a wide variety of relative standards.

The thesis is divided into four parts. *Part One* sets the scene, with Chapter One providing the rationale for considering the topic, together with some background on the use of relative performance standards in environmental policy in Europe and the USA. Chapter Two provides an overview of performance-based credit trading; describing the key features and demonstrating how the generic mechanism can be tailored to implement a number of different types of standard from a range of policy areas. *Part Two* is the heart of the thesis, investigating the properties of performance-based credit trading under the assumption of perfect competition in all markets. The cost efficiency and distributional flexibility of the mechanism are analysed formally in Chapter Three; while the relationship between the price of performance credits and the stringency of the standard is investigated in a series of comparative statics analyses in Chapter Four. The properties of the mechanism are explored further in the next two chapters in the context of two specific policy applications: an industrial energy efficiency standard, and a packaging recovery target under extended producer responsibility. The assumption of perfect competition is relaxed in *Part Three*, in which the implications of strategic interaction between agents are investigated. Following a review of the literature on market-power and trading mechanisms in Chapter Seven, the implications of market power for the design of the mechanism are assessed in Chapter Eight – again in the context of a packaging recovery target. Finally, *Part Four* pulls together the findings of the preceding chapters and draws some general conclusions.

Definitions and notation

The following definitions and notational conventions are used throughout the thesis.

- $\Re$ is the set of real numbers; $\Re_+$ is the set of non-negative real numbers; $\Re_{++}$ is the set of strictly positive real numbers; $\Re_-$ is the set of non-positive real numbers; $\Re_{--}$ is the set of strictly negative real numbers. For any positive integer A, $\prod_{\mathrm{A}} \Re \equiv \Re \times \Re \times \ldots \times \Re$ is the Euclidian space of dimension A.

- $\aleph$ is the set of positive integers (i.e. natural numbers). For any finite set $\mathrm{A} \subset \aleph$, the number of elements in the set is denoted by $|\mathrm{A}|$. Thus, for any two mutually exclusive sets A and B (i.e. $\mathrm{A} \cap \mathrm{B} = \varnothing$), $|\mathrm{A} \cup \mathrm{B}| = |\mathrm{A}| + |\mathrm{B}|$. The union of sets $\mathrm{A}^1 \ldots \mathrm{A}^\mathrm{N}$ is denoted by $\bigcup_{i \in \mathrm{N}} \mathrm{A}^i = \mathrm{A}^1 \cup \mathrm{A}^2 \cup \ldots \cup \mathrm{A}^\mathrm{N}$.

- Variables for individual agents (i.e. consumers, firms) are denoted by lower-case letters (e.g. *x*, *y*, etc.). Aggregate variables are denoted by upper-case letters (e.g. *X*, *Y*, etc). Arbitrary (non-specific) fixed values of individual and aggregate variables, and exogenous parameters, are denoted by letters in normal type-face (e.g. x, Y, etc. and a, b, etc.). Specific values of variables, such as solution values, are identified by a superscript (e.g. x^*, $\mathrm{Y}^\#$, etc).

- Vectors are denoted by lower-case or upper-case letters in bold (e.g. $\boldsymbol{x}$, $\boldsymbol{Y}$, etc.) depending on whether they relate to individual variables, aggregate variables, or parameters. When comparing the values of two vectors:

  $\mathbf{x} \geq \mathbf{y}$ means that all elements of vector $\mathbf{x}$ are greater than or equal to the corresponding elements of vector $\mathbf{y}$ (i.e. includes $\mathbf{x} = \mathbf{y}$);

  $\mathbf{x} > \mathbf{y}$ means that all elements of vector $\mathbf{x}$ are greater than or equal to the corresponding elements of vector $\mathbf{y}$, with at least one element strictly greater (i.e. excludes $\mathbf{x} = \mathbf{y}$);

  $\mathbf{x} >> \mathbf{y}$ means that all elements of vector $\mathbf{x}$ are strictly greater than the corresponding elements of vector $\mathbf{y}$. In particular, if $\mathbf{x} >> \mathbf{0}$ then all of its elements are strictly positive.

  Corresponding interpretations apply for $\mathbf{x} \leq \mathbf{y}$, $\mathbf{x} < \mathbf{y}$, and $\mathbf{x} << \mathbf{y}$.

- Unless stated otherwise, input quantities are represented as negative real numbers, and output quantities as positive real numbers. So, for example, $w_{ki} \in \Re_{-}$ is the quantity of commodity $k \in \mathrm{K} \subset \aleph$ used as an input by firm $i \in \mathrm{I} \subset \aleph$.

Part One

Introduction

Chapter 1 Background and motivation

The primary objective of this opening chapter is to provide the rationale for the topic addressed by the thesis. That is, to explain the reasons why a study of the theoretical properties of a market-based implementation mechanism for relative performance standards is relevant and interesting. In addition to this, the chapter has two secondary objectives. The first is to provide some general background on the use of relative performance standards in environmental policy. The second is to provide a framework for the analysis, and to define some of the terminology that will be used throughout the thesis.

The chapter starts by addressing the last of these objectives. A general framework for analysing and evaluating environmental policy is set out, in which a distinction is made between environmental objectives, regulatory targets and implementation mechanisms. This in turn allows a distinction to be made between alternative definitions of environmental effectiveness and economic efficiency.

The second section of the chapter considers some of the theoretical arguments that can be made to justify the need for a cost-efficient implementation mechanism for relative performance standards. In particular, it is shown that when one moves away from the ideal "first-best" world – in which there is perfect information, markets are all perfectly competitive, and there are no distortionary taxes, the use of an efficiently implemented performance standard may be preferable to other forms of regulatory intervention based (explicitly or implicitly) on an absolute limit.

Even if there were no theoretical arguments for doing so, the widespread use of relative performance standards in practice, provides a strong practical argument for the development of a cost-efficient implementation mechanism; since it is likely to be easier to change the existing implementation mechanism (to make it more efficient) than to move to a completely new form of regulatory intervention based on absolute limits. With this in mind, the third section of the chapter provides a brief review of the broad range of environmental policy areas in which relative performance standards are being (or have been) used.

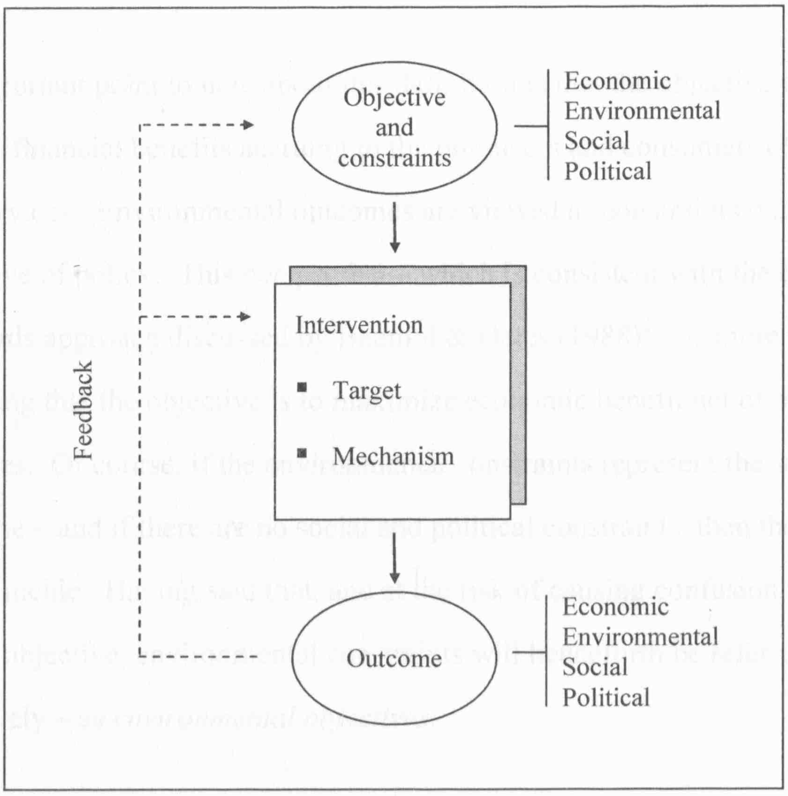
Where relative performance standards have been used, they have generally been implemented either by the imposition of fixed individual standards, or by negotiated agreements with industry associations. However, there are a small number of examples where market-based, trading mechanisms have been used, and three of these examples are discussed in the fourth section of the chapter.

1.1 Framework for analysing and evaluating environmental policy

1.1.1 Analysis framework

Figure 1.1 provides a schematic representation of the environmental policy process. This model is of course a highly simplified and stylised representation, of what is in practice a much more complex and unstructured process. However, it provides a useful framework for the definition, analysis and evaluation of the various components of environmental policy, and for clarifying the meaning of the some of the terminology that is used.

Figure 1.1 Environmental policy process



The model can be applied to environmental policy as a whole, or to individual policy areas (e.g. waste, air quality, water quality, climate change, etc.). In either case it is assumed that the *overall policy objective* is to:

- maximize gross economic benefit / minimize gross economic cost

subject to satisfying: a) identified environmental constraints

b) identified social constraints

c) identified political constraints

For example, one of the environmental constraints might be that emissions of greenhouse gases must not exceed a certain level; one of the social constraints, that energy poverty must not be worsened; and one of the political constraints, that a powerful industry lobby must not be alienated.

An important point to note about this definition is that the objective includes only the private financial benefits accruing to the producers and consumers of marketed goods and services. Environmental outcomes are viewed as constraints on, rather than as an objective of policy. This perspective – which is consistent with the environmental standards approach discussed by Baumol & Oates (1988)[1] – is more realistic than assuming that the objective is to maximize economic benefit net of environmental damages. Of course, if the environmental constraints represent the socially optimal outcome – and if there are no social and political constraints, then the two definitions will coincide. Having said that, and at the risk of causing confusion with the overall policy objective, environmental constraints will henceforth be referred to – more intuitively – as *environmental objectives.*

For each environmental objective, there will usually be a *regulatory intervention.* That is, an action on the part of government aimed at changing the behaviour of firms and / or individuals.[2] In some cases there may be several complementary interventions. Alternatively, a single intervention may support more than one environmental objective – possibly in different policy areas. For example, a target for the amount of biodegradable municipal waste that is sent to landfill would support environmental objectives in two areas – waste management and climate change.

Each intervention has two components – a *target* (or a set of targets) and an *implementation mechanism.* The first defines what is to be achieved, while the second

[1] It actually extends the approach, in that it also explicitly takes into account social and political constraints.

[2] It is possible that for some, non-critical environmental objectives that the government may rely on unilateral voluntary action by firms or individuals, rather than intervening directly.

defines how it is to be achieved. As can be seen in Table 1.1 – which provides some illustrative examples – the target is made up of three parts; a target variable, and target value, and a condition.

Table 1.1 Examples of regulatory targets

Target variable	Condition	Target value
▪ Aggregate quantity of biodegradable waste sent to landfill by all waste disposal authorities in England and Wales	≤	L tonnes
▪ Individual emissions of NO_X by each plant regulated under IPPC	≤	L tonnes
▪ Individual quantity of recycled newsprint used by each newspaper publisher, divided by its total newsprint use.	≥	r %
▪ Aggregate emissions of carbon dioxide by all power generators	≤	Efficient level

The *target variable* defines the object of the intervention – i.e. the physical flow that is to be controlled. It may be defined as an aggregate for a specified group of agents, or it may be defined for an individual agent. It may be defined in absolute terms, or relative to another flow. Throughout this thesis, regulatory targets that are defined in absolute terms are referred to as *limits*, while those defined in relative terms are referred to as *performance standards*.

The *target value* may be common – i.e. the same value for all agents (or groups of agents), or it may be differentiated to reflect individual circumstances (e.g. the costs of abatement, or the environmental sensitivity of the locale). While in most cases it will be defined explicitly, this need not always be so, and the target value may have to be inferred from the underlying environmental objective, and / or the implementation mechanism. For example, if the mechanism is an emissions charge set equal to marginal environmental damages, then the (unspecified) target value is the efficient level of emissions. Finally, the *condition* determines whether the target value represents a maximum that cannot be exceeded, or a minimum that must be achieved.

In many cases, the regulatory target will be the same as the environmental objective. This is the case, for example, when an absolute objective for the aggregate emissions of greenhouse gases is translated into an emissions limit. However, while one would expect the two to be causally related (or linked), they do not necessarily have to be identical. For example, the same environmental objective could be achieved by setting performance standards for the specific energy consumption (i.e. energy consumption per tonne of output) of energy intensive sectors.

There are a wide range of mechanisms that can be used to implement a regulatory intervention. However, essentially these can be classified into the six broad types that are listed in Table 1.2. These represent a spectrum of different levels of prescription – with flexibility generally increasing as one moves down through the different types. Technology-based and consent-based mechanisms are the most prescriptive, and are commonly termed "command and control" mechanisms, in that the regulator specifies a certain technology to be used, or a maximum emissions level, and then monitors the regulated plants to ensure that they comply. Contract-based mechanisms can be viewed as applying command and control to a group of firms, and allowing them to decide how to share the collective burden among themselves. Market-based and price-based mechanisms work by creating financial incentives for agents to change their behaviour. As such they are often referred to as economic mechanisms (or economic instruments). Finally, information-based mechanisms attempt to overcome the information failures that prevent agents from undertaking economically rational abatement activities.

Table 1.2 Classification of implementation mechanisms

Type	Description
▪ Technology-based	The specification of particular production technologies that must (or must not) be used, or of particular technical characteristics for products.
▪ Consent-based	The specification of the maximum quantity of a non-market environmental commodity, or a market commodity, that may be used by an individual plant, either in absolute or relative terms.
▪ Contract-based	The signing of legally-binding contracts with groups of firms (e.g. industry associations), that specify certain collective actions, and / or targets, either in absolute or relative terms.
▪ Market-based	The creation of markets for new commodities that are linked to non-market environmental commodities, or to existing market commodities.
▪ Price-based	The use of taxes and / or subsidies to introduce prices for non-market environmental commodities, or to adjust the prices of market commodities.
▪ Information-based	The provision of information about management best practice, and about the costs and benefits of abatement technologies.

Whatever type of implementation mechanism is used, there is a (regulatory) *control variable*. That is, a variable that is set, and adjusted, by the regulator in order to achieve the regulatory target and / or the environmental objective. For example, this might be the total number of permits issued in a market-based mechanism; or the value of the emissions tax that is set in a price-based mechanism; or the particular technology that is specified in a technology-based mechanism. The control variable should be consistent with the regulatory target, whenever the latter defined explicitly. Thus, if the intervention takes the form of an aggregate emissions limit (L), then the total number of emission permits should be set equal to L.

In practice there is increasing interest in the use of "hybrid" mechanisms, which combine one – or more – of the different "pure" types. For example, under the Climate Change Levy "package" in the United Kingdom, firms in energy intensive sectors have been able to gain a partial exemption from the Levy (a price-based mechanism), by entering into negotiated agreements (a contract-based mechanism) that set targets for specific energy consumption. Furthermore, a firm is allowed to meet its target by purchasing credits from another firm that has beaten its target (a market-based mechanism). Notwithstanding this trend, Table 1.2 provides a useful conceptual taxonomy of the different implementation mechanisms that can be used.

The final element of the policy process is *feedback*. Based on an evaluation of the *outcome* of the regulatory intervention, changes may need to be made to the control variable in order to ensure that the environmental objective is achieved. This is particularly relevant when there is a divergence between the environmental objective and the regulatory target, where achievement of the target may not guarantee that the objective is met. It is also possible, that the environmental objective itself may be revised if the economic cost turns out to be much lower, or much higher than expected.[3]

1.1.2 Evaluation criteria and definitions

The OECD has identified five broad dimensions for evaluating actual, or prospective, implementation mechanisms (OECD, 1991). These are:

[3] This might also lead to a relaxation of the political constraints if it reduces industry's concerns about the potential cost of environmental regulation.

- Economic efficiency
- Environmental effectiveness
- Equity
- Political acceptability
- Administrative feasibility and cost

The first four dimensions relate directly to the policy objective and constraints identified above. The fifth can be viewed as part of the objective function for the policy problem – at least with regard to the cost. It is identified separately here in order to allow any evaluation to distinguish between the deadweight cost of administering the intervention, and the abatement costs borne by producers and consumers. While these evaluation dimensions may seem straightforward, both environmental effectiveness and economic efficiency can be interpreted in different ways, as the OECD report acknowledges (in relation to the latter).

The effectiveness of an implementation mechanism can be assessed either in terms of its performance versus the regulatory target, or in terms of its performance versus the environmental objective. Of course, if the target is the same as the objective, then the distinction is redundant. If the implementation mechanism achieves the target, then the objective is met automatically. However, if the target and the objective are not the same, then this does not necessarily follow. For example, suppose that the target value for specific energy consumption is set so that *ex ante* it achieves an environmental objective for the absolute level of CO_2 emissions. If the actual level of output turns out higher than expected, or if the fuel mix is different, then the objective may not be met, even though the target is achieved.

Economic efficiency can also be interpreted in several different ways. It can relate to the gross economic cost of achieving the regulatory target – i.e. the *cost efficiency of the* *implementation mechanism.* Alternatively, it can relate to the gross economic cost of achieving the environmental objective – i.e. the *cost efficiency of the regulatory intervention.* Both of these interpretations are "conditional" measures of efficiency. An alternative, unconditional measure of efficiency relates to the net economic benefit (or welfare) that is achieved – i.e. the *welfare efficiency of the regulatory intervention.*

Under any of these interpretations, efficiency can be measured in either absolute or relative terms. An implementation mechanism is absolutely cost efficient if it minimizes the gross economic cost of meeting the regulatory target. In this case, the target is said to be efficiently implemented. Similarly, a regulatory intervention is absolutely cost efficient if it minimizes the gross economic cost of achieving the environmental objective, and is absolutely welfare efficient if it maximizes net economic benefits. In many situations however, one is interested in the relative costs of alternative mechanisms or interventions. In this case, one mechanism / intervention is relatively efficient (compared to another) if the resultant gross economic cost is lower.

Of course, if the regulatory target is the same as the environmental objective, then the distinction between the different measures of efficiency is redundant. If a mechanism is cost efficient, then the regulatory intervention is automatically cost efficient (in both absolute and relative terms). Furthermore, if the objective is set at the socially optimal level, then the intervention is also welfare efficient. However, when the target differs from the objective, then the distinction becomes important. In this case, a regulatory intervention may not be cost efficient, even if the target is efficiently implemented.

1.2 Theoretical arguments for using performance standards

It is well established that in an ideal world – where there is perfect information, markets are all perfectly competitive, and there are no distortionary taxes, a regulatory intervention that uses a market-based mechanism to implement an absolute target for aggregate emissions is cost efficient, and – potentially – welfare efficient.

It is thirty years since Montgomery (1972) demonstrated that a system of tradable ambient permits would minimize the gross economic cost of meeting a set of location-specific, absolute deposition targets for a linear air pollutant such as sulphur dioxide.[4] Furthermore, the outcome is independent of the initial allocation of permits between polluters. While the implementation efficiency of the mechanism was proved in the context of a non-mixable pollutant, the result applies equally to the case of a perfectly-mixable pollutant such as carbon dioxide (CO_2).[5] As was noted in the previous section, provided that the targets reflect the underlying environmental objectives, and that these are set optimally, then the result can be extended directly to the cost efficiency, and the welfare efficiency of the regulatory intervention. The analysis was extended by Spulber (1985) for the case of a perfectly-mixable pollutant, to allow for the entry and exit of firms. He demonstrated that, if the total number of permits is set equal to the socially optimal level of emissions, then the mechanism will induce the optimal long-run structure for the industry. That is, it will provide the correct incentives for market exit.

In contrast, to date little consideration has been given to properties of regulatory interventions that use a market-based mechanism to implement a relative performance standard. Dinan (1992) describes how a credit trading mechanism could be used to implement a relative performance standard for the percentage of virgin pulp used in the production of newsprint (relative to total pulp).[6] While he claims that the mechanism

[4] A pollutant is linear if there is a linear relationship between emissions at source locations, and depositions or ambient concentrations at receptor locations (i.e. all source-receptor coefficients are exogenous).

[5] A perfectly-mixable pollutant can be interpreted as a special case, where there is a single receptor point, and the source-receptor coefficients are all equal.

[6] The scheme is described in section 1.4.

would minimize the cost of achieving the target – he does not prove that this is actually the case.

The welfare efficiency of an intervention that uses a system of tradable credits to implement a relative standard for specific emissions (i.e. emissions per unit output) is considered by Fischer (2001), and by Gielen *et al* (2002). The two analyses differ slightly in their representation of the firms' decision variables.[7] However, they come to the same conclusion – that it is not possible for the intervention to be welfare efficient. That is, unlike a permit-trading scheme for an absolute limit, a credit-trading scheme for a relative standard cannot induce the socially optimal outcome. This is because the mechanism has the effect of subsidizing output. Consequently, at the optimal values of specific emissions / abatement, output is higher than the socially optimal level, and hence so too are emissions. The corollary of this is that in order to achieve the socially optimal level of emissions, it is necessary to impose a more stringent performance standard, at a higher gross economic cost.

Given these findings, it may seem strange to pursue the idea of using a market-based mechanism to implement aggregate performance standards. Surely it would always be better to set an absolute target, and use a tradable permit scheme, or an equivalent emissions tax, as the implementation mechanism? However, as was noted above, all of these findings relate to an ideal world – where regulators make decisions under conditions of certainty; where markets are all perfectly competitive; and where there are no distortionary taxes, such as taxes on labour. When any of these assumptions are relaxed, a rather different picture emerges, and it is possible that welfare can be higher under an efficiently implemented performance standard than under an efficiently implemented limit.

---

[7] In Fischer (2001) the decision variables are the firms' specific emission rates (ε_i) and their output levels (y_i), while in Gielen *et al* (2002), the decision variables are the firms abatement expenditures (a_i), and their output levels (y_i).

The implications of relaxing each of these assumptions will be considered in turn. In each case, three alternative forms of regulatory intervention are compared:

- an absolute limit for aggregate emissions, implemented by a cost-efficient, revenue-raising mechanism;
- an absolute limit for aggregate emissions, implemented by a cost-efficient, revenue-neutral mechanism;
- a relative standard for aggregate specific emissions, implemented by a cost-efficient, revenue-neutral mechanism.

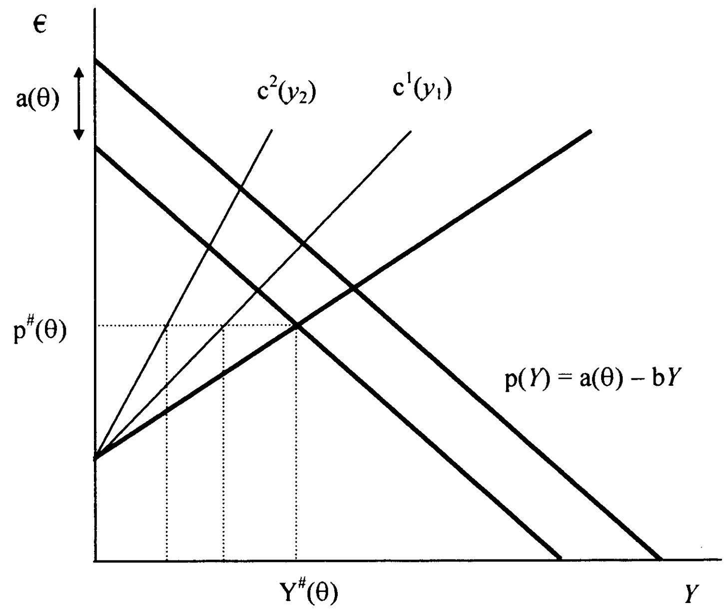
The exact specification of the mechanism in each case is not necessary for any of the following analyses. The only requirement is that it minimizes the gross economic cost of achieving the respective target. However, in order to facilitate the distinction between the first two forms of intervention, it is assumed that the revenue-raising mechanism is an emissions tax. Hence the three forms of intervention are identified as an emissions tax, an absolute limit, and a relative standard.

1.2.1 Uncertainty

The lead times involved in developing new regulatory interventions, or making changes to existing interventions, can often be lengthy. This means that the decision regarding the stringency of the intervention may have to be taken many months, or even years, before it comes into force. While this has the advantage of allowing time for the regulated agents to adapt, it means that the regulator must base the decision on a forecast of the expected economic conditions at the time of implementation, which will be subject to error.

The implications of this uncertainty are explored in the following simple example, in which two price-taking firms produce a homogeneous commodity that has a linear inverse demand curve. The slope of the curve is known with certainty, but the intercept (i.e. the choke price) varies depending on the state of the world, which is represented by the random parameter θ. The probability distribution over the different states of the world is denoted by $\Omega(\theta)$.

Figure 1.2 Pre-regulation market equilibrium



The marginal production costs of the firms are both linear – with a common intercept, but different slopes; and each firm has a constant specific emission rate (i.e. emissions per unit output).[8] Firm 1 has the lower production cost, and hence the higher sales prior to the regulatory intervention (see Figure 1.2). Firm 2 has the lower emission rate. Thus, for a given level of output, a reduction in aggregate emissions requires a

[8] Linear marginal production costs and constant specific emission rates are consistent with the firms having production functions of the form $y_i = \text{Min}\,[w_{1i}, w_{2i}^{1/2} / \alpha_i]$, and emission functions of the form $e_i = \varepsilon_i w_{1i}$.

redistribution of market share from firm 1 to firm 2. The pollutant is perfectly-mixable, and the marginal environmental damage of aggregate emissions is constant.

The regulator knows the marginal cost curves and the emission rates of the two firms, the marginal damage of emissions, and the probability function $\Omega(\theta)$. Consequently, it can determine the values of the emissions tax (t^*), absolute limit (L^*), and relative standard (r^*), that maximize expected net welfare under the respective forms of regulatory intervention. Of course, in the first case, the emissions tax is just set equal to the marginal damage of emissions, which is the same in all states of the world.

Table 1.3 shows the outcomes under the three different forms of intervention for the following parameter values and probability function:

- Marginal cost functions: $c^1(y_1) = 2 + y_1$ $\quad c^2(y_2) = 2 + 4\,y_2$
- Specific emission rates: $\varepsilon_1 = 4$ $\quad \varepsilon_2 = 1$
- Env. damage function: $D(e) = 0.2\,e$
- Inv. demand function: $p(Y) = (9 + \theta) - 0.5\,Y$ $\quad \theta \in \{-1, 0, 1, 2, 3\}$
- Probability function: $\Omega(\theta) = \frac{1}{5} \sum_{x=\theta}^{3} \frac{1}{x+2}$

The expected value of the random parameter is zero, and hence the choke price in the expected state of the world is € 9. The resultant "optimal" values for the regulatory parameters are: $t^* =$ € 0.2 for the tax; $L^* = 22.2$ for the limit; and $r^* = 3.33$ for the standard.

There are several points to note from Table 1.3. Expected net welfare is higher under the emissions tax than under either the absolute limit or the relative standard. Indeed, since the marginal damage of emissions is constant, the outcome under the tax is socially optimal in every state of the world. The relative performance versus the limit is of course to be expected. Weitzman (1974), and Adar & Griffin (1976) have demonstrated that when the marginal damage curve and the marginal abatement cost curve are both linear, and there is uncertainty about the intercept of the latter, expected welfare is higher under the emissions tax if the marginal abatement cost curve has the steeper slope. In this example, the marginal abatement curve is not considered explicitly; being defined implicitly by inverse demand function and the marginal cost functions of the two firms. However, since marginal damages are constant, it is guaranteed to be the steeper of the two.

Table 1.3 Comparison of regulatory interventions under uncertainty

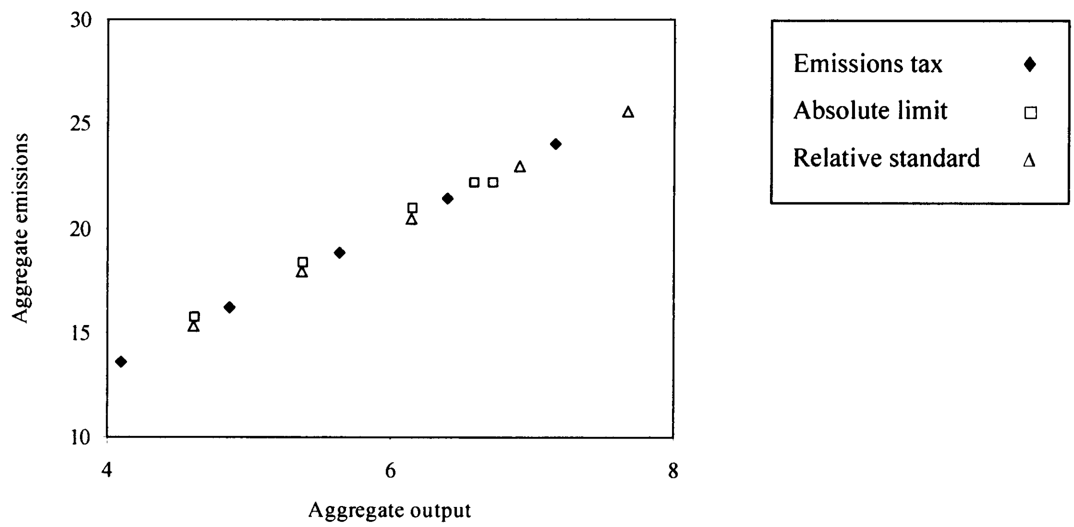
Choke price (a)		**12**	**11**	**10**	**9(*)**	**8**	**Expected**
Probability distribution		0.04	0.09	0.157	0.257	0.457	1.000
Emissions tax	Aggregate surplus	38.25	30.94	24.40	**18.63**	13.63	19.15
	Env. damages	4.80	4.28	3.76	**3.23**	2.71	3.23
	Net welfare	33.44	26.66	20.64	**15.40**	10.92	15.91
	% of max	100.0%	100.0%	100.0%	**100.0%**	100.0%	100.0%
Absolute limit	Aggregate surplus	37.73	31.07	24.62	**18.85**	13.85	19.32
	Env. damages	4.44	4.44	4.18	**3.66**	3.14	3.61
	Net welfare	33.29	26.63	20.43	**15.18**	10.71	15.72
	% of max	99.5%	99.9%	99.0%	**98.6%**	98.0%	98.8%
Relative standard	Aggregate surplus	38.38	31.09	24.56	**18.81**	13.82	19.32
	Env. damages	5.11	4.60	4.09	**3.58**	3.07	3.58
	Net welfare	33.27	26.49	20.47	**15.23**	10.75	15.74
	% of max	99.5%	99.4%	99.2%	**98.9%**	98.4%	98.9%

(*) Expected state of the world

The performance of the relative standard versus the absolute limit depends on the state of the world. In "high" states (i.e. $\theta = 2, 3$), the aggregate surplus is higher under the standard. However, so too is the environmental damage, and consequently net welfare is lower. The situation is reversed in "low" states (i.e. $\theta = -1, 0, 1$), with aggregate surplus and environmental damage both being lower, and net welfare higher. Overall, expected net welfare is slightly higher under the standard. It is also higher in the expected state of the world, and state with the greatest probability of occurrence.

Figure 1.3 provides an additional insight into the relative performance of the three forms of intervention. The diamonds (♦) show the combinations of emissions and aggregate output for the five different states of the world under the emission tax; while the squares (□) show the corresponding actual outcomes under the absolute limit; and the triangles (Δ) show the actual outcomes under the performance standard.

Figure 1.3 Relationship between aggregate emissions and aggregate output in different states of the world



Given the constancy of marginal environmental damage, the outcome under the emissions tax in each of the five states of the world coincides with the social optimum.

In "high" states of the world, the cost of reducing emissions (i.e. the surplus forgone) is greater. Hence it is optimal to allow more emissions, and to have higher output. In "low" states the opposite is true, and it is optimal to allow fewer emissions, and to have lower output.

The outcomes under the performance standard show a similar correlation with the random parameter; with output and emissions being greater in "higher" states of the world. However, in every state, both variables exceed the respective optimal values. In the three lowest states, the absolute limit is not binding, and there is little difference between the outcomes under the two quantitative interventions, although emissions are slightly higher under the limit. In the two highest states however, the absolute limit is binding, and hence emissions are capped at 22.2 tonnes. In the highest state this leads to a significant divergence between the outcomes, with output and emissions being too low (versus the social optimum) under the limit, and too high under the standard.

It is clear therefore that the best form of regulatory intervention in this example would be an (implicit) absolute limit, implemented by an emissions tax. However, if this is not politically feasible (i.e. the political constraint cannot be satisfied), then it is better to use a relative performance standard than an absolute limit. Expected net welfare is higher; as is the net welfare in the expected state of the world. Of course, this conclusion reflects the particular assumptions that have been made about the forms of the supply, demand and probability functions, and the choice of particular values for the parameters. Under other assumptions, the absolute limit would have been superior. However, none of the assumptions are particularly unreasonable, or unusual. In any case, the example demonstrates that when there is uncertainty about the marginal cost of abatement, the superiority of an absolute limit over a relative performance standard is no longer guaranteed.

1.2.2 Market power

The assumption of competitive markets may be a reasonable approximation of reality in many cases. However, there may be applications in which it is less justifiable. In particular, the firms that generate the emissions may act strategically in the "downstream" markets for the commodities that they produce, in order to influence the price that they receive. Ideally, this market failure should be addressed by a separate regulatory intervention. However, if this is not feasible, then it is necessary to take the failure into account when designing the environmental intervention.[9]

The implications of this type of imperfection are explored for the extreme case of a single firm that can exercise monopoly power to set the market price of its output. The firm produces this output using a mix of two Leontief processes ($m = 1, 2$); each having a constant marginal cost (c_m), and a constant specific emissions rate (ε_m).[10] Without any loss of generality, it is assumed that $c_1 < c_2$, and that $\varepsilon_1 > \varepsilon_2 = 0$. Consequently, in the absence of any regulatory intervention, the firm use only process 1 to produce its output, and its total cost function is $C(Y) = c_1 Y$.

Post-regulation, the firm's conditional total cost function depends on the form of the intervention, as follows:

Emissions tax: $$C(Y; t) = (c_1 + \varepsilon_1 t) Y \quad \text{if } t \leq (c_2 - c_1)/\varepsilon_1$$

$$= c_2 Y \quad \text{if } t > (c_2 - c_1)/\varepsilon_1$$

Absolute limit: $$C(Y; L) = c_1 Y \quad \text{if } \varepsilon_1 Y \leq L$$

[9] This was first recognized by Buchanan (1969) in relation to Pigouvian taxes.

[10] This is consistent with an assumption that for each process, emissions are a linear function of production inputs.

$$= \quad c_2 Y \; - \; \left(\frac{c_2 - c_1}{\varepsilon_1}\right) L \qquad \text{if } \varepsilon_1 Y > L$$

Relative standard: $C(Y; r) \quad = \quad \left[c_1\left(\frac{r}{\varepsilon_1}\right) + c_2\left(1 - \frac{r}{\varepsilon_1}\right)\right] Y$

Thus under the emissions tax, if the tax rate is sufficiently high, then the firm is better off switching all of its production to process 2; in which case its emissions fall to zero. Conversely, under the absolute limit, if the limit is sufficiently high (relative to output), then the firm can continue to produce all of its output using process 1. Under the relative standard however, the firm always uses both production processes, although the relative mix depends on the stringency of the standard.

Consequently, for given values of the tax (t), and the standard (r), the respective marginal costs of production are constant – lying somewhere between the costs of the two processes (c_1 and c_2).[11] However, under the absolute limit, the marginal cost has a step at a threshold level of output; being equal to c_1 below the threshold, and equal to c_2 above. The three cases are shown in Figure 1.4, assuming that the values of the tax and the standard have been set to induce the same level of emissions as under the limit.

Two points are clear in Figure 1.4. First, if the marginal revenue curve intersects the marginal cost curve for process 2 to the right of the step, then the equilibrium output under the limit is the same as that under the equivalent tax, and lower than that under the equivalent standard.[12] Second, in this situation the impact of a change in the value

[11] The marginal cost may be equal to c_2 under the tax, but is always strictly less than c_2 for the standard, assuming that this is greater than zero.

[12] Assuming that the absolute limit is strictly positive, then so too is the value of the equivalent relative standard – i.e. the standard that induces the same level of emissions. Consequently, the weighted average

of the regulatory variable depends on the form of the intervention. A decrease in the value of the absolute limit (L) reduces the output threshold at which the step occurs, but does not affect the height of the step. Consequently, the equilibrium output level remains unchanged. Similarly, an increase in the tax rate (t) has no impact on the marginal cost, since this is already at its maximum value. In contrast, a reduction in the relative standard (r) causes the marginal cost to shift upwards, and hence output decreases.

Figure 1.4 Market equilibrium

a) absolute limit **b) emissions tax / relative standard**

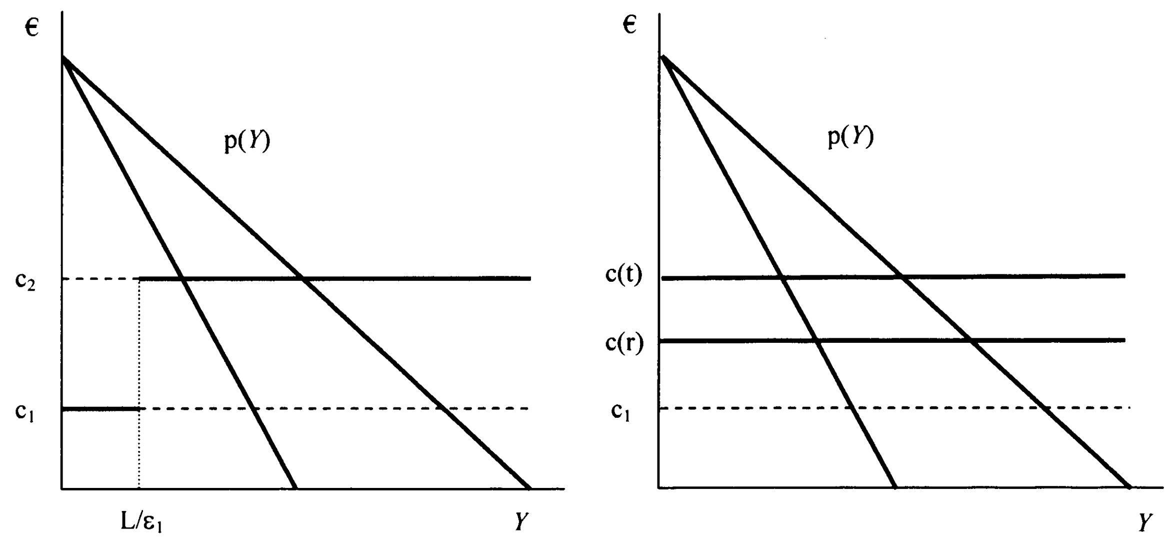


Figure 1.5 and Table 1.4 compare the outcomes of the under the three forms of intervention for the following parameter values, and functional forms for environmental damages and inverse demand:

- Marginal costs of production $c_1 = 10$ $c_2 = 30$
- Emission rates: $\varepsilon_1 = 3.333$ $\varepsilon_2 = 0$

---

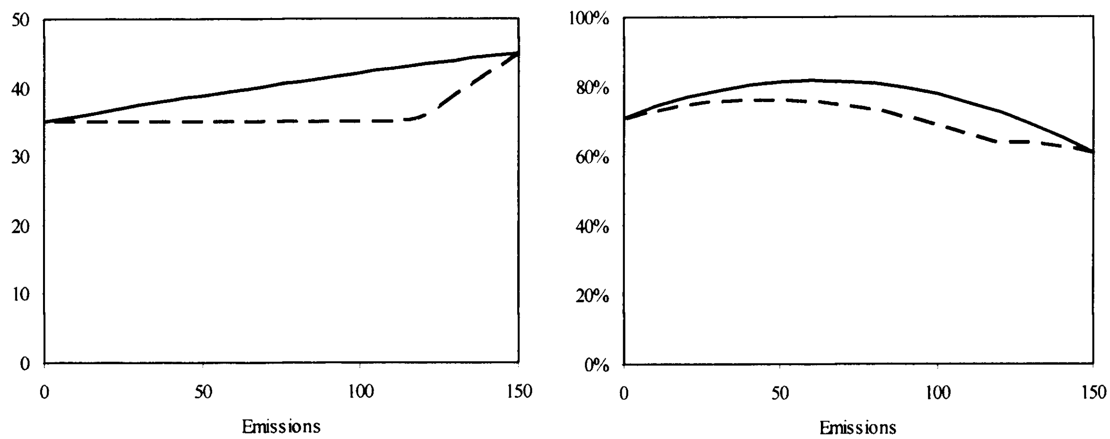
of the two marginal costs is less than c_2. Assuming that the marginal revenue curve is downward sloping,

- Env. Damage function: $D(e) = 0.065\, e^2$
- Inverse demand function: $p(Y) = 100 - Y$

For these parameter values, the socially optimal level of output (i.e. in the absence of any market distortions) is 70,000 units, with emissions equal to 46 tonnes. Prior to the regulatory intervention, the actual output level is 45,000 units, and emissions are equal to 150 tonnes. Thus, while emissions are too high, output is already too low.

Figure 1.5 Simulation results

a) output (000's) **b) % of maximum welfare**



——— relative emissions standard

— — emissions tax / absolute emissions limit

There are several points to note from the simulation results. First, as one would expect, there is an equivalence between the emissions tax and the absolute limit, with the "optimal" tax rate (or permit price) being less than marginal damage of emissions at the

---

this implies that output is higher.

optimum – which is equal to €6.5 per tonne of emissions.[13] However, it should be noted that while attainment of the optimum is guaranteed under the emissions limit, it is not under the tax. This is because, for the optimal value of the tax, the firm is indifferent between the two processes. Depending on the actual process mix that it chooses, emissions can take any value between 0 and 110 tonnes. Second, for any given environmental outcome, output is higher under the relative performance standard. Thus even when the marginal revenue curve intersects the vertical step of the marginal cost curve under the absolute limit, the marginal cost is lower under an equivalent relative standard. Third, this implies that the maximum net welfare that can be achieved (shaded grey in Table 1.4) is higher under the performance standard than under the limit / tax; as are the "optimal" values of output and emissions. However, under all three forms of intervention, the output is less than the pre-regulation level, despite the fact that this is already below the socially optimal level.

Thus in this example, the relative performance standard is superior to both the absolute limit and the tax. The reason for this superiority is the implicit subsidy that it provides to output. As has been discussed above, this is a handicap under perfectly competitive conditions, preventing the attainment of the optimal outcome. However, in this case where output is too low, it is a positive advantage. Again, the relative outcomes under the different forms of intervention reflect the particular assumptions that have been made about the production technology of monopolist, and the particular choices for parameter values. Under different assumptions, the absolute limit / tax would have been superior. In particular, this would be the case if the value of marginal damages was higher – penalising the higher emissions under the relative standard. Nonetheless, it is clear that where a polluter has monopoly power, it may well be preferable to impose a relative standard rather than an absolute limit or an emissions tax.

[13] Barnett (1980) demonstrated that when an emissions tax is applied to a monopolist, the (second-best) optimal value of the tax is lower than the resultant marginal damage of emissions.

Table 1.4 Simulation results

a) Emissions tax or absolute emissions limit

Tax (t)	Output (000 units)	Emissions (tonnes)	Total Surplus	Env. Damage	Net Welfare	% of Maximum
0.0	45	150	3038	1463	1575	61%
1.8	42	140	2898	1274	1624	63%
3.6	39	130	2750	1099	1651	64%
5.4	36	120	2592	936	1656	64%
6.0	35	110	2498	787	1711	66%
6.0	35	100	2438	650	1788	69%
6.0	35	90	2378	527	1851	72%
6.0	35	80	2318	416	1902	73%
6.0	35	70	2258	319	1939	75%
6.0	35	60	2198	234	1964	76%
6.0	35	50	2138	163	1975	76%
6.0	35	40	2078	104	1974	76%
6.0	35	30	2018	59	1959	76%
6.0	35	20	1958	26	1932	75%
6.0	35	10	1898	7	1891	73%
6.0	35	0	1838	0	1838	71%

b) Relative emissions standard

Standard (r)	Output (000 units)	Emissions (tonnes)	Gross Surplus	Env. Damage	Net Welfare	% of Maximum
3.33	45	150	3038	1463	1575	61%
3.15	44	140	2964	1274	1690	65%
2.96	44	130	2889	1098	1791	69%
2.77	43	120	2814	936	1878	73%
2.57	43	110	2738	786	1952	75%
2.37	42	100	2661	650	2011	78%
2.17	42	90	2584	527	2058	79%
1.96	41	80	2506	416	2090	81%
1.74	40	70	2427	318	2108	81%
1.52	40	60	2346	234	2112	82%
1.29	39	50	2265	163	2103	81%
1.05	38	40	2183	104	2079	80%
0.80	37	30	2099	58	2040	79%
0.55	37	20	2013	26	1987	77%
0.28	36	10	1926	6	1920	74%
0.00	35	0	1838	0	1838	71%

Of course, this monopoly case may not appear very relevant to the design of a trading-mechanism, since the problem of determining the efficient allocation of abatement effort between firms does not arise. However, it should be noted that similar results would be obtained if the regulatory intervention applied to a group of non-competing firms, each of which had monopoly power in its own downstream product market.[14]

1.2.3 Distortionary taxes

While it is possible to argue that uncertainty about economic conditions can be reduced by shortening the lead times involved in the development and implementation of regulatory interventions, and that market imperfections can be addressed directly by other regulatory interventions, it is hard to justify the exclusion of distortionary taxes from the analysis. In particular, taxes on labour are (virtually) universal. Consequently, a meaningful comparison of alternative forms of regulatory intervention should really take this into account. There are a small number of studies that have done so; using computable general equilibrium models to compare the efficiency of different forms of intervention in the presence of labour taxes.

Parry & Williams III (1999) assess the relative costs of a number of different regultory interventions to reduce carbon dioxide (CO_2) emissions in the USA, including (*inter alia*) an emissions tax, an efficiently implemented aggregate emissions limit, and an efficiently implemented performance standard for specific emissions.

The model is relatively simple; having a single representative consumer, whose net welfare function is strongly separable in aggregate CO_2 emissions, and weakly separable on two consumption goods and leisure. In addition to the two final good sectors, there are seven other production sectors: three producing fossil fuels; three

[14] Strictly speaking, the results can only be translated directly if the firms are identical, so that the optimal limit (L^*) and the optimal standard (r^*) are the same for each firm. However, by continuity, they will also hold for heterogeneous firms, provided that the differences are not too great.

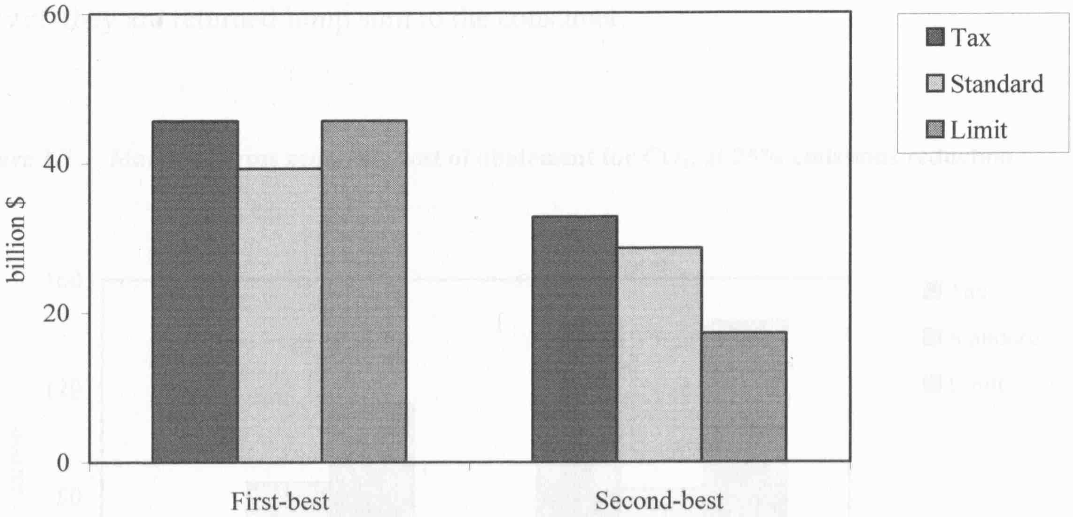
producing energy-intensive intermediate goods; and one sector producing a composite intermediate good with low energy content. While the nine production sectors are heterogeneous, the model assumes that firms are homogeneous within each sector. Emissions arise when the fossil fuels are used in the production of intermediate and final goods, with aggregate emissions being a linear function of the input quantities. The government levies a tax on labour and profits remitted to the consumer (including the rents arising under the absolute limit), and uses the revenue to provide a lump sum transfer to households. Finally, the markets for all goods – final and intermediate – are assumed to be perfectly competitive.

The structure of the model allows for a range of responses to a regulatory intervention. Within the fossil fuel sector, there may be a shift from coal (with a high carbon content) to natural gas (with a low carbon content). Within the intermediate production sector, there may be a shift away from the energy-intensive goods. Finally, there may be a shift in consumption away from the energy-intensive final good, or an increase in leisure at the expense of consumption.

The regulatory intervention applies to all intermediate and final production sectors. For the absolute limit and the relative standard it is assumed that the private marginal cost of emissions reduction is equated across industries. With homogeneity of firms within sectors, this is the same as assuming that the regulatory targets are implemented efficiently in each case.

Figures 1.6 shows the maximum net welfare gain that can be achieved under the three forms of regulatory intervention, assuming that marginal environmental damages are equal to \$100 per ton of carbon. In the "first-best" world (i.e. with no distortionary labour taxes), the maximum gain in net welfare is achieved either by setting the emissions tax equal to marginal damages, or by setting the aggregate number of permits equal to the optimal emissions level. The maximum gain that can be achieved with a performance standard is always less than this – confirming the analytical findings of Ficsher (2001) and Gielen *et al* (2002).

Figure 1.6 Maximum net welfare gain when marginal environmental damages are equal to $100 / ton carbon

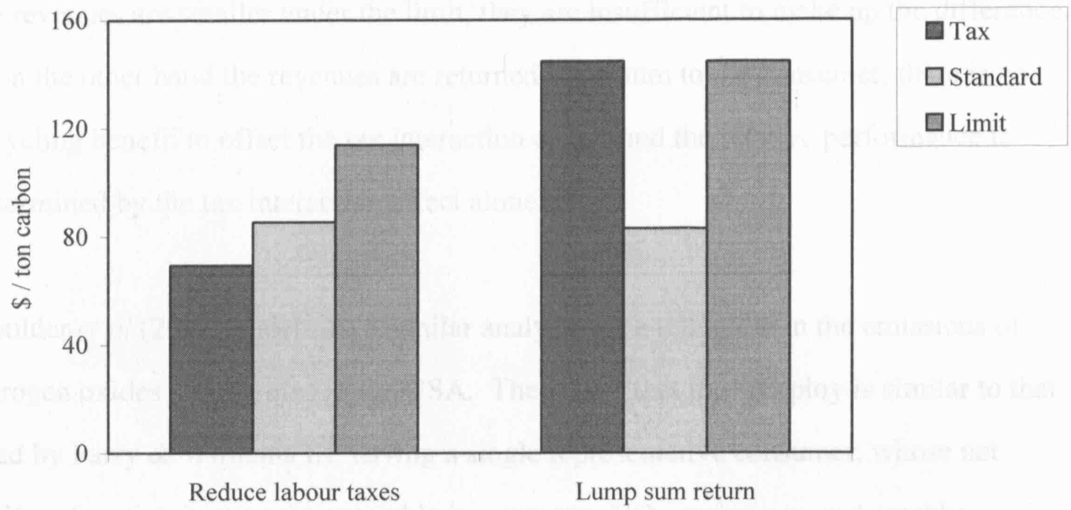


Source: Figure 4 and Figure 5, Parry & Williams III (1999)

However, in the "second-best" world of distortionary labour taxes, the picture is rather different. The first point to note is that, for all three forms of intervention, the maximum welfare gain that can be achieved is lower than in the "first-best" case. That is, the presence of distortionary labour taxes raises the gross economic cost of all forms of regulatory intervention. The second point is that the emissions tax and the absolute limit are no longer equivalent; with the maximum gain under the tax now being almost twice as large as that under the limit. Furthermore, while the maximum gain under the performance standard is still not as great as under the tax, it is considerably greater than that achieved under the absolute limit.

The relative performance of the three forms of regulatory intervention in the second-best setting is driven largely by two tax-related effects: a *tax interaction* effect, and a *revenue recycling* effect. This can be seen more clearly in Figure 1.7, which shows the gross economic cost of achieving a 25% reduction in CO_2 emissions, for two different assumptions regarding the use of emissions tax revenues, In the first case, revenues are used to reduce the distortionary labour taxes (as is the case in Figure 1.6). In the second, they are returned lump sum to the consumer.

Figure 1.7 Marginal gross economic cost of abatement for CO_2, at 25% emissions reduction



Source: Figure 3 and Table 3, Parry & Williams III (1999)

Since all three forms of intervention raise the cost of production, they increase the relative price of consumption goods and reduce the real wage. This causes a reduction in the supply of labour, which exacerbates the deadweight loss arising from the labour tax. The scale of this interaction effect is greater for the emissions tax and the absolute limit, than it is for the relative standard. This however is only one half of the story. The overall impact depends on the extent to which the intervention raises revenues, and the use to which these are put.

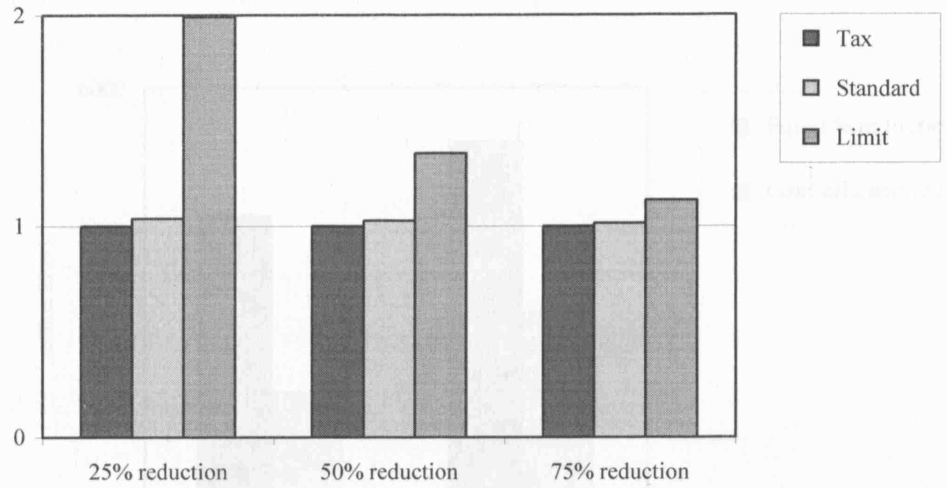
Revenues are raised under both the emissions tax and the absolute limit. In the latter case, these derive from the taxation of the resultant rents (at the pre-existing tax rate) rather than the direct taxation of emissions. In contrast, there is no impact on profits under the performance standard, and hence no revenues are raised.[15] If these revenues are used to reduce the distortionary labour tax, then there is an offsetting benefit under the emission tax and the absolute limit. In the case of the tax, this is sufficiently large to outweigh the tax interaction advantage of the performance standard. However, because the revenues are smaller under the limit, they are insufficient to make up the difference. If on the other hand the revenues are returned lump sum to the consumer, there is no recycling benefit to offset the tax interaction effect, and the relative performance is determined by the tax interaction effect alone.

Goulder *et al* (2000) undertake a similar analysis for a reduction in the emissions of nitrogen oxides (NO_X), also in the USA. The model that they employ is similar to that used by Parry & Williams III; having a single representative consumer, whose net welfare function is strongly separable in aggregate NO_X emissions, and weakly separable in two consumption goods and leisure. This time there are only two intermediate production sectors. One of the intermediate products (termed the dirty good) causes emissions when it is used as a production input, while the other does not. The emissions of each sector depend on the quantity of the dirty good used in production, and its expenditure on end-of-pipe abatement (in the form of additional labour input). As with the previous model, differences between sectors are recognised, but it is assumed that firms are homogeneous within each sector.

[15] This is due to the assumption of nested CES production functions, with constant returns to scale.

Figure 1.8 shows the relative gross economic costs of the alternative forms of intervention for three levels of emissions reduction, assuming that the emission tax revenues (and the revenues raised from the taxation of rents under the absolute limit) are used to reduce labour taxes. The costs are expressed as ratios, relative to the cost under the emissions tax for each level of reduction.

Figure 1.8 Ratio of gross economic cost under each policy intervention to cost under emissions tax, for different levels of emission reduction

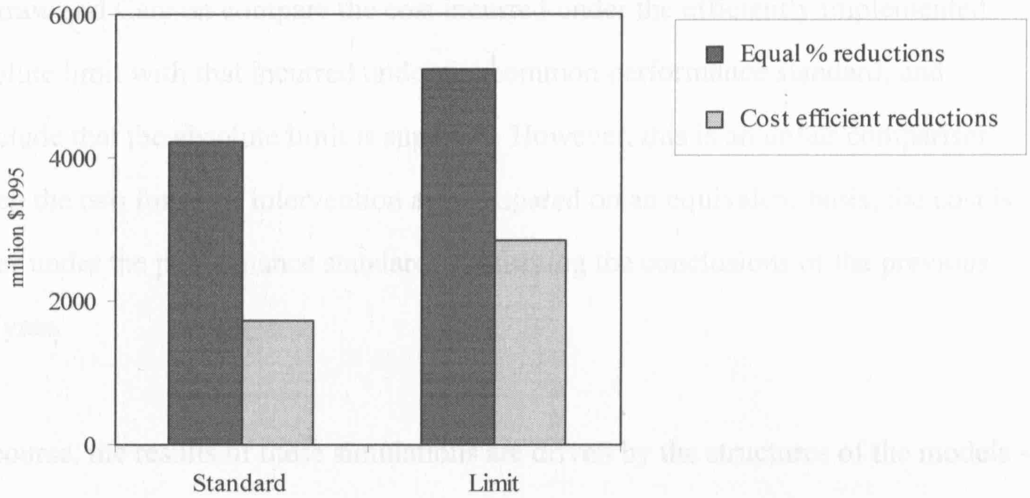


Source: Table 4, Goulder *et al* (1999).

As in the previous study, the gross economic cost is lowest under the emissions tax. However, the cost is only marginally higher under the performance standard. In contrast, the cost under absolute limit is significantly greater, particularly for relatively low levels of emissions reduction. For a 25% reduction, the cost under the absolute limit is almost double that under the performance standard. As the reduction target becomes more stringent, the differences between the three forms of intervention become less marked, and the cost under the limit is only 11% greater for a 75% reduction in emissions.

Burtraw & Cannon (2000) also compare alternative regulatory interventions to achieve reductions in NO_X emissions in the USA, using a model that has five intermediate production sectors. The motivation for their analysis is to investigate how alternative methodologies for representing abatement costs in the model can affect the comparison of the interventions. However, it is possible to use their results to show how the costs incurred under an absolute limit, and a relative standard, are affected by the choice of implementation mechanism.

Figure 1.9 Total gross economic cost of a 25% reduction in NO_X emissions



Source: Table 3, Burtraw & Cannon (2000)[16]

[16] Burtraw & Cannon (2000) consider three alternative ways of representing abatement costs, which they term heterogeneous, homogeneous, and lower envelope. In theory, the outcomes for the emissions tax and the efficiently implemented absolute limit should be the same for the heterogeneous and lower envelope representations, while the outcomes for the common performance standard should be the same under the heterogeneous and homogeneous representations. However, for technical reasons that they discuss in the paper, there is a discrepancy between the values that they show in their Table 3. Consequently, for the equal % reduction, the figure used in Figure 1.9 for the performance standard is equal to the average of the figures for the heterogeneous and the homogeneous abatement cost cases. For the cost efficient reduction, the figure used for the absolute limit is equal to the average of the figures for the heterogeneous and the lower envelope case.

Figure 1.9 shows the gross economic costs incurred under the absolute limit and the relative standard, for two implementation "plans". Under one plan, limits and standards are set for the sectors, so that each reduces its emissions by 25%.[17] Under the other plan, limits and standards are set for each sector, so as to minimize the total gross economic cost of achieving a 25% reduction in aggregate emissions. It is clear from this that the efficiency of the implementation mechanism is a crucial factor in determining the cost of the regulatory intervention. For the absolute limit the cost is 45% lower when the limits are allocated efficiently. For the performance standard, the cost is 60% lower.

Burtraw and Cannon compare the cost incurred under the efficiently implemented absolute limit with that incurred under the common performance standard, and conclude that the absolute limit is superior. However, this is an unfair comparison. When the two forms of intervention are compared on an equivalent basis, the cost is lower under the performance standard; confirming the conclusions of the previous analyses.

Of course, the results of these simulations are driven by the structures of the models – particularly the use of nested-CES production functions and a separable utility function, and the choice of parameter values. However, these assumptions are commonly adopted in computable general equilibrium models of this type, and in each case the models are calibrated to actual data, and use empirically derived estimates of the parameter values. Consequently, it seems reasonable to conclude that in the presence of distortionary labour taxes, the gross economic cost of achieving an absolute environmental objective may be lower under a relative performance standard than under an absolute limit, assuming that both are implemented efficiently.

[17] In the case of the performance standard, this is not quite the same as setting a common standard for all sectors. However, it would be so if all sectors experience the same percentage reduction in their output.

1.3 Performance standards in environmental policy

The previous section provided some theoretical arguments to motivate the analysis of a market-based mechanism for relative performance standards. Another, possibly more compelling reason for doing so, is the popularity of relative performance standards in practice. Numerous examples may be found, in a wide range of different environmental policy areas. Three particular applications are described in some detail in section 1.4 – which reviews the use of trading mechanisms for performance standards. A number of other examples – from both sides of the Atlantic – are outlined briefly in the following paragraphs. The intention is not to provide an exhaustive inventory. Rather, the objective is to illustrate the range of different applications in which performance standards have been, or are being used.

- In Europe, the 1994 EU Directive on Packaging and Packaging Waste set a target recovery rate of 50% for aggregate waste packaging, to be achieved by all member states by the year 2001, with subsidiary target recycling rates of 18% for individual materials (e.g. glass, plastic, etc.).

- In England and Wales, targets for the proportion of household waste that is recycled and composted by local authorities have been set under the so-called "Best Value" legislation[18]. These statutory targets, which apply to the fiscal years 2003/04 and 2005/06, are differentiated between authorities based on their actual recycling rates in 1998/99.

- In the USA, regulations were introduced in the mid 1970s, which imposed corporate average fuel economy (CAFE) standards for all new passenger vehicles.[19] The current standard is 27.5 mile per gallon of gasoline (or equivalent

[18] Local Government (Best Value) Performance Indicators and Performance Standards Order 2001.

[19] Energy Policy and Conservation Act, 1975 (P. L. 94 – 163).

amount of other fuel) for passenger cars, and 20.7 mpg for light trucks. The standards apply to the annual sales of each manufacturer and importer, based on model years. Firms earn credits by selling alternative fuel (and duel fuel) vehicles, which can be used to boost their average performance. They may also carry forward credits from any year in which they exceed the standard, which can be used to boost performance in succeeding years.

- In Europe a target has been set for the average carbon dioxide (CO_2) emissions per kilometre of all new vehicles sold in the European Union. The standard of 140 gCO_2 per km – which is enshrined in a series of negotiated agreements between the European Commission and vehicle manufacturer associations in 1999 and 2000 – is to be achieved by 2008/09.[20]

- In the Netherlands, the Energy Efficiency Policy Document (1990) set a target for industry of a 20% reduction in the specific energy consumption (i.e. energy use per unit output) by the year 2000, compared with a base year of 1989. This was implemented by as series of Long Term Agreements (LTAs) with industry sectors and individual companies. By the end of 1998, forty-one agreements – which are binding under civil law – had been signed, covering a diverse range of industries and service sectors.

- In the United Kingdom, forty-four energy intensive sectors have been granted an 80% exemption from the Climate Change Levy (and industrial energy tax) in return for entering into legally binding negotiated agreements with the government. Thirty-seven of these Climate Change Agreements (CCAs) set

[20] An agreement was signed with the European Automobile Manufacturers Association (ACEA) in February 1999 (1999/125/EC). Further agreements were signed with the Korea Automobile Manufacturers Association (KAMA), and the Japan Automobile Manufacturers Association (JAMA) in April 2000 (2000/ 303/EC and 2000/304/EC). The target year is 2008 for the ACEA agreement, and 2009 for the other two.

targets for specific energy consumption for the year 2010, with interim targets every two years. The structure of an agreement can take several different forms. In the vast majority of cases there is an "umbrella" agreement with the sector association, and a series of "underlying agreements" between the government and the constituent target units, that set individual targets.[21] Performance is reviewed at each of the interim target dates. If the sector target is met, then all units are "re-certified" for the next two years. If it is not, then the exemption is only extended for those units that have met their individual targets.

As can be seen from this brief review, relative performance standards can take many different forms. Furthermore, they may be expressed explicitly, or implicitly as improvements versus a base year value. However, essentially they fall into the following three broad categories:[22]

- proportions $$\frac{w}{w + x + y} \leq r$$
- weighted averages $$\frac{\alpha w + \beta x + \chi y}{w + x + y} \leq r$$
- rates $$\frac{w + x + y}{a + b + c} \leq r$$

While this taxonomy provides a useful framework for classifying performance standards, it is not necessarily mutually exclusive. In certain circumstances it is possible for a target that is expressed as one type, to be reformulated as another. For example, if there are only two components included in the definition of a weighted

[21] A target unit is an industrial facility, or group of facilities, sharing one target.

[22] In all three cases, the target is expressed in terms of a maximum value that must not be exceeded. However, it is always possible to define an equivalent target in terms of a minimum value. For example, a 60% target for the maximum proportion of waste that is sent to landfill is exactly equivalent to a 40% target for the minimum proportion that is diverted.

average, then it is always possible to reformulate the target in terms of a proportion (and vice versa). That is:

$$\frac{\alpha w + \beta x}{w + x} \leq r \qquad \Leftrightarrow \qquad \frac{w}{w + x} \leq \frac{r - \beta}{\alpha - \beta} = \hat{r}$$

Similarly, if the denominator ($w + x + y$) in the proportion and the weighted average is directly proportional to ($a + b + c$), then it is always possible to reformulate a target for a proportion, or a weighted average, in terms of a rate (and vice versa). That is:

$$\frac{\alpha w + \beta x + \chi y}{w + x + y} \leq r \qquad \Leftrightarrow \qquad \frac{\alpha w + \beta x + \chi y}{a + b + c} \leq \left(\frac{w + x + y}{a + b + c}\right) r = \tilde{r}$$

and

$$\frac{w}{w + x + y} \leq r \qquad \Leftrightarrow \qquad \frac{w}{a + b + c} \leq \left(\frac{w + x + y}{a + b + c}\right) r = \tilde{r}$$

In practice, the first circumstance is unlikely to arise, as there will usually be many component terms in the weighted average. For example, in the case of the CAFE standard, the number of terms is equal to the number of different vehicle models that the manufacturer produces. The second circumstance is more likely to apply, particularly in the case of targets related to the recycling of products and materials. For example, in the case of packaging recovery targets, the total quantity of packaging discarded in any year is (approximately) equal to the amount produced. Even for durable products, such as vehicles or electronic equipment, there may be a predictable relationship between annual sales and disposals, if the "survival profile" is stable and sales are growing at a constant rate.[23] In these cases, a recycling target expressed as a proportion of the total amount discarded is completely equivalent to one expressed as a rate relative to the total amount produced.

---

[23] The survival profile shows the percentage of products sold at time T that are still in use at time T + t

1.4 Trading mechanisms with performance standards

To date the majority of relative performance standards have been implemented either through a system of fixed individual standards – which may be common or differentiated between agents, or by negotiated agreements with industry associations that do not specify individual targets. While the implementation of the CAFE standards exhibits some of the characteristics of a trading system – i.e. the averaging of a manufacturer's performance over its entire model range, and the carrying forward of credits for over-performance, there is no flexibility between manufacturers. However, there are a small number of examples where "external" trading mechanisms have been introduced, or proposed, for the implementation of relative performance standards.

In recent years in the United Kingdom, trading mechanisms have been introduced to implement the energy efficiency targets in the Climate Change Agreements, and to implement the recovery targets for waste packaging set down in the EU Packaging Directive. These are described in detail in Chapter 5 and Chapter 6 respectively. Three other examples are described briefly here. The first is the *lead credit trading* programme that operated in the USA between 1982 and 1987. This is probably the most well known example of trading with a relative target. The second is the credit trading scheme for *relative NO_X emissions* that is due to start in the Netherlands in 2004. The final example, relating to the *recycled content of newsprint*, has not been implemented, although it was given serious consideration by the US Congress in the early 1990's.

Lead credit trading in the USA

In the mid 1970s, the US Environmental Protection Agency (EPA) introduced a regulation[24] which required that the average lead content of all gasoline sold in the USA

[24] EPA, Control of Lead Additive in Gasoline, 38 Fed. Reg. 33 734 (1973) (final note)

be reduced from 1.7 grams per gallon after 1st January 1975, to 0.5 grams per gallon by 1st January 1979 – with the rates applying to the quarterly production of individual refineries. In 1982, the rate was tightened to 1.1 grams per gallon for leaded gasoline. At the same time, the EPA introduced an "inter-refinery averaging" programme, commonly known as the *lead trading* programme. The rational for introducing the programme was that the added flexibility would alleviate the difficulty of meeting the new standard, particularly for smaller refineries.

Under the programme, if a refinery produced gasoline with a lead content below the standard, then it was entitled to sell lead rights (or credits), up to an amount equal to the difference between amount of lead allowed and the amount actually used. In contrast, if the lead content of its production exceeded the amount allowed, then it would have to purchase sufficient credits to make up the difference. For example, if refinery A produced 100 million gallons with an average lead content of 0.9 grams per gallon, and refinery B produced 50 million gallons with an average content of 1.5 grams per gallon, then refinery A could sell (up to) 20 million grams of credits (i.e. $(1.1 - 0.9) \times 100$), while refinery B was required to purchase (at least) 20 million grams of credits (i.e. $(1.1 - 1.5) \times 50$).

At the end of each quarter, refineries were required to report to the EPA their gasoline production and lead usage, together with information on the quantity of credits bought and / or sold, and the identity of their trading partners. Credits expired at the end of the quarter in which they were generated, and hence unused credits were not allowed to be carried forward. In addition to refineries, the programme also applied to gasoline importers, and to blenders who could generate credits by adding ethanol to gasoline, and hence reduce its lead content.

In 1985, the EPA lowered the standard to 0.5 grams per gallon, and announced that it would be further reduced to 0.1 grams per gallon form the beginning of 1986, at which time the trading programme would be terminated. However, later in the year, the EPA extended the life of the programme, when it retrospectively introduced *lead banking* for credits generated during 1985, and announced that refineries would be allowed to trade and use banked credits until the end of 1987.

The general consensus is that the programme was a success, and that it allowed the lead content of gasoline to be phased out much sooner than would otherwise have been feasible (USEPA, 2001; Hahn & Hester, 1989). The market in lead credits was very active, with levels of activity increasing through the life of the progamme. By the end of 1987, the quantity of credits traded was more than 40% of the total lead use; around 60% of refineries had engaged in lead trading; and around 90% had taken advantage of the banking provision. While the cost savings realized by the programme have never been fully calculated, the EPA has estimated that the banking provision alone resulted in savings of $226 million (USEPA, 2001).

Hahn & Hester (1989) provide a comprehensive evaluation of the programme, and draw some conclusions regarding the reasons for its success. Their analysis of the quarterly returns to the EPA revealed that – contrary to the initial expectations – the market for lead credits was not one-sided, with the large refineries dominating supply. Up until the end of 1985, small refineries were just as likely to participate in the market as sellers of lead credits as they were to participate as buyers. However, for the final two years of the programme, there was a significant net flow of credits from large to small refineries, reflecting the much greater use of the banking provision by large refineries in 1985.

They attribute the success of the programme to three factors. The first two relate to the design of the programme. First, there were only minimal restrictions imposed on trading – particularly after the introduction of banking in 1985. Second, the monitoring and reporting requirements did not impose a significant administrative burden. However, probably more important than these factors was the fact the refineries already traded feedstocks and gasoline additives. Consequently, the trading of lead credits did not represent a major innovation.

NO_X trading in the Netherlands

In response to the EU Directive on national emission ceilings[25], the Dutch government set a target for aggregate NO_X emissions by large industrial installations of 55 kilotonnes per annum, to be achieved by the year 2010. This represents a 39% reduction versus the emissions level in 2000. While this environmental objective is defined in absolute terms, the regulatory intervention that has been adopted takes the form of a relative performance standard, implemented by a credit trading scheme. The underpinning legislation for the scheme was launched by the Dutch Ministry of the Environment (VROM) in March 2003, and it is expected to come into force in the middle of 2004 (Dekkers, 2003).

Under the proposed scheme, a common performance standard rate (PSR) – defined in terms of annual NO_X emissions per unit of energy input – will apply to every industrial installation with energy consumption exceeding 20 MWth. On the basis of a projected energy consumption of 1100 Petajoules in 2010, the PSR for that year has been set at 50 g/GJ. Thus, provided that the projection turns out to be correct, the environmental objective will be achieved. Intermediate targets have been set for the years 2004-2009; starting at 65 g/GJ, and declining steadily to the final target. The scheme allows for the projected energy use figure to be reviewed in 2006, and for the PSR values for the final four years to adjusted downwards if necessary, in order to ensure that the 55 kilotonne objective is met in 2010.

The scheme will apply to around 250 facilities with installed thermal capacity exceeding 20 MWth. If a facility achieves an emissions rate that is lower than the PSR for that

[25]Directive on National Emission Ceilings for Certain Atmospheric Pollutants (2001/81/EC)

year, then it is allowed to generate credits equal to the difference, multiplied by its energy consumption. These credits are denominated in tonnes of NO_X. For example, if a facility using 4 Petajoules of energy achieves an emissions rate of 40 g/GJ in 2010, then it will be allowed to generate 40 tonnes of NO_X credits (i.e. $(50 - 40) \times 4$). In contrast, if its emissions rate is 55 g/GJ, it will have to purchase 20 tonnes of NO_X credits from other facilities that have beaten the target.

In addition to the trading of credits between facilities, the proposed scheme allows for a limited amount of banking and borrowing between years. Up to 5% of a facility's annual allocation (i.e. PSR × energy consumption) may be transferred to the preceding, and the succeeding year. At the end of each year, each facility will have to provide a report to the Netherlands Emission Authority (NEA), giving details of its NO_X emissions, energy consumption, and any credits that it has purchased or sold.[26]

An interesting aspect of the proposed scheme is that it will operate along side existing legislation, rather than replace it. In particular, the facilities will still be subject – where appropriate – to regulation under the Large Combustion Plant (LCP) and Integrated Pollution Prevention and Control (IPPC) Directives, and will have to comply with all emission limits specified in their operating permits under these regulations. On the assumption that the annual allocations under the 50 g/GJ target are more stringent than the existing limits, there will still be gains from trading. However, by imposing a ceiling on the demand for NO_X credits, the constraints may prevent all of the potential gains from being achieved.

[26] There will be a three month reconciliation period after the end of each year, in order to allow the facilities time to purchase any credits that they may need to meet the target.

Recycled content of newsprint

Targets for the recycled content of newsprint have been set on both sides of the Atlantic.[27] For example, in the United Kingdom, the government has entered into a negotiated agreement with the newspaper publishers, which sets a 70% collective target for the minimum proportion of newsprint that is produced from recycled pulp by the end of 2006. In the USA, individual newspaper publishers and printers in California are legally mandated to ensure that 50% of the newsprint that they consume is classified as recycled-content newsprint (RCN) – i.e. at least 40% of the pulp used in its production is from post-consumer waste.

These targets are implemented respectively by a contract-based mechanism and a consent-based mechanism. However, in both countries, consideration has been given to the idea of using a market-based mechanism. In the United Kingdom, a Private Member's Bill on Newspaper and Magazine Recycling was introduced in 2000, which *inter alia* would have imposed aggregate percentage targets for the collection and recycling of waste newspapers and magazines, and for the recycled content of newspapers. During the Bill's debate, the Department of the Environment, Transport and the Regions (DETR) floated the idea of using a system of tradable credits to implement the targets.[28] However, the Bill was eventually defeated, with the government entering into the negotiated agreement with the newspaper publishers, and the trading idea was not developed.

Eight years earlier, a bill was considered by the US Congress that would have required the producers and importers of newsprint to use a steadily increasing percentage of recycled pulp, with the target implemented by a tradable credit scheme. However, unlike the UK proposal where the newspaper publishers would have had the

[27] Newsprint is the blank paper on which newspapers are printed. That is, it is an input to the production of newspapers.

[28] ENDS Report, Issue No. 303, April 2000

responsibility for achieving the targets, the scheme would have been applied to the producers of newsprint.

The proposal – which again did not reach the statute books – is described and analysed by Dinan (1992). A minimum recycled content standard (RCS) would be set for the proportion of newsprint produced from recycled pulp in the USA, or imported form overseas. Initially this was to be set at 20%, but after 18 months it would rise at 2% points per annum, until it reached 30%. Under the trading scheme, a newsprint producer would be granted property rights to 0.8 credits for every ton of newsprint that it produced using recycled pulp, and would have to redeem to the authorities 0.2 credits for every ton produced using virgin pulp.[29] Thus, if the actual recycled content (ARC) for a producer was equal to the RCS, then the quantity of rights that it would be granted would exactly match its obligation. However, if it deviated from the RCS, then the producer would either be able to sell its surplus credits, or would have to purchase additional credits to make up the shortfall.

This "gross" credit approach may appear rather cumbersome compared to the "net" credit approach used in the two previous examples – where rights to credits were only generated when a producer beat its target, and obligations to purchase credits only arose when there was a shortfall. However, as Dinan notes, it has the advantage of emphasising the equivalence of the trading scheme to a combined tax / subsidy scheme; where a tax (equal to 0.2 × the credit price) is imposed on newsprint made from virgin pulp, and a subsidy (equal to 0.8 × the credit price) is paid for newsprint made from recycled pulp. It is also clear – although not recognized by Dinan – that this price-based mechanism would be revenue neutral, since the equilibrium quantity of virgin newsprint is four times the quantity produced from recycled pulp.

[29] The scaling parameters used to calculate property rights per ton, and obligations per ton, reflect the target proportion at the start of the scheme (i.e. 20%).

Dinan also considers some of the issues that might undermine the efficiency of the trading mechanism – in particular, the possibility of strategic behaviour in the market for credits. Due to the economies of scale involved in the de-inking of old newspapers, he calculates that there would be between 12 and 15 sellers of credits (i.e. producers of recycled newsprint), with around 50 buyers. In the worst-case scenario, the two largest sellers would control over 40% of the total supply of credits, and hence there would be a significant danger of strategic behaviour undermining the potential efficiency gains.

1.5 Summary

The attainment of a particular environmental objective (or set of objectives) will usually require some form of regulatory intervention, in order to change the behaviour of firms and / or individuals. An intervention has two basic components – a target and an implementation mechanism. The target may be defined in terms of an absolute limit, or in terms of a relative performance standard; it may apply to each agent individually, or to a group of agents; it may be specified explicitly, or it may be implied. Similarly, there are a wide range of mechanisms that can be used to implement the target. Essentially, these can be classified into six broad types, including (*inter alia*) market-based mechanisms and price-based mechanisms, which both work by creating financial incentives for agents to change their behaviour.

The distinction between the underlying environmental objective and the regulatory target means that the environmental effectiveness of an implementation mechanism can be assessed against either yardstick. Of course, if the target is the same as the objective, then the distinction between the two measures is redundant. However, if there is a divergence between the two, then it is possible for the mechanism to be environmentally effective in terms of the target, but not the objective.

The distinction also gives rise to several different interpretations for economic efficiency. It can relate to the gross economic cost of achieving the regulatory target – i.e. the cost-efficiency of the implementation mechanism; to the gross economic cost of achieving the environmental objective – i.e. the cost-efficiency of the regulatory intervention; or to the net economic benefit that is achieved – i.e. the welfare-efficiency of the regulatory intervention. In the first-best world (with perfect competition in all markets), these three definitions are hierarchical in the sense that each is necessary (but not sufficient) for the next. That is, the regulatory intervention cannot be welfare-efficient unless it is cost-efficient, and this in turn requires that the implementation mechanism is cost-efficient. Again, if the target and the underlying objective coincide, then the distinction between the first two measures of efficiency is redundant.

In an ideal, first-best world there are clear theoretical arguments for preferring a regulatory target that takes the form of an (explicit or implicit) absolute aggregate limit over one that takes the from of a relative aggregate performance standard. When the respective implementation mechanisms are both cost-efficient, the gross economic cost of achieving any given environmental objective is higher under the performance standard. Consequently, while it is possible to achieve the socially optimal outcome (i.e. maximize net welfare) under the aggregate limit, it is not possible to do so under the aggregate performance standard.

However, when one moves away from this rather unrealistic view of the world, to one in which there is uncertainty about the economic conditions that will apply when the regulatory intervention is implemented, regulated firms are able to act strategically in their output markets, and the government uses distortionary labour taxes to raise revenue, then the picture becomes more complex. Net welfare is higher under the absolute limit if the implementation mechanism raises revenue, provided that this is used to reduce labour taxes. However, if revenues are returned in the form of lump sum payments, or if the mechanism used to implement the absolute limit is revenue-neutral, then "potential" net welfare can be higher under the performance standard. Therefore, if revenue-neutrality is a necessary condition for the political acceptability of an intervention, it may be preferable that the target takes the form of an aggregate performance standard – provided of course that this can be implemented efficiently.

Notwithstanding the theoretical arguments, relative performance standards are used in a wide range of environmental policy areas – including waste / resource management, fuel / energy efficiency, climate change and air / water quality. This provides a strong practical argument for the development of a cost-efficient implementation mechanism, as it is likely to be easier to change the existing mechanism than to move to a completely new form of regulatory intervention based on an absolute limit.

To date the majority of relative performance standards have been implemented either through a system of fixed individual standards, or by negotiated agreements with industry associations that do not specify individual targets. However, there are a small number of examples where market-based, trading mechanisms have been used. The earliest example, and probably the most well known, was the lead credit trading programme that operated in the USA in the mid-1980s. More recently, trading mechanisms have been introduced in the United Kingdom to implement industrial energy efficiency targets, and recovery targets for waste packaging. In the Netherlands, a credit-trading scheme for specific NO_X emissions is due to start in 2004.

Thus, there is clearly an interest on the part of practitioners in using a trading mechanism to implement relative performance standards. Unfortunately, there is little that is known about the theoretical properties of such a mechanism – other than its inability to achieve the social optimum. For example:

- Can all of the many different types of relative performance standard be implemented by a trading mechanism?

- Does a trading mechanism achieve the aggregate performance standard at the lowest possible gross economic cost?
- What impact does the trading-mechanism have on the prices and quantities of commodities used and produced by the regulated agents?
- What are the distributional impacts of the mechanism, and is it possible to manipulate these impacts?
- What information does the trading mechanism provide about the cost of the regulatory intervention?
- Is there an equivalence (symmetry) between the trading mechanism and a price-based mechanism?
- What are the implications of market imperfections for the environmental effectiveness and economic efficiency of the mechanism?

These are the questions that will be addressed in the following seven chapters; starting in the next chapter with the question of existence.

Chapter 2 Overview of performance-based credit trading

This chapter introduces a market-based mechanism that can be used to implement any regulatory target that can be expressed in terms of a linear aggregate performance rule. This includes a wide range of targets in different policy areas, including the three forms of relative performance standard identified in the previous chapter. It also demonstrates that the mechanism is environmentally effective, in that it ensures that the regulatory target is achieved.

The first section of the chapter provides a generic definition for an aggregate performance rule, and then uses four specific examples to show how the rule can be adapted to different forms of regulatory target, by setting appropriate values for the various rule parameters. The second section introduces the main elements of the mechanism, and shows how the aggregate rule is translated into a series of performance rules that are applied to individual agents. It also shows how the detailed design of the mechanism can be tailored to meet other policy constraints, by varying the values of two sets of design parameters that affect the distribution of obligations and / or property rights between agents.

The environmental effectiveness of the mechanism is considered in the third section of the chapter, and it is demonstrated that the satisfaction of the individual rules ensures that the aggregate performance rule is also satisfied. Finally, the chapter concludes by considering the application of the mechanism to a hybrid regulatory target, where the target value is made up of two components – one fixed, the other varying with output.

2.1 Aggregate performance rule

Performance-based credit trading can be used to implement any regulatory target that can be expressed in terms of a linear *aggregate performance rule*. This rule defines allowable combinations of aggregate outputs and / or inputs of specified market commodities – indexed by $k \in K \subset \aleph$, and / or aggregate inputs of specified non-market environmental services[1] – indexed by $m \in M \subset \aleph$, for a specified group of economic agents – indexed by $i \in I \subset \aleph$.[2]

The aggregate performance rule takes the form of the following generic linear constraint:

$$\sum_{k \in K} \alpha_k \sum_{i \in I^{k+}} y_{ki} + \sum_{k \in K} \beta_k \sum_{i \in I^{k-}} w_{ki} + \sum_{m \in M} \chi_m \sum_{i \in I^{m-}} z_{mi} + \delta \geq 0 \qquad \ldots (2.1)$$

where:

$I^{k+} \subset I$ is the set of agents that produce market commodity $k \in K$ as an output

$I^{k-} \subset I$ is the set of agents that use market commodity $k \in K$ as an input

$I^{m-} \subset I$ is the set of agents that use non-market commodity $m \in M$ as an input

[1] Emissions of pollutants are treated as inputs of non-market (i.e. un-priced) services provided by the environment.

[2] Depending on the particular application, the agents may be firms, public authorities, or individual consumers.

$y_{ki} \in \Re^{+}$ is the quantity of market commodity $k \in$ K produced by firm $i \in \mathrm{I}^{k+}$

$w_{ki} \in \Re^{-}$ is the quantity of market commodity $k \in$ K used by firm $i \in \mathrm{I}^{k-}$

$z_{mi} \in \Re^{-}$ is the quantity of non-market commodity $j \in$ J used by firm $i \in \mathrm{I}^{j-}$

and $\alpha_k \in \Re$, $\beta_k \in \Re$, $\chi_m \in \Re^{+}$, and $\delta \in \Re^{+}$ are exogenous parameters

There are no restrictions on the number of elements in each commodity set. Indeed, as the examples outlined below illustrate, in many applications at least one of the commodity sets will be empty – i.e. the performance rule will include only market commodities, or only non-market commodities. There are however two restrictions on the values of the rule parameters. First, the parameters for the non-market commodities, and the constant term, must all be non-negative. Second, for any market commodity $k \in$ K, either $\alpha_k = 0$ or $\beta_k = 0$. That is, the performance rule cannot include both the outputs and the inputs of a particular commodity.

Notwithstanding these restrictions, the general formulation of the performance rule is very flexible. By choosing appropriate definitions for the commodity sets K and / or M, and appropriate values for the parameter vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\chi}$, and the scalar δ, it is possible to represent a variety of different forms of regulatory target with a wide range of potential policy applications – including the three types of relative performance standard identified in Chapter 1. While it is the generic rule's ability to represent a range of relative standards that is of particular interest, it can also incorporate regulatory targets that take the form of an absolute limit – as the first of the following examples illustrates.[3]

[3] The second example – industrial energy efficiency, and an application to waste recovery targets under extended producer responsibility, are explored in more detail in chapters 5 and 6 respectively.

Example 1 Greenhouse gas emissions limit

Commodity sets: $K = \varnothing$

$M = \{$ six greenhouse gases $\}$

Parameter values: $\chi_m = c_m$ for $m \in M$, where c_m is the carbon dioxide equivalence factor

$\delta = L$ where L is the aggregate absolute limit in tonnes of carbon dioxide

Performance rule: $$\sum_{m \in M} c_m \sum_{i \in I^{m-}} z_{mi} + L \geq 0$$

Regulatory target: $$\sum_{m \in M} c_m \sum_{i \in I^{m-}} (-z_{mi}) \leq L$$

That is, the aggregate emissions of a basket of six greenhouse gases (measured as tonnes of carbon dioxide equivalent) should not exceed L tonnes.

Example 2 Industrial energy efficiency standard for the chemicals sector

Commodity sets: $K = K^E \cup K^C$ where $K^E = \{$ energy products $\}$

$K^C = \{$ chemical products $\}$

$M = \varnothing$

Parameter values: $\alpha_k = 0$ for $k \in K^E$

$= r\, a_k$ for $k \in K^C$, where a_k is the unit weight in tonnes, and $r > 0$ is the target rate for specific energy consumption (SEC) in joules per tonne

$\beta_k = b_k$ for $k \in K^E$, where b_k is unit energy content in joules

$= 0$ for $k \in K^C$

$\delta = 0$

Performance rule: $$\sum_{k \in K^C} r\, a_k \sum_{i \in I^{k+}} y_{ki} \;+\; \sum_{k \in K^E} b_k \sum_{i \in I^{k-}} w_{ki} \;\geq\; 0$$

Regulatory target: $$\frac{\sum_{k \in K^E} b_k \sum_{i \in I^{k-}} (-w_{ki})}{\sum_{k \in K^C} a_k \sum_{i \in I^{k+}} y_{ki}} \;\leq\; r$$

That is, the aggregate specific energy consumption rate of firms in the chemicals sector should not exceed r joules per tonne.

Example 3 *Recycled content of newsprint standard*

Commodity sets: K = { recycled content newsprint, virgin content newsprint }

$M = \varnothing$

Parameter values: $\alpha_k = 0$ for $k \in K^1$

$\beta_1 = r - 1$ where $0 < r < 1$ is the relative standard for the minimum proportion of recycled content newsprint

$\beta_2 = r$

$\delta = 0$

Performance rule: $$(r-1) \sum_{i \in I^{1-}} w_{1i} \;+\; r \sum_{i \in I^{2-}} w_{2i} \;\geq\; 0$$

Regulatory target: $$\frac{\sum_{i \in I^{1-}} (-w_{1i})}{\sum_{i \in I^{1-}} (-w_{1i}) \;+\; \sum_{i \in I^{2-}} (-w_{2i})} \;\geq\; r$$

That is, the proportion of aggregate newsprint coming from recycled sources should exceed r percent.

Example 4 *Vehicle Fuel Efficiency standard*

Commodity sets: $K = \{ \text{vehicle models} \}$

$M = \varnothing$

Parameter values: $\alpha_k = r - r_k$ for $k \in K$, where r is the relative standard for the minimum average vehicle fuel efficiency (in km per litre), and r_k is the actual fuel efficiency of the vehicle model.

$\beta_k = 0$ for $k \in K$

$\delta = 0$

Performance rule: $$\sum_{k \in K} (r - r_k) \sum_{i \in I^{k+}} y_{ki} \geq 0$$

Regulatory target: $$\frac{\sum_{k \in K} r_k \sum_{i \in I^{k+}} y_{ki}}{\sum_{k \in K} \sum_{i \in I^{k+}} y_{ki}} \geq r$$

That is, the weighted average fuel efficiency of all vehicles produced should not be less than r kilometres per litre.

It is clear from these four examples that the performance rule parameters represent a combination of exogenous, commodity-specific *scaling parameters* (i.e. c_m, a_k, b_k and r_k), and *regulatory target values* (i.e. L and r). Since the performance rule will often combine variables that have different units of measurement, the scaling parameters play an important role in standardising the units of measurement. In the case of an absolute limit, the denomination of the performance rule is the same as that of the regulatory target. In the case of a relative standard, it is the same as that of the numerator of the target. So for example, in the third application – the relative standard for the recycled content of newsprint, the performance rule is denominated in tonnes of newsprint.

2.2 The trading system

Performance-based credit trading operates by introducing a new market commodity – *performance credits*, and then linking this commodity to the market and / or non-market commodities included in the aggregate performance rule. These links are provided by a series of *individual performance rules*, which are defined for each agent $i \in \mathrm{I}$:

$$\sum_{k \in \mathrm{K}^{i+}} \rho_k y_{ki} \quad + \quad \sum_{k \in \mathrm{K}^{i-}} \sigma_k w_{ki} \quad + \quad \sum_{m \in \mathrm{M}^{i-}} \chi_m z_{mi} \quad + \quad \delta_i \quad \geq \quad v_i \qquad \ldots (2.2)$$

where

$\mathrm{K}^{i+} \subset \mathrm{K}$ is the subset of regulated market commodities produced by agent $i \in \mathrm{I}$ as an output

$\mathrm{K}^{i-} \subset \mathrm{K}$ is the subset of regulated market commodities used by agent $i \in \mathrm{I}$ as an input

$\mathrm{M}^{i-} \subset \mathrm{M}$ is the subset of regulated non-market commodities used by agent $i \in \mathrm{I}$ as an input

$v_i \in \Re$ is the quantity of performance credits sold / purchased by agent $i \in \mathrm{I}$

and $\rho_k \in \Re$, $\sigma_k \in \Re$, $\chi_m \in \Re^+$ and $\delta_i \in \Re$ are exogenous parameters

The value of the left hand side of the rule may be positive or negative, depending on the values of the various parameters and the agent's input and output quantities. If it is positive, then the rule defines the maximum net quantity of performance credits that the agent is allowed to sell. Alternatively, if it is negative, then it defines the minimum net quantity that it must purchase.[4] The sum of the positive quantities represents the

[4] There is no restriction on the gross sales and purchases of performance credits.

aggregate supply of performance credits, while the sum of negative quantities represents the aggregate demand.

If the left hand side of (2.2) is identically equal to zero, then the rule reduces to a prohibition on the net sale of performance credits. While the agent is allowed to purchase credits, it will not wish to do so if the price of credits is positive, since they are of no value (other than for resale). In this case, the agent is said to be an *inactive* participant in the trading scheme. It should be noted that this is different to the case where the left hand side happens to be equal to zero for the equilibrium values of the inputs and outputs of an *active* agent.

All transfers of performance credits between agents must be registered with the authorities, and at the end of the period each agent must submit a (verified) return, providing information on the quantities of the regulated commodities that it has used or produced. These returns can then be reconciled with the information held on the register regarding the transfers of performance credits, to ensure that each agent has satisfied its individual rule.

From the definition of the aggregate performance rule, it follows that the set I includes all those agents that produce or use any of the market or non-market commodities included in the rule. However, it is also possible for the set to include some agents that do not produce or use any of these commodities. Respectively, these two subsets are termed the *regulated sector* (I^R) and the *auxiliary sector* (I^A), with:

$$I^R \equiv \{ i \in I \mid M^{i-} \neq \varnothing \text{ or } K^{i-} \neq \varnothing \text{ or } K^{i+} \neq \varnothing \}$$

$$I^A \equiv \{ i \in I \mid M^{i-} = \varnothing \text{ and } K^{i-} = \varnothing \text{ and } K^{i+} = \varnothing \}$$

The inclusion of the auxiliary sector provides additional flexibility in varying the distributional impact of the regulatory intervention. In particular, it allows agents that suffer economic costs, or enjoy economic benefits as a result of the introduction of the performance rule, even though they do not produce or use any of the commodities included in the rule, to be included in the trading scheme.[5] It also allows the regulator itself to participate in the scheme.

While the individual rules have the same basic structure as the aggregate performance rule, there are a number of important differences. First, as has already been noted, the individual rules contain an additional variable – the quantity of performance credits purchased or sold by the agent. Second, the values of the parameters for the market commodities are not the same as those in the aggregate rule. However, there is a fixed relationship between the two sets of parameters, which is defined by the following three sets of identities:

$$\rho_k \equiv (1 - \theta_k)\,\alpha_k - \theta_k\,\beta_k \qquad \text{for all } k \in \mathrm{K}$$

$$\sigma_k \equiv (1 - \theta_k)\,\beta_k - \theta_k\,\alpha_k \qquad \text{for all } k \in \mathrm{K}$$

$$\delta_i \equiv \delta / \mathrm{I} + \varepsilon_i \qquad \text{for all } i \in \mathrm{I}$$

where $\theta_k \in [0,1]$. Thus, whereas in the aggregate performance rule it is not possible for both α_k and β_k to be non-zero for any market commodity $k \in \mathrm{K}$, it is possible for both ρ_k and σ_k to be non-zero in the individual rules. However, since for any particular agent $i \in \mathrm{I}$, the commodity cannot be both an output and an input, only one of the two parameters (or neither) will appear in the individual performance rule for that agent.

[5] This could be because they produce commodities that are used by the regulated agents, or use commodities that are produced by them. Alternatively, it could be because they produce or use commodities that are close substitutes or compliments to the regulated commodities.

The constant term δ_i is made up of two parts. The first part is the same for all agents, being equal to the constant in the aggregate rule divided by the total number of agents in the trading system. The second part is agent-specific, and it allows the individual performance rules of the agents to be differentiated. There is no restriction on the values of these individual parameters. However, in aggregate, they are required to sum to zero, i.e.

$$\sum_{i \in I} \varepsilon_i = 0 \qquad \text{... (2.3)}$$

This, in turn, implies that the sum of the constant terms in the individual rules is equal to the constant term in the aggregate performance rule, i.e.

$$\sum_{i \in I} \delta_i = \delta \qquad \text{... (2.4)}$$

Apart from this requirement, there are no restrictions on the values of the individual constant terms. In particular, they may be positive or negative, unlike the constant term in the aggregate rule.

The parameters $\boldsymbol{\theta} = (\theta_1, \dots, \theta_K)$ and $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_I)$ can be thought of as *regulatory design variables*, that allow the regulator to vary the detailed design of the trading scheme in order to meet the other policy constraints. Interpretation of these parameters is facilitated by reformulating the individual performance rule in terms of obligations and property rights. This also highlights the intrinsic nature of performance credits – i.e. the benefit that they bestow on the holder.

Expanding and re-arranging the individual performance rule (2.2) for agent $i \in I$ gives the following inequality:

$$
\begin{aligned}
&\left[\sum_{k \in K_{+}^{i+}} (1-\theta_k)\, \alpha_k\, y_{ki} + \sum_{k \in K_{+}^{i-}} \theta_k\, \alpha_k\, (-w_{ki}) \right] + \\
&\left[\sum_{k \in K_{-}^{i-}} (1-\theta_k)\, \beta_k\, w_{ki} + \sum_{k \in K_{-}^{i+}} \theta_k\, \beta_k\, (-y_{ki}) \right] + \\
&\left[\frac{\delta}{I} \right] + \left[-v_i \right] + \left[\varepsilon_i \right] \\
&\qquad \geq \left[\sum_{k \in K_{+}^{i-}} (1-\theta_k)\, \beta_k\, (-w_{ki}) + \sum_{k \in K_{+}^{i+}} \theta_k\, \beta_k\, y_{ki} \right] + \\
&\qquad\quad \left[\sum_{k \in K_{-}^{i+}} (1-\theta_k)\, \alpha_k\, (-y_{ki}) + \sum_{k \in K_{-}^{i-}} \theta_k\, \alpha_k\, w_{ki} \right] + \\
&\qquad\quad \left[\sum_{m \in M^{i-}} \chi_m\, (-z_{mi}) \right] \qquad \ldots (2.5)
\end{aligned}
$$

where

$K_{+} \subset K$ is the subset of regulated market commodities that have a positive performance rule parameter (either $\alpha_k > 0$ or $\beta_k > 0$)

$K_{-} \subset K$ is the subset of regulated market commodities that have a negative performance rule parameter (i.e. either $\alpha_k < 0$ or $\beta_k < 0$)

By construction, all of the summation terms inside the various square brackets are non-negative. Consequently, both sides of the inequality are non-negative.[6] The right hand side represents the total *obligation* of agent $i \in I$ under the performance rule. This obligation is discharged by redeeming performance credits to the authorities at the end

[6] This was not the case under the original formulation of the individual performance rule (2.2).

of the period. The left hand side represents the quantity of *property rights* for performance credits held by the agent (i.e. the number to which it has legal title). Thus, the reformulated performance rule requires that agents must hold sufficient property rights to meet their respective obligations. Put another way, a performance credit entitles the holder to discharge one unit of its obligation.

An obligation arises whenever a transaction occurs that involves a commodity that has a positive input parameter in the aggregate performance rule (i.e. $\beta_k > 0$ or $\chi_m > 0$), or a market commodity that has a negative output parameter (i.e. $\alpha_k < 0$).[7] The size of the resultant obligation is equal to the magnitude of the transaction (i.e. the number of units involved) multiplied by the magnitude of parameter, and it is shared between the two parties to the transaction according to the value of θ_k.

The number of property rights held by an agent is determined by four factors: (1) its share of any endogenous property rights that are generated; (2) its (equal) share of any exogenous property rights that are allocated under the aggregate performance rule; (3) any transfers of rights to or from other agents; and (4) the value of the parameter ε_i.

An endogenous property right is generated whenever a transaction occurs that involves a market commodity that has a positive output parameter in the aggregate performance rule (i.e. $\alpha_k > 0$), or a negative input parameter (i.e. $\beta_k < 0$). As with the obligations, the number of rights generated depends on number of units involved in the transaction and the value of the parameter, and they are shared between the two parties according to the value of θ_k. The final factor determining an agent's holding of property rights is the value the parameter ε_i. Increasing the value of this parameter for a particular agent increases its holding of property rights. However, since the values of these parameters

[7] Of course, in the case of a non-market commodity, the transaction is one-sided. That is, there is only a "purchaser".

must sum to zero, this must be offset by a reduction in the value of the parameter for another agent (or agents), with the opposite effect.

Thus, the two design parameters operate in different ways. The value of the vector θ determines the initial *assignment* of any endogenous property rights and / or obligations that arise from transactions involving regulated market commodities (see Table 2.1). In contrast, the value of vector ε determines the *reallocation* of property rights between agents.

Table 2.1 Interpretation of assignment parameter θ_k

	$\alpha_k < 0$	$\alpha_k = 0$	$\alpha_k > 0$
$\beta_k < 0$		Share of property right assigned to seller	
$\beta_k = 0$	Share of obligation assigned to purchaser		Share of property right assigned to purchaser
$\beta_k > 0$		Share of obligation assigned to seller	

While the reformulated individual performance rule (2.5) may appear rather complex, in most practical applications the values of the parameters mean that it reduces down to a much simpler expression, as the following examples illustrate. These relate to the first two policy applications highlighted in section 2.1.

Example 1 Greenhouse gas emissions limit

In this application there are no market commodities. Consequently, the assignment parameters θ are redundant, and it is only the performance adjustment factors (ε) that can affect the individual performance rules. Consider the case where:

$$\varepsilon_i = \gamma_i L - \frac{L}{I} \quad \text{for all } i \in I \qquad \text{with} \qquad \sum_{i \in I^{m-}} \gamma_i = 1$$

which satisfies requirement (2.3). The resultant individual performance rules are:

$$\gamma_i L + [-v_i] \geq \sum_{m \in M^{i-}} c_m(-z_{mi}) \qquad \text{for } i \in I^R$$

$$\gamma_i L + [-v_i] \geq 0 \qquad \text{for } i \in I^A$$

Since only non-market environmental services enter into the aggregate performance rule, the entire obligation is borne by the regulated agent. There are no endogenous property rights, and the exogenous property rights are shared between the regulated and auxiliary agents, according to the values of the parameters γ_i. If $\gamma_i = 0$ for all $i \in I^R$, then all of the rights are assigned to the auxiliary sector. In contrast, if $\gamma_i = 0$ for all $i \in I^A$, then all of the rights are assigned to the regulated sector. If the auxiliary sector contains only one agent – the regulator, then the first case is equivalent to the auctioning of permits in a "cap & trade" system, while the second case is equivalent to the free allocation of permits.

Example 2: Industrial energy efficiency standard for chemicals sector

In this application $\delta = 0$. Assuming that the performance adjustment factors are all set equal to zero (i.e. $\boldsymbol{\varepsilon} = \mathbf{0}$), then the individual performance rules depend only on the values of the assignment parameter vector $\boldsymbol{\theta}$. For simplicity, it is assumed that the regulated sector is divided into three mutually exclusive sub-sectors: chemical producers (I^C), energy producers (I^E), and downstream chemical users (I^D), where:[8]

[8] For simplicity it is assumed that chemical producers do not produce any energy, and do not use any chemical products as inputs. In reality, neither of these assumptions is valid.

$I^C = \quad \{ i \in I \mid K^{i-} \subset K^1 \text{ and } K^{i+} \subset K^2 \}$

$I^E = \quad \{ i \in I \mid K^{i-} = \varnothing \text{ and } K^{i+} \subset K^1 \}$

$I^D = \quad \{ i \in I \mid K^{i-} \subset K^2 \text{ and } K^{i+} = \varnothing \}$

Three alternative cases are considered. In the first case $\theta_k = 0$ for all $k \in K^1$, and $\theta_k = 0$ for all $k \in K^2$, and the resultant performance rules are:

$$r \sum_{k \in K^{i+}} a_k y_{ki} + [-v_i] \quad \geq \quad \sum_{k \in K^{i-}} e_k(-w_{ki}) \qquad \text{for all } i \in I^C$$

$$[-v_i] \quad \geq \quad 0 \qquad \text{for all } i \in I^E \cup I^D \cup I^A$$

All of the obligations are assigned to the chemical producers; as are all of the endogenously generated property rights. Consequently, the rules for the agents in the other three sectors reduce to a prohibition on the net sales of performance credits.

In the second case, $\theta_k = 1$ for all $k \in K^1$, and $\theta_k = 1$ for all $k \in K^2$, and the resultant performance rules are:

$$[-v_i] \quad \geq \quad \sum_{k \in K^{i+}} e_k y_{ki} \qquad \text{for all } i \in I^E$$

$$r \sum_{k \in K^{i-}} m_k(-w_{ki}) + [-v_i] \quad \geq \quad 0 \qquad \text{for all } i \in I^D$$

$$[-v_i] \quad \geq \quad 0 \qquad \text{for all } i \in I^C \cup I^A$$

In this case, all of the obligations are assigned to the producers of the energy inputs, and all of the endogenous property rights are assigned to the users of the chemical products.

The rules for the chemical producers reduce to a prohibition on net sale of performance credits, and hence they are now inactive.

In the third case, $\theta_k = ½$ for all $k \in K^1$, and $\theta_k = ⅓$ for all $k \in K^2$, and the resultant performance rules are:

$$⅔ \; r \sum_{k \in K^{i+}} a_k y_{ki} \;+\; [-v_i] \quad \geq \quad ½ \sum_{k \in K^{i-}} e_k(-w_{ki}) \qquad \text{for all } i \in I^C$$

$$[-v_i] \quad \geq \quad ½ \sum_{k \in K^{i+}} e_k y_{ki} \qquad \text{for all } i \in I^E$$

$$⅓ \; r \sum_{k \in K^{i-}} a_k(-w_{ki}) \;+\; [-v_i] \quad \geq \quad 0 \qquad \text{for all } i \in I^D$$

$$[-v_i] \quad \geq \quad 0 \qquad \text{for all } i \in I^A$$

In this case, the obligations are shared between the chemical producers and the energy producers; with two thirds of the total being assigned to the former, and one third to the latter. Similarly, the endogenous property rights are shared equally between the chemical producers and the chemical users. Hence, agents in all three sectors are active participants in the market for performance credits.

2.3 Environmental effectiveness of the trading system

With so much flexibility in the design of the individual performance rules, it may not be clear that they are consistent with the aggregate performance rule. That is, does the satisfaction of the individual trading rules ensure that the aggregate performance rule is satisfied, and hence that the underlying regulatory target is achieved. Fortunately, it is straightforward to show that this is indeed the case, provided that all of the commodity markets clear at positive prices.[9]

If all of the individual performance rules (2.2) are satisfied, then summing over all agents gives:

$$\sum_{i\in I} v_i \quad \leq \quad \sum_{i\in I}\left[\ \sum_{k\in K^{i+}} \rho_k y_{ki} + \sum_{k\in K^{i-}} \sigma_k w_{ki} + \sum_{m\in M^{i-}} \chi_m z_{mi} + \delta_i\right]$$

$$= \quad \sum_{k\in K}\sum_{i\in I^{k+}} \left[(1-\theta_k)\alpha_k - \theta_k\beta_k\right] y_{ki} + \sum_{k\in K}\sum_{i\in I^{k-}} \left[(1-\theta_k)\beta_k - \theta_k\alpha_k\right] w_{ki}$$

$$+ \quad \sum_{m\in M}\sum_{i\in I^{m-}} \chi_m z_{mi} + \delta$$

$$= \quad \sum_{k\in K}\alpha_k \sum_{i\in I^{k+}} y_{ki} + \sum_{k\in K}\beta_k \sum_{i\in I^{k-}} w_{ki} + \sum_{m\in M}\chi_m \sum_{i\in I^{m-}} z_{mi} + \delta$$

$$- \quad \sum_{k\in K}\theta_k(\alpha_k+\beta_k)\left[\sum_{i\in I^{k+}} y_{ki} + \sum_{i\in I^{k-}} w_{ki}\right]$$

Noting that the market clearing conditions for market commodities, and for performance credits are respectively:

$$\sum_{i\in I^{k+}} y_{ki} + \sum_{i\in I^{k-}} w_{ki} \geq 0 \qquad p_k\left[\sum_{i\in I^{k+}} y_{ki} + \sum_{i\in I^{k-}} w_{ki}\right] = 0 \qquad \text{for all } k \in K$$

$$\sum_{i\in I} v_i \geq 0 \qquad\qquad q\left[\sum_{i\in I} v_i\right] = 0$$

where p_k is the price of market commodity $k\in K$, and q is the price of performance credits, then it follows that a sufficient condition for the aggregate performance rule

[9] It is shown in Chapter 3 that, under a number of mild assumptions, this will always be the case.

(2.1) to be satisfied is that $p_k > 0$ for all commodities $k \in K$. Thus, performance based credit trading is environmentally effective under the "narrow" definition identified in Chapter 1. That is, it achieves the regulatory target.

2.4 Hybrid regulatory targets

In the four examples highlighted in section 2.1, the aggregate performance rule represented either an absolute limit, or a relative standard. However, this polarisation is not necessary, and it is possible to have *hybrid regulatory targets*, that combine the two forms. For example, one could amend the first example to have

Commodity sets:	K	=	{ products }
	M	=	{ six greenhouse gases }
Parameter values:	α_k	=	$r\, a_k$
	χ_m	=	c_m
	δ	=	L

Performance rule: $$\sum_{k \in K} r\, a_k \sum_{i \in I^{k+}} y_{ki} \; + \; \sum_{m \in M} c_m \sum_{i \in I^{m-}} z_{mi} \; + \; L \; \geq \; 0$$

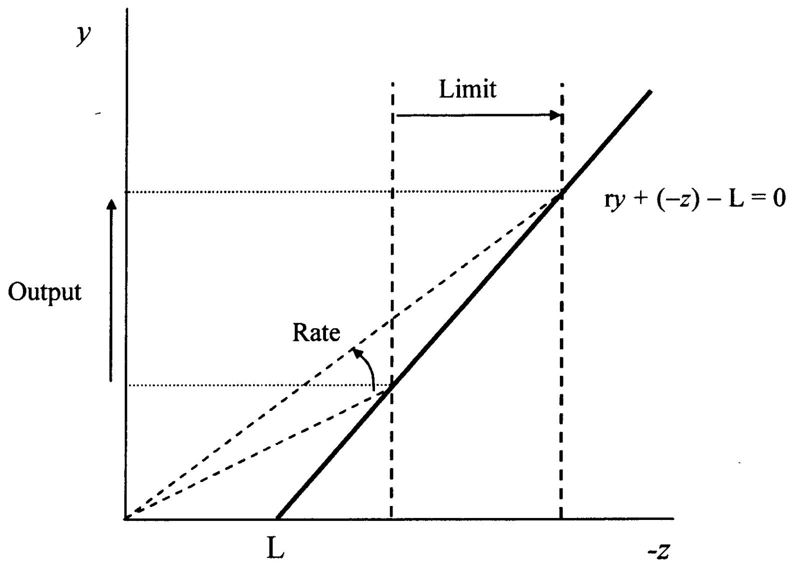
Regulatory target: $$\frac{\sum_{m \in M} c_m \sum_{i \in I^{m-}} (-z_{mi})}{\sum_{k \in K} a_k \sum_{i \in I^{k+}} y_{ki}} \;\leq\; r \;+\; \frac{L}{\sum_{k \in K} a_k \sum_{i \in I^{k+}} y_{ki}} \qquad \text{or}$$

$$\sum_{m \in M} c_m \sum_{i \in I^{m-}} (-z_{mi}) \;\leq\; L \;+\; r \sum_{k \in K} a_k \sum_{i \in I^{k+}} y_{ki}$$

In this case the performance rule can be interpreted either in terms of a relative standard, or in terms of an absolute limit. However, under either interpretation the target value varies with the level of output (i.e. the target is "soft" or "flexible"). As can be seen in

Figure 2.1, under the first interpretation, the rate reduces as output increases; while under the second, the limit increases.

Figure 2.1 A hybrid regulatory target



At first sight this may not seem a very plausible formulation for a regulatory target (under either interpretation). However, if one assumes that while all of the regulated agents emit some greenhouse gases, the products are only produced by a subset of the agents, i.e.

$$\mathrm{I}^1 = \{ i \in \mathrm{I} \mid \mathrm{M}^{i-} \neq \varnothing \text{ and } \mathrm{K}^{i+} = \varnothing \}$$

$$\mathrm{I}^2 = \{ i \in \mathrm{I} \mid \mathrm{M}^{i-} \neq \varnothing \text{ and } \mathrm{K}^{i+} \neq \varnothing \}$$

and that the values of the design parameters are:

$$\theta_k = 0 \qquad \text{for all } k \in \mathrm{K}$$

$$\varepsilon_i = \gamma_i \mathrm{L} - \frac{\mathrm{L}}{\mathrm{I}} \qquad \text{for all } i \in \mathrm{I}, \text{ with } \sum_{i \in \mathrm{I}} \gamma_i = 1, \text{ and } \gamma_i = 0 \text{ for all } i \in \mathrm{I}^2 \cup \mathrm{I}^\mathrm{A}$$

then the individual performance rules (2.2) reduce to:

$$v_i \quad \leq \quad \sum_{m \in M} c_m z_{mi} \; + \; \gamma_i L \qquad \text{for all } i \in I^1$$

$$v_i \quad \leq \quad r \sum_{k \in K^{i+}} a_k y_{ki} \; + \; \sum_{k \in K^{i-}} c_m z_{mi} \qquad \text{for all } i \in I^2$$

$$v_i \quad \leq \quad 0 \qquad \text{for all } i \in I^A$$

Thus, in the absence of any trading (i.e. $v_i = 0$ for all $i \in I$), the agents in sector I^1 are subject to absolute limits ($\gamma_i L$) on their emissions, while those in sector I^2 are subject to a common performance standard (r) for their specific emission rates. When viewed from this perspective, the hybrid target looks more familiar. In particular, the United Kingdom Emissions Trading Scheme (UKETS) for greenhouse gases, partitions regulated agents into absolute and relative sectors in this way.[10] While a hybrid regulatory target has not been specified explicitly for the UKETS, it is implied by the design of the implementation mechanism.

2.4 Summary

This chapter has described a market-based mechanism that can be used to implement any regulatory target that can be expressed in terms of a linear aggregate performance rule. This includes a wide range of targets in different policy areas. In particular, it includes the three forms of relative performance standard identified in Chapter 1 – i.e. proportions, weighted averages, and rates.

---

[10] The scheme – which is described in Chapter 5 – is more complex than the simple two-sector system represented here. In particular, the performance standards relate to specific energy consumption, and there are restrictions on the sales of credits between the two sectors. Nonetheless, the general point remains valid.

The mechanism operates by introducing a new commodity – performance credits – and then linking this to the commodities included in the aggregate performance rule. These links are provided by a series of individual performance rules, which determine the maximum number of credits that an agent is allowed to sell, or the minimum number that it must purchase. The detailed design of the scheme can be tailored to meet other policy constraints, or to ameliorate the effect of market distortions, by varying the values of two sets of design parameters that affect the distribution of obligations and / or property rights between agents.[11]

It has been demonstrated that satisfaction of the individual performance rules ensures that the aggregate rule is also satisfied, irrespective of the values that are chosen for the design variables. Hence, the mechanism is environmentally effective, under the "narrow" definition identified in Chapter 1 – i.e. that the regulatory target is achieved.

Of course, so far no consideration has been given to the financial impacts of the mechanism. In particular, the economic efficiency of performance-based credit trading – in terms of achieving the regulatory target at lowest gross economic cost – has yet to be established. Also, while it has been shown how the distribution of obligations and property rights can be varied by the use of the design variables, it has not been demonstrated that this necessarily translates into a redistribution of the financial impacts. For example, if the redistribution affects the market prices of performance credits or market commodities, then there may be no correlation between the two. Both of these important aspects are addressed in the formal analysis of the mechanism that is under taken in Chapter 3.

[11] The implications of market power for the design of a performance-based trading scheme are addressed in Part 3 of the thesis.

Part Two

Performance-based credit trading under perfect competition

Chapter 3 Formal analysis of performance-based credit trading

The last chapter provided an overview of performance-based credit trading, and showed how it can be used to implement a wide range of regulatory targets, including the three main forms of relative standard – proportions, weighted averages, and rates. In this chapter, the cost efficiency, and the distributional flexibility, of the mechanism are analysed formally.

The chapter is divided into three sections. The first section sets out the model that underpins the analysis, and provides definitions of the various variables, vectors and sets that are used in the notation. It also explicitly identifies the assumptions that have been made about the structure of the economic system; the technical properties of production and emission functions; and the performance rule parameters.

There are a wide range of models that could be adopted for the analysis. It is useful to think of these as lying on a (multi-dimensional) spectrum. At one end of the spectrum are atomistic, dynamic, general equilibrium models with relatively general functional forms, which allow free entry and exit in individual markets (on either side). At the other end are aggregate, static, partial equilibrium models with specific functional forms, which assume fixed numbers of agents in each individual market. The choice of appropriate location on this spectrum will depend on the underlying purpose of the analysis. For a generic analysis of the theoretical properties of an implementation mechanism, one would like the model to be as abstract as possible, with restrictions only imposed where they are needed to ensure tractability. In contrast, for a specific analysis of the mechanism in a particular policy application, one would like the model to reflect the salient characteristics of the application. In the first case there are strong arguments for using an atomistic, general equilibrium model; while in the second, an aggregate, partial equilibrium model is likely to be more appropriate.[1]

The model adopted here can be viewed as a compromise between these two extreme positions. It is an atomistic, general equilibrium model, with relatively few restrictions on the characteristics of production technologies. However, it is static and short-run. An interesting feature of the model is that, by choosing appropriate values for certain exogenous parameters, it can mimic a partial equilibrium model for a group of related commodities. This allows the general theoretical properties that are derived in this chapter (and the next) to be transferred directly into the partial equilibrium framework that is adopted for the specific applications that are considered in chapters 5 and 6.

The second section characterises the regulated cost minimum – i.e. the allocation that minimizes the aggregate gross economic cost of achieving the regulatory target embodied in the aggregate performance rule. The necessary and sufficient conditions for the cost minimum are derived and discussed. The section concludes by considering the interpretation of the shadow value of the aggregate performance rule in the context of a relative standard for specific emissions, and it is shown that it is not in general equal to the marginal cost of (absolute) aggregate emissions reduction.

The third section characterises the market equilibrium following the introduction of performance-based credit trading. The necessary and sufficient conditions for the

[1] Although it should be noted that Montgomery (1972) adopted a partial equilibrium model for his generic analysis of ambient permit trading with absolute limits.

market equilibrium are derived, and the cost efficiency of the equilibrium is established; as are the impacts of the two types of regulatory design parameter on the prices of market commodities and performance credits, and on the distribution of the aggregate cost. The interpretation of the market price of performance credits is considered – again in the context of a relative standard for specific emissions, and it is shown that the mechanism does not in general lead to the equalization of the marginal cost of (absolute) emissions reduction across agents. Finally, the section concludes by demonstrating the equivalence of performance-based credit trading to a price-based mechanism comprising a system of commodity taxes and subsidies.

3.1 Analytical framework

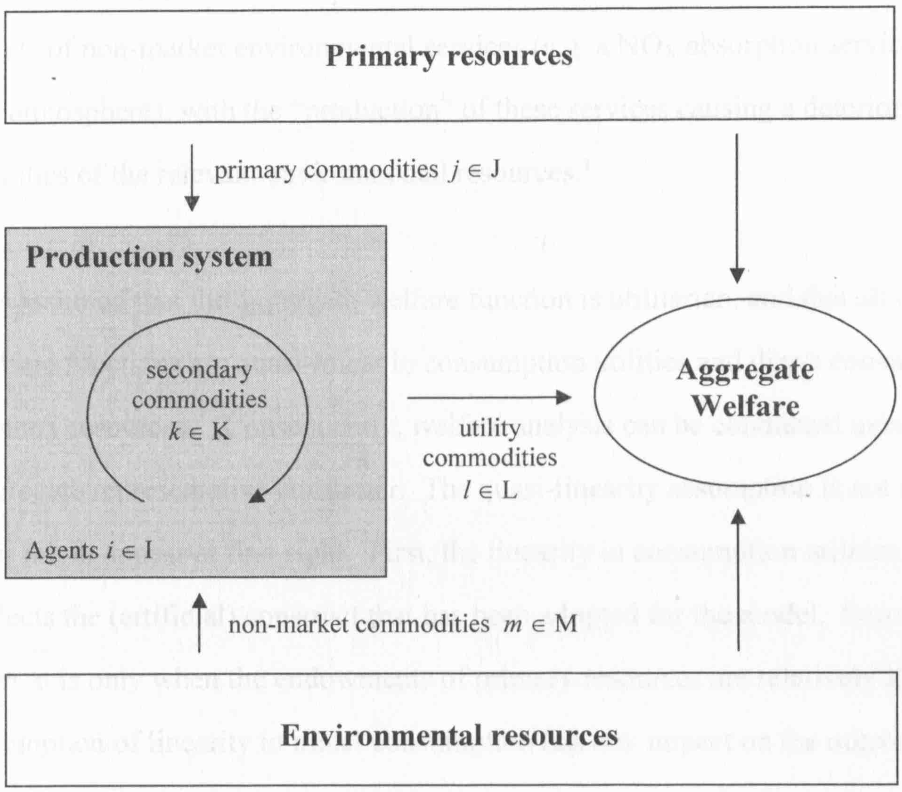
Figure 3.1 provides a schematic representation of the general equilibrium model that underpins the formal analysis presented in this chapter. In this short-run, static model, endowments of primary resources (e.g. person hours, etc.) and of environmental resources (e.g. air quality, water quality, etc.) are fixed, and these may either be used to produce commodities (i.e. goods and services), or may be "consumed" directly to generate welfare.

The production system comprises a fixed number of agents (indexed by $i \in \mathrm{I} \subset \aleph$). These may be individuals, households, firms, or other institutions (such as local authorities). The agents produce and / or use four types of commodities:

- *primary* commodities, indexed by $j \in \mathrm{J} \subset \aleph$, which have fixed endowments and cannot be produced;
- *secondary* commodities, indexed by $k \in \mathrm{K} \subset \aleph$, which are produced from primary and other secondary commodities, and used within the economic system;

- *utility* commodities, indexed by $l \in L \subset \aleph$, which are produced from primary and secondary commodities;
- *non-market* commodities, indexed by $m \in M \subset \aleph$, which are produced from environmental resources.

Figure 3.1 General equilibrium model



The model is slightly unusual in two respects. First, consumption is treated as a production activity – undertaken by individuals, that uses primary and secondary commodities as inputs to "produce" utility commodities (or consumption utilities), which enter into the individuals' welfare functions, along with their direct consumption of primary resources and environmental resources. Thus, a distinction is made between an individual's use of a primary resource in the production of utility, and their direct (or pure) consumption of the resource. While the benefit derived from the direct consumption of time is obvious (e.g. from time spent sleeping, or meditating), the benefit that is derived from the direct consumption of other types of primary resource is less clear. However, if one interprets direct consumption as preservation of (part of) the resource, then the benefit represents a form of "existence value".

Second, the emissions of pollutants arising from the production of secondary commodities (e.g. NO_X emissions caused by the generation of electricity) are treated as inputs of non-market environmental services (e.g. a NO_X absorption service provided by the atmosphere), with the "production" of these services causing a deterioration in the qualities of the relevant environmental resources.[2]

It is assumed that the aggregate welfare function is utilitarian, and that all individual welfare functions are quasi-linear in consumption utilities and direct consumption of primary resources.[3] Consequently, welfare analysis can be conducted using an aggregate representative consumer. The quasi-linearity assumption is not as restrictive as it might appear at first sight. First, the linearity in consumption utilities merely reflects the (artificial) construct that has been adopted for the model. Second, as will be seen, it is only when the endowments of primary resources are relatively large that the assumption of linearity in direct consumption has any impact on the outcome. If the endowments are relatively small, then the resources are used entirely for production, and consequently the assumption becomes redundant.

---

[2] For example, if the atmospheric concentration of a particular pollutant is given by $A = \mathrm{A}^{\mathrm{N}} + e / \delta$, where A^{N} is the natural concentration, e is the emissions rate, and δ is the natural removal rate; and if air quality is measured by $Q = \mathrm{A}^0 - A$, where A^0 is some arbitrary reference level; then the implied production function for the atmospheric absorption of emissions is $e = [\mathrm{Q}^{\mathrm{N}} - Q] \times \delta$.

[3] That is, $W = \boldsymbol{\phi}'\boldsymbol{B} + \boldsymbol{\gamma}'\boldsymbol{A} + \mathrm{E}(\boldsymbol{Q})$, where $\boldsymbol{B}$ is a vector of consumption utilities; $\boldsymbol{A}$ is a vector of the direct consumptions of the primary resources; $\boldsymbol{Q}$ is a vector of environmental resource qualities; and ϕ and γ are vectors of exogenous parameters.

Table 3.1 defines the notation that is used to represent the various subsets of potential input and output commodities for individual agents, and the subsets of agents that potentially use or produce individual commodities.[4]

Table 3.1 Definitions of commodity and agent subsets

Subset	**Definition**
$J^{i-} \subset J$	primary commodities used as inputs by agent $i \in I$
$J^{I-} = \bigcup_{i \in I} J^{i-}$	primary commodities used as inputs by any agent
$K^{i+} \subset K$	secondary commodities produced as outputs by agent $i \in I$
$K^{i-} \subset K$	secondary commodities used as inputs by agent $i \in I$
$K^{I+} = \bigcup_{i \in I} K^{i+}$	secondary commodities produced as outputs by any agent
$K^{I-} = \bigcup_{i \in I} K^{i-}$	secondary commodities used as inputs by any agent
$L^{i+} \subset L$	utility commodities produced as outputs by agent $i \in I$
$L^{I+} = \bigcup_{i \in I} L^{i+}$	utility commodities produced as outputs by any agent
$M^{i-} \subset M$	non-market commodities used as inputs by agent $i \in I$
$M^{I-} = \bigcup_{i \in I} M^{i-}$	non-market commodities used as inputs by any agent
$I^{j-} \subset I$	agents that use primary commodity $j \in J$ as an input
$I^{J-} = \bigcup_{j \in J} I^{j+}$	agents that use any primary commodity as an input
$I^{k+} \subset I$	agents that produce secondary commodity $k \in K$ as an output
$I^{k-} \subset I$	agents that use secondary commodity $k \in K$ as an input
$I^{K+} = \bigcup_{k \in K} I^{k+}$	agents that produce any secondary commodity as an output
$I^{K-} = \bigcup_{k \in K} I^{k+}$	agents that use any secondary commodity as an input
$I^{l+} \subset I$	agents that produce utility commodity $l \in L$ as an output
$I^{L+} = \bigcup_{l \in L} I^{l+}$	agents that produce any utility commodity as an output
$I^{m-} \subset I$	agents that use non-market commodity $m \in M$ as an input
$I^{M-} = \bigcup_{m \in M} I^{m-}$	agents that use any non-market commodity as an input

[4] The elements of these subsets are determined by the agents' respective production technologies.

While it is intended that the framework should be as general as possible, it is necessary to impose some structure. Consequently, the following assumptions are made regarding the properties of the commodity and agent subsets.

Assumption A1: $|\mathrm{L}^{i+}| + |\mathrm{K}^{i+}| = 1$ for all $i \in \mathrm{I}$.

Assumption A2: $\mathrm{L}^{\mathrm{I}+} \neq \varnothing$ and $\mathrm{J}^{\mathrm{I}-} \neq \varnothing$

Assumption A3: $\mathrm{K} = \mathrm{K}^{\mathrm{I}+} = \mathrm{K}^{\mathrm{I}-} \neq \varnothing$

Assumption A4: $\mathrm{K}^{i-} \neq \varnothing$ for all $i \in \mathrm{I}^{\mathrm{L}+}$.

Assumption A5: $\mathrm{K}^{i+} \cap \mathrm{K}^{i-} = \varnothing$, and $\mathrm{J}^{i-} \neq \varnothing$ for all $i \in \mathrm{I}^{\mathrm{K}+}$.

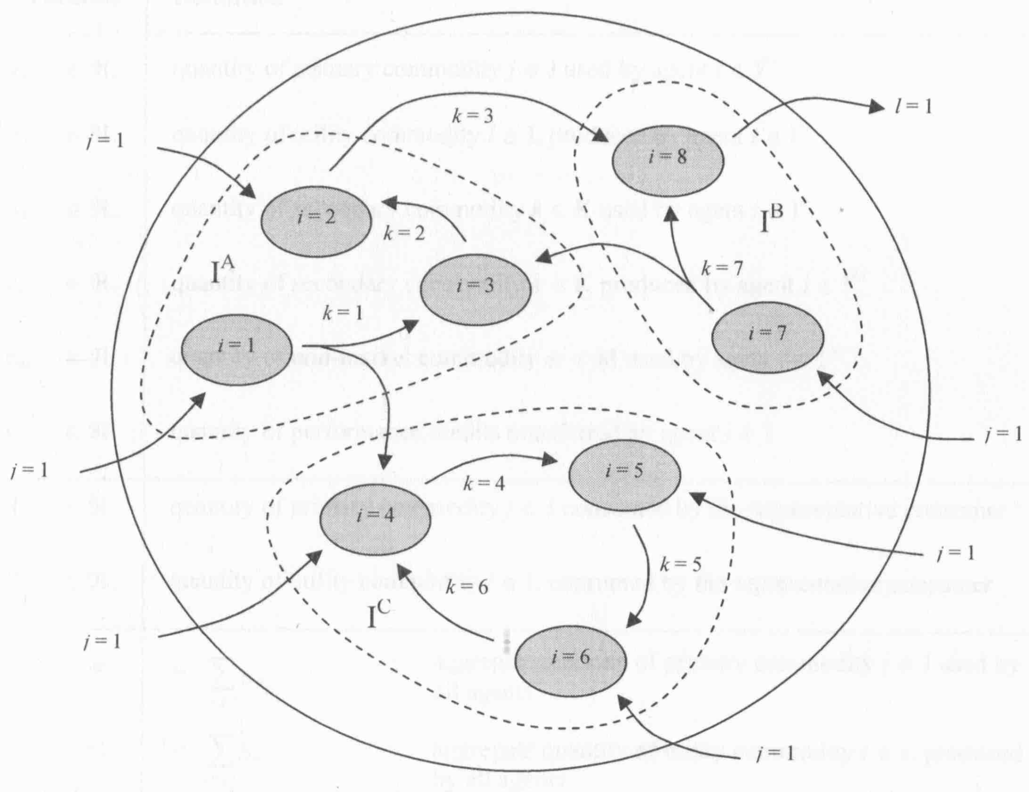
Assumption A6: $|\mathrm{I}^{j-}| > 1$ for some $j \in \mathrm{J}$

Assumption A7: for all $\tilde{\mathrm{I}} \subset \mathrm{I}^{\mathrm{K}+}$, there exists some $k \in \bigcup_{i \in \tilde{\mathrm{I}}} \mathrm{K}^{i+}$ such that $\mathrm{I}^{k-} \not\subset \tilde{\mathrm{I}}$.

Assumption A1 requires that each agent produces one output only – which may be a secondary commodity or a utility commodity. This is not as restrictive as it might appear at first sight. For example, if one assumes that firms comprise a number of independent operating plants (or divisions), then by defining the plants to be agents, it is possible to have firms producing more than one output. Assumption A2 requires that the system uses at least one primary input, and produces at least one utility commodity, while assumption A3 requires that all secondary commodities are both produced and used within the economy. While these two assumptions are trivial, in conjunction with the first assumption they imply that the number of agents must be greater than the number of secondary commodities.

Assumptions A4 and A5 respectively require that the subset of inputs for any agent that produces a utility commodity must include at least one secondary commodity, and that the subset of inputs for any agent that produces a secondary commodity must include at least one primary commodity. Assumption A6 requires that at least one of the primary commodities must be used as an input by more than one agent. While assumption A4 is trivial, assumptions A5 and A6 may appear rather restrictive. However, since most production processes require the input of labour – which is a primary commodity, they are not unduly onerous. Finally, assumption A7 requires that for any group of agents producing secondary commodities, at least one of the commodities produced by that group must be a potential input for at least one other agent in the economy. Again, since this does not preclude the "internal" use of commodities produced by the group, it is not very restrictive.

Figure 3.2 Hypothetical production system



Assumptions A1-A7 are illustrated in Figure 3.2 for a hypothetical production system in which there are eight agents – partitioned into three subsets, one primary commodity, one utility commodity, and seven secondary commodities. If the agents within I^A and I^C represent operating plants, then these subsets can be interpreted as firms; with firm I^A producing two secondary commodities (1 and 3), and so on. The system satisfies assumptions A1-A4 and A6 – each agent produces a single commodity, with agents 1 – 7 producing secondary commodities, and agent 8 producing consumption utility. However, it contravenes the assumptions A5 and A7. First, agent 3 produces a secondary commodity without using any primary input. Second, none of the secondary commodities produced by the constituent agents of group I^C are used by any other agent in the system.

Table 3.2 Definition of real system variables

Variable	Definition
$a_{ji} \in \Re_-$	quantity of primary commodity $j \in$ J used by agent $i \in I^{j-}$
$b_{li} \in \Re_+$	quantity of utility commodity $l \in$ L produced by agent $i \in I^{l+}$
$w_{ki} \in \Re_-$	quantity of secondary commodity $k \in$ K used by agent $i \in I^{k-}$
$y_{ki} \in \Re_+$	quantity of secondary commodity $k \in$ K produced by agent $i \in I^{k+}$
$z_{mi} \in \Re_-$	quantity of non-market commodity $m \in$ M used by agent $i \in I^{m-}$
$v_i \in \Re$	quantity of performance credits transferred by agent $i \in$ I
$A_j \in \Re_-$	quantity of primary commodity $j \in$ J consumed by the representative consumer
$B_l \in \Re_-$	quantity of utility commodity $l \in$ L consumed by the representative consumer
$a_j \in \Re_-$	$= \sum_{i \in I^{j-}} a_{ji}$ aggregate quantity of primary commodity $j \in$ J used by all agents
$b_l \in \Re_+$	$= \sum_{i \in I^{l+}} b_{li}$ aggregate quantity of utility commodity $l \in$ L produced by all agents
$x_k \in \Re$	$= \sum_{i \in I^{k+}} y_{ki} + \sum_{i \in I^{k-}} w_{ki}$ aggregate <u>net</u> quantity of secondary commodity $k \in$ K produced by all agents

Table 3.2 provides the definitions for all the real variables that are used in the analysis. Other variables – such as market and shadow prices – are defined in the text when they are introduced. As can be seen, production outputs are represented by (weakly) positive values, while production inputs and consumption are represented by (weakly) negative values. In addition to the basic variables listed in Table 3.2, a number vectors are defined for notational convenience. These vectors – which are defined in Table 3.3 – represent input and output plans (or allocations) for the various agents, and the aggregate consumption plans for primary and utility commodities.

Table 3.3 Definition of notational vectors

Vector	Definition
$\boldsymbol{a}_i \equiv (\dots a_{ji} \dots)$	primary input plan for agent $i \in \mathrm{I}$
$\boldsymbol{b}_i \equiv (\dots b_{li} \dots)$	utility output plan for agent $i \in \mathrm{I}$
$\boldsymbol{w}_i \equiv (\dots w_{ki} \dots)$	secondary input plan for agent $i \in \mathrm{I}$
$\boldsymbol{y}_i \equiv (\dots y_{ki} \dots)$	secondary output plan for agent $i \in \mathrm{I}$
$\boldsymbol{z}_i \equiv (\dots z_{mi} \dots)$	non-market input plan for agent $i \in \mathrm{I}$
$\widetilde{\boldsymbol{a}} \equiv (\boldsymbol{a}_1 \mid \dots \mid \boldsymbol{a}_{\mathrm{I}})$	primary input plan for all agents
$\widetilde{\boldsymbol{b}} \equiv (\boldsymbol{b}_1 \mid \dots \mid \boldsymbol{b}_{\mathrm{I}})$	utility output plan for all agents
$\widetilde{\boldsymbol{w}} \equiv (\boldsymbol{w}_1 \mid \dots \mid \boldsymbol{w}_{\mathrm{I}})$	secondary input plan for all agents
$\widetilde{\boldsymbol{y}} \equiv (\boldsymbol{y}_1 \mid \dots \mid \boldsymbol{y}_{\mathrm{I}})$	secondary output plan for all agents
$\widetilde{\boldsymbol{v}} \equiv (v_1 , \dots , v_{\mathrm{I}})$	performance credit plan for all agents
$\boldsymbol{a} \equiv (a_1 , \dots , a_{\mathrm{J}})$	aggregate input plan for primary commodities
$\boldsymbol{b} \equiv (b_1 , \dots , b_{\mathrm{L}})$	aggregate output plan for utility commodities
$\boldsymbol{x} \equiv (x_1 , \dots , x_{\mathrm{K}})$	aggregate net output plan for secondary commodities
$\boldsymbol{A} \equiv (A_1 , \dots , A_{\mathrm{J}})$	aggregate consumption plan for primary commodities
$\boldsymbol{B} \equiv (B_1 , \dots , B_{\mathrm{L}})$	aggregate consumption plan for utility commodities

The dimension of each vector is equal to the number of elements in the corresponding commodity / agent subset (as defined in Table 3.1). Consequently, for an individual agent $i \in \mathrm{I}$, one or more of the vectors $\boldsymbol{a}_i$, $\boldsymbol{b}_i$, $\boldsymbol{w}_i$ and $\boldsymbol{y}_i$ may have zero dimension (if the corresponding commodity set is empty for that agent). Indeed, under assumption A1 this must be the case for one of the two output vectors. Thus, while the *production plan* for agent $i \in \mathrm{I}$ is denoted by the concatenated vector $(\boldsymbol{a}_i \mid \boldsymbol{b}_i \mid \boldsymbol{w}_i \mid \boldsymbol{y}_i)$, this is done only to simplify notation and does not imply that both types of inputs are used, or that both types of outputs produced.

The *system production plan* is denoted by the concatenated vector $(\tilde{\boldsymbol{a}} \mid \tilde{\boldsymbol{b}} \mid \tilde{\boldsymbol{w}} \mid \tilde{\boldsymbol{y}})$.[5] Under assumptions A1-A6, all of the component input and output plans have at least one element, while the dimension of the system production plan is strictly greater than the combined number of agents, primary commodities, and secondary commodities (i.e. $\dim(\tilde{\boldsymbol{a}} \mid \tilde{\boldsymbol{b}} \mid \tilde{\boldsymbol{w}} \mid \tilde{\boldsymbol{y}}) > |\mathrm{I}| + |\mathrm{J}| + |\mathrm{K}|$).

The set of technically feasible production plans for agent $i \in \mathrm{I}$ (i.e. the agent's production set) is denoted by:[6]

$$\mathrm{G}^i \quad = \quad \{ (\boldsymbol{a}_i \mid \boldsymbol{b}_i \mid \boldsymbol{w}_i \mid \boldsymbol{y}_i) \mid \mathrm{g}^i(\boldsymbol{a}_i \mid \boldsymbol{b}_i \mid \boldsymbol{w}_i \mid \boldsymbol{y}_i) \leq 0 \} \qquad \ldots (3.1)$$

where

$$\mathrm{g}^i(\boldsymbol{a}_i \mid \boldsymbol{b}_i \mid \boldsymbol{w}_i \mid \boldsymbol{y}_i) \quad \equiv \quad b_{li} + \mathrm{f}^i(\boldsymbol{a}_i \mid \boldsymbol{w}_i) \qquad \text{for all } i \in \mathrm{I}^{l+}$$

$$\equiv \quad y_{ki} + \mathrm{f}^i(\boldsymbol{a}_i \mid \boldsymbol{w}_i) \qquad \text{for all } i \in \mathrm{I}^{k+}$$

[5] It should be noted that the system production plan does <u>not</u> represent the aggregate inputs and outputs of the various commodities.

[6] In order to simplify notation, it is assumed that the production set G^i is defined by single transformation function $\mathrm{g}^i(\cdot)$. However, this is not necessary, and in particular applications it may be defined by multiple functions, as is the case in the waste management application considered in Chapter 6.

The following assumptions are made regarding the properties of the production function of each agent $i \in \mathrm{I}$:

Assumption A8: f^i is continuous and convex for all $(\boldsymbol{a}_i \mid \boldsymbol{w}_i) \leq 0$, with $\mathrm{f}^i(\boldsymbol{a}_i \mid \boldsymbol{w}_i) \leq 0$ and $\mathrm{f}^i(\mathbf{0}) = 0$.

Assumption A9: the gradient vector $\nabla\mathrm{f}^i$ is continuous for all $(\boldsymbol{a}_i \mid \boldsymbol{w}_i) \ll \mathbf{0}$, with $\nabla\mathrm{f}^i \geq \mathbf{0}$ for all $(\boldsymbol{a}_i \mid \boldsymbol{w}_i) \leq \mathbf{0}$, and $\nabla\mathrm{f}^i \gg \mathbf{0}$ for all $(\boldsymbol{a}_i \mid \boldsymbol{w}_i) \ll \mathbf{0}$.

Assumption A10: if $\mathrm{f}^i(\boldsymbol{a}_i \mid \boldsymbol{w}_i) < 0$, then for all values $\zeta \in \Re_{++}$ there exist values δ, $\varepsilon \in \Re_{--}$, with $\delta < \varepsilon$, such that: $\mathrm{f}^i_j(\boldsymbol{a}_i \mid \boldsymbol{w}_i) > \zeta$ for any $j \in \mathrm{J}^{i-}$ with $a_{ji} > \varepsilon$; $\mathrm{f}^i_j(\boldsymbol{a}_i \mid \boldsymbol{w}_i) < \zeta$ for any $j \in \mathrm{J}^{i-}$ with $a_{ji} < \delta$; $\mathrm{f}^i_k(\boldsymbol{a}_i \mid \boldsymbol{w}_i) > \zeta$ for any $k \in \mathrm{K}^{i-}$ with $w_{ki} > \varepsilon$; and $\mathrm{f}^i_k(\boldsymbol{a}_i \mid \boldsymbol{w}_i) < \zeta$ for any $k \in \mathrm{K}^{i-}$ with $w_{ki} < \delta$.

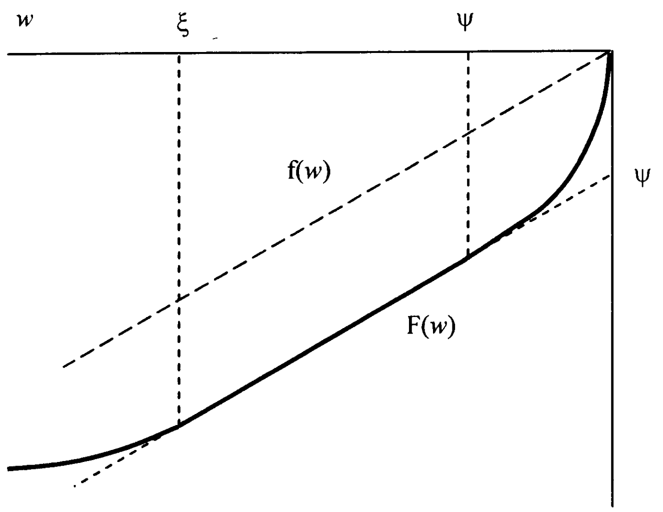
Assumption A11: for all values $\zeta \in \Re_{++}$, the directional derivative $\mathrm{D}_{\mathbf{u}}\mathrm{f}^i(\mathbf{0}) > \zeta$ for any unit vector $\mathbf{u} \ll \mathbf{0}$.

In addition, the production functions of each agent $i \in \mathrm{I}^{\mathrm{K}+}$ are all assumed to satisfy the following assumption:

Assumption A12: $\mathrm{f}^i(\boldsymbol{a}_i \mid \boldsymbol{w}_i) = 0$ if $\boldsymbol{a}_i = \mathbf{0}$.

Assumptions A8 – A11 are fairly standard, and are satisfied for a range of commonly used production functions. While assumptions A10 and A11 do not rule out constant returns to scale, they do imply that a production function can only be linear in any of its arguments over a finite range of values $\xi_{ki} \leq w_{ki} \leq \psi_{ki} < 0$. However, since the values of $|\xi_{ki}|$ and $|\psi_{ki}|$ can be arbitrarily large and small respectively, it is always possible to approximate a linear production function by a "composite" function which satisfies the assumptions (see Figure 3.3). Assumption A12 means that the production of any secondary commodity requires the input of some primary commodity. That is, primary commodities are – in aggregate – essential to production.[7] This is not very restrictive since – as was noted above – most production requires the input of labour, which is a primary commodity.

Figure 3.3 Approximating a linear production function with a single input



$$F(w) = \begin{cases} f^1(w) & \text{for } \psi < w \\ f(w) + \psi & \text{for } \xi \le w \le \psi \\ f^2(w) & \text{for } w < \xi \end{cases}$$

It follows from assumptions A8 and A9 that each production set G^i is closed and convex, with a non-empty interior. The *aggregate production set* for the system is denoted by:

$$G^I \quad = \quad \{\, (\boldsymbol{a} \mid \boldsymbol{b} \mid \boldsymbol{x}) \mid (\boldsymbol{a}_i \mid \boldsymbol{b}_i \mid \boldsymbol{w}_i \mid \boldsymbol{y}_i) \in G^i \text{ for all } i \in I \,\} \qquad \ldots (3.2)$$

It follows from assumptions A8 and A12 that G^I is convex and irreversible; that $\mathbf{0} \in G^I$; and that $G^I \cap \Re_+^{J+K+L} = \{0\}$.

The set of *feasible allocations* for the economy is denoted by:

[7] An input, or group of inputs, is essential if a strictly negative quantity is necessary for any output to be produced.

$$G = \{ (\boldsymbol{A} \mid \boldsymbol{B} \mid \tilde{\boldsymbol{a}} \mid \tilde{\boldsymbol{b}} \mid \tilde{\boldsymbol{w}} \mid \tilde{\boldsymbol{y}}) \mid (\boldsymbol{a}_i \mid \boldsymbol{b}_i \mid \boldsymbol{w}_i \mid \boldsymbol{y}_i) \in G^i \text{ for all } i \in I ;$$

$$A_j^0 + A_j + \sum_{i \in I^{j-}} a_{ji} \geq 0 \text{ for all } j \in J ;$$

$$\sum_{i \in I^{l+}} b_{li} + B_l \geq 0 \text{ for all } l \in L ;$$

$$\sum_{i \in I^{k+}} y_{ki} + \sum_{i \in I^{k-}} w_{ki} \geq 0 \text{ for all } k \in K \} \qquad \ldots (3.3)$$

where A_j^0 endowment of primary resource $j \in J$. It follows from the properties of the individual production sets (3.1) and aggregate production set (3.2) that G is convex and compact, with a non-empty interior.[8]

The quantity of non-market commodity $m \in M^{i-}$ that is as used by agent $i \in I$ is not a direct choice variable. Rather it is determined by the quantities that the agent chooses for its various secondary inputs $k \in K^{i-}$; being defined by the "emissions function" $z_{mi} = h^{mi}(\boldsymbol{w}_i)$. The following two assumptions are made regarding the properties of the emissions function for each non-market commodity $m \in M^{i-}$ used by agent $i \in I$:

Assumption A13: h^{mi} is continuously differentiable and concave, with $h^{mi}(\boldsymbol{w}_i) \leq 0$ for all $\boldsymbol{w}_i \leq \boldsymbol{0}$ and $h^{mi}(\boldsymbol{0}) = 0$.

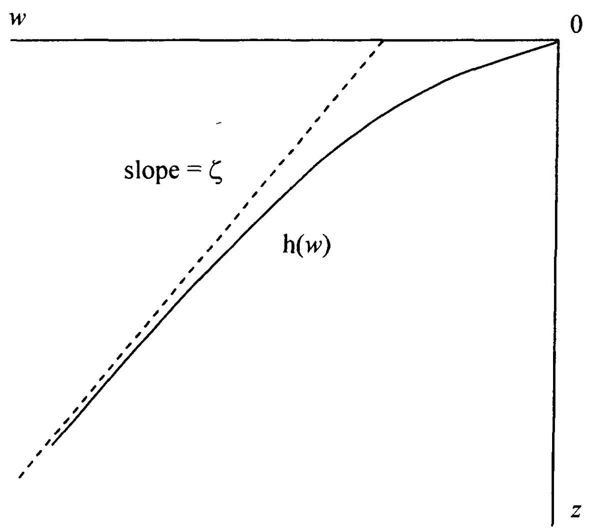
Assumption A14: there exists a value $\zeta \in \Re_+$ such that $| h^{mi}_k | \leq \zeta$ for all $\boldsymbol{w}_i \leq \boldsymbol{0}$

These assumptions are not very restrictive. While the magnitudes of the partial derivatives of the emissions function are bounded, no restriction is imposed on their sign. Consequently, the use of secondary commodities to abate emissions is allowed (in which case the respective partial derivatives are negative). They are illustrated in

[8] The convexity and non-emptiness of G follow directly from the properties of the individual production sets. The compactness of G follows from the closed-ness of the individual production sets and the properties of the aggregate production set (see Proposition 16.AA.1 in Mas-Colell *et al*, 1995).

Figure 3.4, for the case where emissions are an increasing function of a single market input.

Figure 3.4 Example of an emissions function



The set of *allowable allocations* for the economy under the aggregate performance rule is denoted by:

$$R = \{ (A \mid B \mid \tilde{a} \mid \tilde{b} \mid \tilde{w} \mid \tilde{y}) \mid \Omega(\tilde{w} \mid \tilde{y} ; \alpha \mid \beta \mid \chi) \leq \delta \} \qquad \dots (3.4)$$

where $\delta \geq 0$, and:

$$\Omega(\tilde{w} \mid \tilde{y} ; \alpha \mid \beta \mid \chi) = -\sum_{k \in K} \alpha_k \sum_{i \in I^{k+}} y_{ki} - \sum_{k \in K} \beta_k \sum_{i \in I^{k-}} w_{ki} - \sum_{m \in M} \chi_m \sum_{i \in I^{m-}} h^{mi}(w_i)$$

The performance rule parameters α_k, β_k and χ_m are set equal to zero for all un-regulated secondary and non-market commodities. For regulated commodities, the following restrictions are imposed on the values of the parameters:

Assumption A15: $\chi_m \geq 0$ for all $m \in M$

Assumption A16: $\alpha_k \, \beta_k = 0$ for all $k \in \mathrm{K}$

Assumption A17: $\alpha_k \left(\sum_{m \in \mathrm{M}^{i-}} \chi_m \, \mathrm{h}_k^{mi}(\boldsymbol{w}_i) \right) = 0$ for all $k \in \mathrm{K}$ and $i \in \mathrm{I}^{m-}$

Assumption A18: $\beta_k \left(\sum_{m \in \mathrm{M}^{i-}} \chi_m \, \mathrm{h}_k^{mi}(\boldsymbol{w}_i) \right) = 0$ for all $k \in \mathrm{K}$ and $i \in \mathrm{I}^{m-}$

None of these restrictions is very onerous, and it is hard to envisage any practical applications in which they would have any impact. Assumption A15 implies that the value of the constraint function does not decrease when the use of any non-market commodity $m \in \mathrm{M}$ increases in magnitude (i.e. when emissions increase). That is, emissions are always treated as a "bad". Assumption A16 implies that for each secondary commodity $k \in \mathrm{K}$, either $\alpha_k = 0$ or $\beta_k = 0$ (or both). As such it prohibits performance rules that include both the inputs and the outputs of a particular commodity.

Finally, assumptions A17 and A18 imply that if a secondary commodity enters into the emissions function of a pollutant that is included in the performance rule, then that commodity cannot also be included directly in the rule – either as an output, or as an input. However, the assumptions do not preclude the inclusion of the outputs, or inputs, of commodities that do not directly affect the emissions. For example, while assumption A18 would prohibit a rule that included both the inputs of coal and the emissions of carbon dioxide, it would allow a rule that combined the emissions with inputs of electricity.

Since the emissions functions h^{mi} are all continuous and concave, the set R is closed and convex, with a non-empty interior. Finally, the aggregate performance rule is defined to be *technically feasible* if the set $\mathrm{G} \cap \mathrm{R}$ has a non-empty interior. It is assumed that this is always the case.

The set of allowable "augmented" production plans for agent $i \in \mathrm{I}$ under its individual performance rule is:

$$\mathrm{R}^i \;=\; \{\, (\boldsymbol{a}_i \,|\, \boldsymbol{b}_i \,|\, \boldsymbol{w}_i \,|\, \boldsymbol{y}_i \,|\, v_i) \;\mid\; \Omega^i(\boldsymbol{w}_i \,|\, \boldsymbol{y}_i \,|\, v_i \,;\, \boldsymbol{\rho} \,|\, \boldsymbol{\sigma} \,|\, \boldsymbol{\chi}\,) \;\le\; \delta_i \,\} \qquad \ldots (3.5)$$

where

$$\Omega^i(\boldsymbol{w}_i \,|\, \boldsymbol{y}_i \,|\, v_i \,;\, \boldsymbol{\rho} \,|\, \boldsymbol{\sigma} \,|\, \boldsymbol{\chi}\,) \;=\; -\sum_{k \in \mathrm{K}^{i+}} \rho_k\, y_{ki} \;-\; \sum_{k \in \mathrm{K}^{i-}} \sigma_k\, w_{ki} \;-\; \sum_{m \in \mathrm{M}^i} \chi_m\, \mathrm{h}^{mi}(\boldsymbol{w}_i) \;+\; v_i$$

The value of the parameter χ_m is the same as in the aggregate performance rule; while the parameters ρ_k, σ_k and δ_i are related to α_k, β_k and δ by the following identities:

$$\rho_k \;\equiv\; (1-\theta_k)\,\alpha_k \;-\; \theta_k\,\beta_k \qquad \text{for all } k \in \mathrm{K}$$

$$\sigma_k \;\equiv\; (1-\theta_k)\,\beta_k \;-\; \theta_k\,\alpha_k \qquad \text{for all } k \in \mathrm{K}$$

$$\delta_i \;\equiv\; (\delta\,/\,\mathrm{I}) \;+\; \varepsilon_i \qquad \text{for all } i \in \mathrm{I}$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_{\mathrm{K}})$ and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_{\mathrm{I}})$ are vectors of regulatory design parameters.[9] The sets of allowable values for these parameter vectors are denoted respectively by:

$$\Theta \;=\; \{\, (\theta_1, \ldots, \theta_{\mathrm{K}}) \;\mid\; \theta_k \in [0,1] \subset \Re_+ \text{ for all } k \in \mathrm{K} \,\}$$

$$\mathrm{E} \;=\; \{\, (\varepsilon_1, \ldots, \varepsilon_{\mathrm{I}}) \;\mid\; \varepsilon_i \in \Re \text{ for all } i \in \mathrm{I}, \text{ and } \sum_{i \in \mathrm{I}} \varepsilon_i = 0 \,\}$$

Again, since each emissions function h^{mi} is concave, Ω^i is convex, and hence each set R^i is closed and convex, with a non-empty interior.

---

[9] The interpretation of the parameter vectors $\boldsymbol{\theta}$ and $\boldsymbol{\varepsilon}$ is discussed in Chapter 2.

Under the assumptions that have been made about the properties of the aggregate and individual welfare functions, aggregate gross economic welfare is given by:

$$W(\boldsymbol{A} \mid \boldsymbol{B}) \quad = \quad - \sum_{l \in L} \varphi_l B_l \quad - \sum_{j \in J} \gamma_j A_j \qquad \ldots (3.6)$$

This can be interpreted as the welfare function for a single representative consumer. Without any loss of generality, it is assumed that $\phi_l = 1$ for all $l \in$ L, and that $\gamma_j = 1$ for all $j \in$ J.

In the absence of any regulatory intervention, the maximum aggregate benefit that can be generated by the system – denoted by W^u – is determined by solving the *unregulated benefit problem* (UB):

$$\text{Maximize} \quad W(\boldsymbol{A} \mid \boldsymbol{B}) \qquad \text{subject to} \qquad (\boldsymbol{A} \mid \boldsymbol{B} \mid \tilde{\boldsymbol{a}} \mid \tilde{\boldsymbol{b}} \mid \tilde{\boldsymbol{w}} \mid \tilde{\boldsymbol{y}}) \in G$$

Since $W(\boldsymbol{A} \mid \boldsymbol{B})$ is continuous and the constraint set G is compact and non-empty, the solution set for this problem is also non-empty and compact (i.e. a solution exists and all solutions are finite).

It is assumed that markets exist for all commodities $j \in$ J, $k \in$ K and $l \in$ L, and that the representative consumer and all productive agents $i \in$ I are price-takers. Furthermore, the system is not subject to any regulatory interventions – other than that represented by the aggregate performance rule. In particular, there is symmetry between the buying and selling prices of each market commodity.[10] Consequently, the gross economic benefit for agent $i \in$ I is given by:

$$\Pi^i(\boldsymbol{a}_i \mid \boldsymbol{b}_i \mid \boldsymbol{w}_i \mid \boldsymbol{y}_i \mid v_i) \quad = \quad \sum_{j \in J^{i-}} p_j a_{ji} \; + \; \sum_{l \in L^{i+}} p_l b_{li} \; + \; \sum_{k \in K^{i-}} p_k w_{ki} \; + \; \sum_{k \in K^{i+}} p_k y_{ki} \; + \; q\, v_i \qquad \ldots (3.7)$$

where $\mathbf{p} \in \Re_+^{J+K+L}$ is the vector of market prices for the primary, secondary and utility commodities, and $q \in \Re_+$ is the market price of performance credits. The gross economic benefits of all production agents accrue to the representative consumer.

3.2 Aggregate cost minimum

The *aggregate gross economic cost* of any feasible allocation $(\boldsymbol{A} \mid \boldsymbol{B} \mid \tilde{\boldsymbol{a}} \mid \tilde{\boldsymbol{b}} \mid \tilde{\boldsymbol{w}} \mid \tilde{\boldsymbol{y}})$ is defined to be $W(\boldsymbol{A} \mid \boldsymbol{B}) - W^u$, where W^u is the maximal value of the aggregate benefit function in the unregulated benefit problem (UB). By construction, aggregate gross economic cost is non-positive. Consequently, the *regulated aggregate cost minimization problem* (RC) is to:

$$\text{Maximize} \quad W(\boldsymbol{A} \mid \boldsymbol{B}) - W^u \quad \text{subject to} \quad (\boldsymbol{A} \mid \boldsymbol{B} \mid \tilde{\boldsymbol{a}} \mid \tilde{\boldsymbol{b}} \mid \tilde{\boldsymbol{w}} \mid \tilde{\boldsymbol{y}}) \in G \cap R$$

The Lagrangian for this problem is (in expanded form):

$$\begin{aligned}
\mathcal{L} \;=\; & -\sum_{l \in L} B_l \;-\; \sum_{j \in J} A_j \;-\; W^u \\
& -\sum_{j \in J} \sum_{i \in I^{j+}} \mu_i \left[b_{ji} + f^j(\boldsymbol{a}_i \mid \boldsymbol{w}_i) \right] \;-\; \sum_{j \in J} \sum_{i \in I^{k+}} \mu_i \left[y_{ki} + f^i(\boldsymbol{a}_i \mid \boldsymbol{w}_i) \right] \\
& +\sum_{l \in L} \varphi_l \left[\sum_{i \in I^{l+}} b_{li} + B_l \right] \;+\; \sum_{j \in J} \gamma_j \left[A_j^0 + A_j + \sum_{i \in I^{j-}} a_{ji} \right] \\
& +\sum_{k \in K} \eta_k \left[\sum_{i \in I^{k+}} y_{ki} + \sum_{i \in I^{k-}} w_{ki} \right] \\
& +\lambda \left[\sum_{k \in K} \alpha_k \sum_{i \in I^{k+}} y_{ki} + \sum_{k \in K} \beta_k \sum_{i \in I^{k-}} w_{ki} + \sum_{m \in M} \chi_m \sum_{i \in I^{m-}} h^{mi}(\boldsymbol{w}_i) + \delta \right]
\end{aligned}$$

---

[10] The impact of relaxing this assumption is demonstrated in Chapter 6, while the implications of strategic price-setting behaviour are considered in Part 3 of the thesis.

Suppose that the allocation $(\mathbf{A}^* \mid \mathbf{B}^* \mid \tilde{\mathbf{a}}^* \mid \tilde{\mathbf{b}}^* \mid \tilde{\mathbf{w}}^* \mid \tilde{\mathbf{y}}^*)$ is a solution to the regulated cost problem, and that the rank of the Jacobian matrix of the binding constraints with respect to the non-zero variables is equal to the number of binding constraints.[11] Then there exists a unique vector $(\mu^* \mid \phi^* \mid \gamma^* \mid \eta^* \mid \lambda^*)$ of finite, non-negative shadow prices, such that $(\mathbf{A}^* \mid \mathbf{B}^* \mid \tilde{\mathbf{a}}^* \mid \tilde{\mathbf{b}}^* \mid \tilde{\mathbf{w}}^* \mid \tilde{\mathbf{y}}^*)$ and $(\mu^* \mid \phi^* \mid \gamma^* \mid \eta^* \mid \lambda^*)$ satisfy the necessary Kuhn-Tucker first order conditions RC1-RC12 set out in Table A3.1.1 of Appendix A3.1.

Since the objective function is concave, and each constraint function is convex, conditions RC1-RC12 are also sufficient. Hence, if there exists a production plan $(\mathbf{A}^* \mid \mathbf{B}^* \mid \tilde{\mathbf{a}}^* \mid \tilde{\mathbf{b}}^* \mid \tilde{\mathbf{w}}^* \mid \tilde{\mathbf{y}}^*)$ satisfying the first-order conditions, then this solution is a global maximizer of the regulated cost minimization problem (RC). Moreover, the existence of a finite solution is assured if – as is assumed – the aggregate performance rule is technically feasible. That is:

<u>Proposition 3.1</u>

The solution set of the regulated cost minimization problem (RC) – denoted by Ψ^* – is non-empty and compact.

Proof: *The objective function in RC is continuous, and the constraint set $G \cap R \subset G$ is non-empty if the aggregate performance rule is technically feasible. Furthermore, since it is the intersection of two closed sets, and the subset of a bounded set, the constraint set is compact. Consequently, Weierstrass's Theorem ensures that a solution exists. Noting that $\Psi^* = W^* \cap (G \cap R) \subset G$, where $W^* = \{ (\boldsymbol{A} \mid \boldsymbol{B} \mid \tilde{\boldsymbol{a}} \mid \tilde{\boldsymbol{b}} \mid \tilde{\boldsymbol{w}} \mid \tilde{\boldsymbol{y}}) \mid W(\boldsymbol{A} \mid \boldsymbol{B}) \geq W^* \}$ is closed, it follows directly that the solution set is also compact.*

[11] This ensures that the non-degenerate constraint qualification (NDCQ) is satisfied. See Theorem 19.12, Simon & Blume (1994).

Characterization of the regulated cost minimum

Under the various assumptions set out in section 3.1, it is possible to show that in any solution, the system production plan and the utility consumption plan must both be strictly non-zero (i.e. <u>all</u> elements take non-zero values), and that all production constraints and all commodity constraints hold with equality. That is:

<u>Proposition 3.2</u>

If the allocation $(\mathbf{A}^* \mid \mathbf{B}^* \mid \tilde{\mathbf{a}}^* \mid \tilde{\mathbf{b}}^* \mid \tilde{\mathbf{w}}^* \mid \tilde{\mathbf{y}}^*)$ and the shadow price vector $(\mu^* \mid \phi^* \mid \gamma^* \mid \eta^* \mid \lambda^*)$ satisfy conditions RC1-RC12, and assumptions A1-A18 are satisfied,

then (i) $(\mathbf{B}^* \mid \tilde{\mathbf{a}}^* \mid \tilde{\mathbf{w}}^*) << \mathbf{0}$ and $(\tilde{\mathbf{b}}^* \mid \tilde{\mathbf{y}}^*) >> \mathbf{0}$

(iii) $\mu_i^* > 0 \qquad b_{li}^* + f^i(\mathbf{a}_i^* \mid \mathbf{w}_i^*) = 0 \qquad$ for all $l \in$ L and $i \in \mathrm{I}^{l+}$

$\mu_i^* > 0 \qquad y_{ki}^* + f^i(\mathbf{a}_i^* \mid \mathbf{w}_i^*) = 0 \qquad$ for all $k \in$ K and $i \in \mathrm{I}^{k+}$

(iv) $\phi_j^* > 0 \qquad \sum_{i \in \mathrm{I}^{j-}} a_{ji}^* + A_j^0 + A_j^* = 0 \qquad$ for all $j \in$ J

$\gamma_l^* > 0 \qquad \sum_{i \in \mathrm{I}^{l+}} b_{li}^* + B_l^* = 0 \qquad$ for all $l \in$ L

$\eta_k^* > 0 \qquad \sum_{i \in \mathrm{I}^{k+}} y_{ki}^* + \sum_{i \in \mathrm{I}^{k-}} w_{ki}^* = 0 \qquad$ for all $k \in$ K

Proof: see Appendix A3.2

It follows from assumptions A1-A6 and A16-A18 that the number of non-zero variables in the solution is at least as great as the number of binding constraints, and that the gradient vectors associated with the constraints are linearly independent.[12] Hence, the rank of the Jacobian matrix of constraint functions is equal to the number of constraints, as assumed.

An important point to note about point (ii) of Proposition 3.1 is that it does not include the consumption plan for primary resources. Unlike the consumption plan for utility commodities, it is perfectly possible for some, or all, of the elements of this consumption plan to be equal to zero, and hence that the respective primary resources are used wholly for production. In this case, the primary resource is said to be scarce.

Since the system production plan and the utility consumption plan are both strictly non-zero, the shadow price of the production constraint (μ_i^*) can be eliminated to give the following four sets of marginal product conditions for agent $i \in$ I:

$$f_j^{i\,*} = \phi_j^* \geq 1 \qquad \text{for all } l \in \mathrm{L}^{i+}, j \in \mathrm{J}^{i-} \quad \ldots (3.8.a)$$

$$f_k^{i\,*} = \eta_k^* + \lambda^* \left(\beta_k + \sum_{m \in \mathrm{M}^i} \chi_m \mathrm{h}_k^{mi\,*} \right) \qquad \text{for all } l \in \mathrm{L}^{i+}, k \in \mathrm{K}^{i-} \quad \ldots (3.8.b)$$

$$(\eta_{k'}^* + \lambda^* \alpha_{k'})\, f_k^{i\,*} = \phi_j^* \geq 1 \qquad \text{for all } k' \in \mathrm{K}^{i+}, j \in \mathrm{J}^{i-} \quad \ldots (3.8.c)$$

$$(\eta_{k'}^* + \lambda^* \alpha_{k'})\, f_k^{i\,*} = \eta_k^* + \lambda^* \left(\beta_k + \sum_{m \in \mathrm{M}^{i-}} \chi_m \mathrm{h}_k^{mi\,*} \right) \qquad \text{for all } k' \in \mathrm{K}^{i+}, k \in \mathrm{K}^{i-} \quad \ldots (3.8.d)$$

where $\quad f_k^{i\,*} = f_k^i(a_i^* | w_i^*) \quad$ and $\quad h^{mi\,*}_k = h^{mi}_k(w_i^*)$

[12] Noting that $\mathbf{B}^* >> 0$, assumptions A1-A6 ensure that the minimum number of non-zero variables in the solution is | I | + | J | + | K | + | L | + 1; which is equal to the maximum number of binding constraints. Assumptions A1, A3 and A16-A18 ensure that the constraints are linearly independent.

Under assumption A1, only two of these sets of conditions will apply to any individual agent. If the agent produces a utility commodity, then the first two sets apply (since $K^{i+} = \varnothing$). If it produces a secondary commodity, then the last two sets apply. In either case the right hand side represents the total shadow cost of the input commodity to the agent, while the left hand side represents the marginal product of that input valued at the total shadow price of output.

While the shadow prices of all utility commodities are constant (equal to one), the shadow prices of primary resources ($\boldsymbol{\phi}^*$) may – in general – vary for different values of the performance rule parameters. However, if the initial endowment of each primary resource $j \in J$ is sufficiently large, then under assumption A10 it must be the case that $\mathbf{A}^* \gg \mathbf{0}$. That is, some of each primary resource is "consumed" directly. In this case the shadow prices of all primary commodities are also constant (and equal to one).

Noting that for each secondary commodity $k \in K$, either $\alpha_k = 0$, or $\beta_k + \Sigma\, \chi_m\, h_k^{mi*} = 0$, or both, it is clear from conditions (3.8) that cost efficiency requires that a wedge be driven between the input price and output price of each secondary commodity that is included (either directly or indirectly) in the aggregate performance rule. Depending on the specific values of the elements of the parameter vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\chi}$, the input price of a particular commodity may be higher than the output price, or it may be lower. For example, if the aggregate performance rule represents a relative standard for vehicle fuel efficiency, then the output price of model $k \in K^V \subset K$ is $\eta_k^* + \lambda^*(r_k - r)$, where r_k is the fuel efficiency of the model and r is the standard.[13] Hence the output price is higher than the input price for fuel efficient models (i.e. for all $k \in K^V$ with $r_k > r$), and lower for inefficient models.

[13] See Chapter 2.

It should be noted that this observation relates to the difference between the input and output prices in the solution to the regulated problem (RC). In general, it is not possible to say whether the value of either of these prices is higher, or lower, than the corresponding price in the solution to the unregulated benefit problem (UB). However, it may be possible to draw such conclusions if additional assumptions are made about the production technologies of certain agents, as is done in the applications considered in Chapters 5 and 6.

Interpretation of Lagrange multiplier λ^*

Before moving on to analyse the market equilibrium under performance-based credit trading, it is instructive to consider the interpretation of the Langrange multiplier λ^*. Denoting the value function for the regulated cost minimization problem by V(L, r), and the maximal value of the aggregate benefit function by Π^*, it follows that: [14]

$$dW^* = V_L\, dL + V_r\, dr$$

$$= \lambda^* \left[dL + \Omega_r(\tilde{w}^* \mid \tilde{y}^*; r)\, dr \right] \qquad \dots (3.9)$$

Thus λ^* is a measure of the increase in aggregate benefit that would arise from a relaxation of the aggregate performance rule – via marginal changes to the target value parameters L and r. However, this interpretation is not very intuitive, and in many applications one would like to interpret the shadow price in terms of a change in the aggregate input or output value of an "appropriate commodity" (or group of commodities).[15] For example, if the performance rule relates to the aggregate emissions of greenhouse gases, then one would like to interpret λ^* in terms of the impact of a

---

[14] As was noted in Chapter 2, the performance rule parameters (α, β, χ and δ) represent a combination of exogenous scaling parameters and regulatory target values (L and r). Since the scaling parameters are fixed, only the two target values are shown as arguments of the value function.

change in the aggregate value of the non-market environmental inputs relating to the absorption of greenhouse gases.

Of course, the identification of the "appropriate commodity" will depend on the particular application. Consequently, in order to proceed, it is necessary to make some assumption about the definition of the aggregate performance rule, and hence the underlying regulatory target. Staying with the example of greenhouse gases, let us assume that the aggregate performance rule is:

$$\mathrm{r}\,Y \;+\; Z \;+\; \mathrm{L} \;\geq\; 0$$

where:

$$Z \;=\; \sum_{m \in \mathrm{M}} \mathrm{c}_m \sum_{i \in \mathrm{I}^{m-}} \mathrm{h}^{mi}(\boldsymbol{w}_i) \quad < \quad 0 \qquad \text{denotes aggregate emissions}$$

$$Y \;=\; \sum_{k \in \mathrm{K}} \mathrm{a}_k \sum_{i \in \mathrm{I}^{j+}} y_{ki} \quad > \quad 0 \qquad \text{denotes aggregate output}$$

This rule incorporates three different forms of regulatory target. If r = 0 and L > 0, then the rule represents an absolute limit for aggregate emissions; if r > 0 and L = 0, then it represents a relative standard for aggregate specific emissions; and if r > 0 and L > 0, then it represents a hybrid target for aggregate emissions (see section 2.4). As was noted above, for this performance rule the obvious candidate for the "appropriate commodity" is aggregate emissions, in which case λ^* is interpreted in terms of the marginal cost of emissions reduction (or abatement), i.e. dW^*/dZ^*.[16]

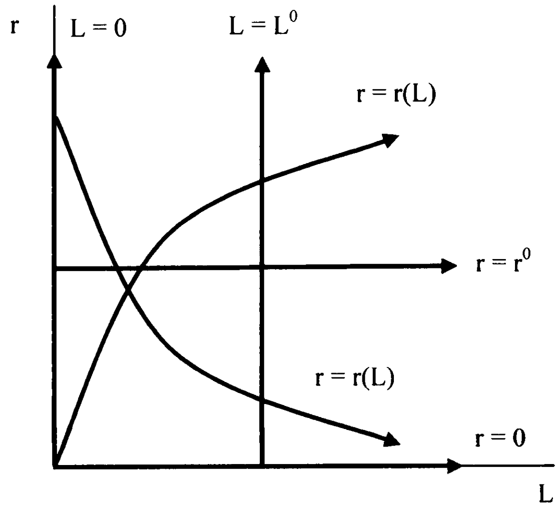
---

[15] In some applications there may not be an appropriate commodity. This is the case, for example, if the performance rule represents a relative standard for the weighted average fuel efficiency of motor vehicles.

[16] It should be noted that an increase in Z^* is equivalent to a decline in the magnitude of emissions. Thus, the differential dZ^* represents marginal abatement.

It is also necessary to stipulate the path that the changes to the two regulatory parameters will follow – i.e. how L and r change together. Figure 3.5 shows a number of possible paths. As can be seen, the value of either parameter may be held constant, or the two values may be related according to some specified function r(L).[17]

Figure 3.5 Relationship between regulatory parameters L and r



The various possibilities can be divided into two cases. In the first case, the value of r remains constant as L changes (i.e. $r = r^0$). Since conditions RC1-RC12 are necessary and sufficient, the second order sufficiency conditions are satisfied. Consequently, the implicit function theorem is applicable, and the conditions can be solved explicitly to give $Z^* = Z(r^0, L)$ and $Y^* = Y(r^0, L)$. Assuming that the first of these functions is strictly monotone in the neighbourhood of the solution, it can be inverted to give $L = L(Z^*; r^0)$ with $L' = 1 / Z_L$. Hence:

$$W^* \quad = \quad V(r^0, L(Z^*; r^0))$$

[17] The functions shown in Figure 3.5 are both monotone. However, this need not necessarily be the case, and r(L) may be increasing for some values of L, and decreasing for others.

$$dW^* = V_L \, L' \, dZ^*$$

$$= \lambda^* \, (1 / Z_L) \, dZ^*$$

Assuming that $\lambda^* > 0$, then the performance rule constraint will remain binding for marginal changes to L. Consequently, the constraint can be expressed as:

$$r\,Y(r^0, L) \;+\; Z(r^0, L) \;+\; L \;\equiv\; 0.$$

Since this is an identity, it can be differentiated to yield:

$$r\,Y_L \;+\; Z_L \;+\; 1 \;=\; 0$$

which implies that $\dfrac{1}{Z_L} = -\left[1 + r\dfrac{Y_L}{Z_L}\right]$

Noting that for $r = r^0$, $dZ^* = Z_L\, dL$, and that $dY^* = Y_L\, dL$, it follows that:

$$\lambda^* \;=\; -\frac{1}{(1+\xi^*)}\,\frac{dW^*}{dZ^*} \qquad \text{where} \quad \xi^* = r\,\frac{dY^*}{dZ^*}$$

Thus, λ^* is equal to the marginal cost of emissions reduction multiplied by a factor that reflects the relative responsiveness of output and emissions to the changes in L when r is held constant. As will be seen in Chapter 5, it does not necessarily follow that the changes to output and emissions will be in the same direction (i.e. that $\xi^* > 0$). However, in most cases this is likely to be so, in which case λ^* understates the marginal cost of abatement. If $r = 0$, then λ^* is equal to the marginal cost of abatement, which – as was noted above – is the usual interpretation in the context of an absolute limit.

In the second case, the value of the parameter r varies, while the value of L may be held constant, or may vary according to the function L(r). Consequently, the explicit solutions are $Z^* = Z(r, L(r))$ and $Y^* = Y(r, L(r))$. Again, assuming that the first of these functions is strictly monotone in the neighbourhood of the solution, it can be inverted to give $r = r(Z^*)$, with $r'(Z^*) = 1 / Z'(r, L(r))$.[18] Hence:

$$W^* \quad = \quad V(r(Z^*), L(r(Z^*)))$$

$$dW^* \quad = \quad [\, V_r\, r' + V_r\, L'\, r' \,]\, dZ^*$$

$$\quad = \quad \lambda^* [\, Y^* + L' \,] / Z'\, dZ^*$$

Again, assuming that $\lambda^* > 0$, then the performance rule constraint will remain binding for marginal changes to r. Consequently, it can be expressed as the identity:

$$r\,Y(r, L(r)) \; + \; Z(r, L(r)) \; + \; L(r) \; \equiv \; 0$$

which can be differentiated to yield:

$$Y^* \; + \; r\,Y' \; + \; Z' \; + \; L' \; = \; 0$$

Dividing through by Z', and rearranging gives: $\dfrac{Y^* + L'}{Z'} = -\left[\, 1 + r\dfrac{Y'}{Z'} \,\right]$

Noting that for $L = L(r)$, $dZ^* = Z'\, dr$ and $dY^* = Y'\, dr$, it follows that:

$$\lambda^* \quad = \quad - \frac{1}{(1+\xi^*)}\,\frac{dW^*}{dZ^*} \qquad \text{where} \quad \xi^* = r\,\frac{dY^*}{dZ^*}$$

Thus, as in the previous case, λ^* is equal to the marginal cost of abatement multiplied by a factor that reflects the relative responsiveness of output and emissions. However, while the expression is the same, it should be noted that the value of ξ^* will (in general) be different.

[18] Note that this rules out certain forms for the function L(r). In particular, it is not possible to have a

As was noted above, this case includes the possibility of holding the value of L constant. In particular it includes the case L = 0. Thus, it follow that unless the optimal value of aggregate output is unaffected by changes to the value of r (i.e. $Y' = 0$), the value of λ^* will not be equal to the marginal cost of abatement under a relative standard for specific emissions.

3.3 Market equilibrium

A competitive market equilibrium with tradable performance credits comprises a vector of non-negative prices $(\mathbf{p} \mid q)$; a consumption plan $(\mathbf{A}^\# \mid \mathbf{B}^\#)$; and an augmented system production plan $(\tilde{\mathbf{a}}^\# \mid \tilde{\mathbf{b}}^\# \mid \tilde{\mathbf{w}}^\# \mid \tilde{\mathbf{y}}^\# \mid \tilde{\mathbf{v}}^\#)$ such that:

- given market prices $\mathbf{p}$, initial endowments of primary resources $\mathbf{A}^0$, and income Π, the consumption plan $(\mathbf{A}^\# \mid \mathbf{B}^\#)$ solves welfare maximization problem of the representative consumer:

$$\text{Maximize} \quad W(\boldsymbol{A} \mid \boldsymbol{B})$$

$$\text{subject to} \quad \Pi^\# + \sum_{j \in J} p_j (A_j + A_j^0) + \sum_{l \in L} p_l B_l \geq 0$$

- given market prices $(\mathbf{p} \mid q)$, the augmented production plan $(\mathbf{a}_i^\# \mid \mathbf{b}_i^\# \mid \mathbf{w}_i^\# \mid \mathbf{y}_i^\# \mid v_i^\#)$ solves the regulated benefit maximisation problem (RB_i) of agent $i \in I$:

$$\text{Maximize} \quad \Pi^i(\boldsymbol{a}_i \mid \boldsymbol{b}_i \mid \boldsymbol{w}_i \mid \boldsymbol{y}_i \mid v_i)$$

$$\text{subject to} \quad (\boldsymbol{a}_i \mid \boldsymbol{b}_i \mid \boldsymbol{w}_i \mid \boldsymbol{y}_i \mid v_i) \in G^i \cap R^i$$

- the market clearing conditions are satisfied for all primary, secondary and utility commodities, and for performance credits:

---

relationship between L and r which results in aggregate emissions remaining constant as r changes.

$$\sum_{i \in I^{j-}} a_{ji} + A_j^0 + A_j \geq 0 \qquad p_j \left[A_j^0 + A_j + \sum_{i \in I^{j-}} a_{ji} \right] \qquad \text{for all } j \in J$$

$$\sum_{i \in I^{l+}} b_{li} + B_l \geq 0 \qquad p_l \left[\sum_{i \in I^{l+}} b_{li} + B_l \right] \qquad \text{for all } l \in L$$

$$\sum_{i \in I^{k+}} y_{ki} + \sum_{i \in I^{k-}} w_{ki} \geq 0 \qquad p_k \left[\sum_{i \in I^{k+}} y_{ki} + \sum_{i \in I^{k-}} w_{ki} \right] = 0 \qquad \text{for all } k \in K$$

$$\sum_{i \in I} v_i \geq 0 \qquad q \left[\sum_{i \in I} v_i \right] = 0$$

The constraint qualification is satisfied for each of the individual optimisation problems. Therefore, if the consumption plan $(\mathbf{A}^{\#} \mid \mathbf{B}^{\#})$ is a solution for the representative consumer, and the augmented production plan $(\mathbf{a}_i^{\#} \mid \mathbf{b}_i^{\#} \mid \mathbf{w}_i^{\#} \mid \mathbf{y}_i^{\#} \mid v_i^{\#})$ is a solution for agent for $i \in I$, then there exists a vector of shadow prices $(\mu^{\#} \mid \lambda^{\#} \mid \varpi^{\#})$, such that $(\mathbf{A}^{\#} \mid \mathbf{B}^{\#})$, $(\tilde{\mathbf{a}}^{\#} \mid \tilde{\mathbf{b}}^{\#} \mid \tilde{\mathbf{w}}^{\#} \mid \tilde{\mathbf{y}}^{\#} \mid \tilde{\mathbf{v}}^{\#})$, $(\mu^{\#} \mid \lambda^{\#} \mid \varpi^{\#})$ and $(\mathbf{p} \mid q)$ satisfy the Kuhn-Tucker conditions ME1-ME10 and ME15 given in Table A3.2.2 in appendix A3.1. [19]

Since the objective function of each optimization problem is concave, and the constraint functions are convex, these conditions are also sufficient, and therefore if there exist vectors $(\mathbf{A}^{\#} \mid \mathbf{B}^{\#})$ for the representative consumer, and $(\mathbf{a}_i^{\#} \mid \mathbf{b}_i^{\#} \mid \mathbf{w}_i^{\#} \mid \mathbf{y}_i^{\#} \mid v_i^{\#})$ for each agent $i \in I$ satisfying the respective first-order conditions, then these are global maxima. Consequently, the Kuhn-Tucker conditions ME1-ME15 are necessary and sufficient for the existence of a market equilibrium.

As with the regulated aggregate cost minimum, it is possible to show that for any market equilibrium, the system production plan and the utility consumption plan are both strictly non-zero; the budget constraint and all production constraints hold with equality; and all commodity market clear. That is:

[19] The shadow price $\varpi^{\#}$ relates to the representative consumer's budget constraint.

Proposition 3.3

If the allocation $(\mathbf{A}^{\#} \mid \mathbf{B}^{\#} \mid \tilde{\mathbf{a}}^{\#} \mid \tilde{\mathbf{b}}^{\#} \mid \tilde{\mathbf{w}}^{\#} \mid \tilde{\mathbf{y}}^{\#})$, and the price vectors $(\mu^{\#} \mid \lambda^{\#} \mid \varpi^{\#})$ and $(\mathbf{p} \mid \mathrm{q})$ satisfy conditions ME1-ME15, and assumptions A1-A18 are satisfied,

then (i) $(\mathbf{B}^{\#} \mid \tilde{\mathbf{a}}^{\#} \mid \tilde{\mathbf{w}}^{\#}) \ll \mathbf{0}$ and $(\tilde{\mathbf{b}}^{\#} \mid \tilde{\mathbf{y}}^{\#}) \gg \mathbf{0}$

(ii) $\mu_i^{\#} > 0 \qquad b_{li}^{\#} + f^i(\mathbf{a}_i^{\#} \mid \mathbf{w}_i^{\#}) = 0 \qquad$ for all $l \in \mathrm{L}$ and $i \in \mathrm{I}^{l+}$

$\mu_i^{\#} > 0 \qquad y_{ki}^{\#} + f^i(\mathbf{a}_i^{\#} \mid \mathbf{w}_i^{\#}) = 0 \qquad$ for all $k \in \mathrm{K}$ and $i \in \mathrm{I}^{k+}$

(iii) $\varpi^{\#} > 0 \qquad \Pi^{\#} + \sum_{j \in \mathrm{J}} \mathrm{p}_j (A_j + \mathrm{A}_j^0) + \sum_{l \in \mathrm{L}} \mathrm{p}_l B_l = 0$

(iv) $\mathrm{p}_j^{\#} > 0 \qquad \sum_{i \in \mathrm{I}^{j-}} \mathrm{a}_{ji}^{\#} + \mathrm{A}_j^0 + \mathrm{A}_j^{\#} = 0 \qquad$ for all $j \in \mathrm{J}$

$\mathrm{p}_l^{\#} > 0 \qquad \sum_{i \in \mathrm{I}^{l+}} \mathrm{b}_{li}^{\#} + \mathrm{B}_l^{\#} = 0 \qquad$ for all $l \in \mathrm{L}$

$\mathrm{p}_k^{\#} > 0 \qquad \sum_{i \in \mathrm{I}^{k+}} \mathrm{y}_{ki}^{\#} + \sum_{i \in \mathrm{I}^{k-}} \mathrm{w}_{ki}^{\#} = 0 \qquad$ for all $k \in \mathrm{K}$

Proof: see Appendix A3.3

As usual with general equilibrium models, in order to identify the price vector $(\mathbf{p} \mid \mathrm{q})$ it is necessary to impose a normalization condition. In this case, an obvious condition is that $\mathrm{p}_l = 1$ for some $l \in \mathrm{L}$. Since $\mathbf{B}^{\#} \gg 0$, it follows directly from condition ME2 that $\varpi^{\#} = 1$, and hence that $\mathrm{p}_l = 1$ for all $l \in \mathrm{L}$.

Existence and cost efficiency of market equilibrium

For any given vectors of design parameters $\theta \in \Theta$ and $\varepsilon \in \mathrm{E}$, let $\Psi^{\#}(\theta \mid \varepsilon)$ be the set of allocations for which there exist a performance credit plan, and market and shadow price vectors, such that the necessary and sufficient conditions for a market equilibrium

ME1-ME15 are satisfied. Then it can be shown that each of these sets contains all of the possible solutions to the regulated cost problem (RC). That is:

Proposition 3.4

$\Psi^* \subset \Psi^{\#}(\theta \mid \varepsilon)$ for all $\theta \in \Theta$ and $\varepsilon \in E$.

Proof: see Appendix A3.4

Furthermore, it can also be shown that each of these sets is contained in the set of solutions to the regulated cost minimum problem. That is:

Proposition 3.5

For all $\theta \in \Theta$ and $\varepsilon \in E$, $\Psi^{\#}(\theta \mid \varepsilon) \subset \Psi^*$.

Proof: see Appendix A3.5

Thus, any regulated aggregate cost minimum can be attained as a market equilibrium for any choice of values for the design parameters θ and ε. Furthermore, for any choice of values for the design parameters, any resultant market equilibrium is a regulated cost minimum. The cost efficiency of the market equilibrium can be seen if one eliminates the shadow prices of the production constraint and the individual performance rules to give the following four sets of marginal product conditions for agent $i \in I$:

$$f_j^{i\#} = p_j \qquad \text{for all } l \in L^{i+}, j \in J^{i-} \quad \ldots (3.10.a)$$

$$f_k^{i\#} = p_k + q\left(\sigma_k + \sum_{m \in M^{i-}} \chi_m h_k^{mi\#}\right) \qquad \text{for all } l \in L^{i+}, k \in K^{i-} \quad \ldots (3.10.b)$$

$$(p_{k'} + q\,\rho_{k'})f_k^{i\#} = p_j \qquad \text{for all } k' \in K^{i+}, j \in J^{i-} \quad \ldots (3.10.c)$$

$$(p_{k'} + q\,\rho_{k'})f_k^{i\,\#} \;=\; p_k + q\left(\sigma_k + \sum_{m\in M^{i-}} \chi_m h_k^{mi\,\#}\right) \qquad \text{for all } k' \in K^{i+},\ k \in K^{i-} \quad \ldots (3.10.d)$$

where $\quad f^i_k{}^\# = f^i_k(\mathbf{a}_i^\# | \mathbf{w}_i^\#) \quad$ and $\quad h^{mi}{}_k{}^\# = h^{mi}{}_k(\mathbf{w}_i^\#)$

Again, only two of these sets of conditions will apply to an individual agent. Recalling that $\rho_k = \alpha_k - \theta_k(\alpha_k + \beta_k)$ and that $\sigma_k = \beta_k - \theta_k(\alpha_k + \beta_k)$, it is clear that any production plan $(\tilde{\mathbf{a}}^\# \mid \tilde{\mathbf{b}}^\# \mid \tilde{\mathbf{w}}^\# \mid \tilde{\mathbf{y}}^\#)$ which satisfies these conditions will also satisfy the corresponding marginal product conditions (3.8) for a regulated cost minimum, with $\lambda^* = q$, $\phi_j^* = p_j$, and $\eta_k^* = p_k - \theta_k(\alpha_k + \beta_k)q$. Furthermore, as was demonstrated in Chapter 2, since the markets for all endogenous commodities clear, satisfaction of the individual performance rules ensures that the aggregate performance rule is also satisfied.

Two corollaries follow directly from Propositions 3.1, 3.4 and 3.5.

Corollary 3.1

$$\Psi^\#(\theta^1 | \varepsilon^1) \;\equiv\; \Psi^\#(\theta^2 | \varepsilon^2) \;\equiv\; \Psi^\# \;\equiv\; \Psi^* \qquad \text{for all } \theta^1, \theta^2 \in \Theta, \text{ and } \varepsilon^1, \varepsilon^2 \in E.$$

Corollary 3.2

$\Psi^\#$ is non-empty and compact.

Thus, a market equilibrium is guaranteed to exist for any values of the design parameters θ and ε. Furthermore, any market equilibrium that is attained under one vector of values for the design parameters, can be attained under any other vector of values.

Market prices, performance credits and distribution of aggregate cost

It follows directly from Corollary 3.2 that the aggregate cost is unaffected by the choice of values for the design parameters θ and ε. However, of more interest is the impact of these parameters on the distribution of the aggregate cost between the individual agents within the production system. Since the equilibrium quantities of the agents' inputs and outputs are unaffected by the values of the design parameters, any distributional effect must be driven either by the impact of these parameters on the equilibrium prices of the market commodities and / or performance credits, or by their impact on the quantities of performance credits transferred between agents.

Starting with the potential impact on prices, it can be shown that while the market prices of primary and utility commodities, and performance credits are independent of the values of the two design parameters, the price of any secondary commodity $k \in \mathrm{K}$ that has a non-zero performance rule parameter (i.e. either $\alpha_k \neq 0$ or $\beta_k \neq 0$) depends on the value that is chosen for the assignment parameter θ_k. That is:

Proposition 3.6

For any $(\mathbf{A}^{\#} \mid \mathbf{B}^{\#} \mid \tilde{\mathbf{a}}^{\#} \mid \tilde{\mathbf{b}}^{\#} \mid \tilde{\mathbf{w}}^{\#} \mid \tilde{\mathbf{y}}^{\#}) \in \Psi^{\#}$

Let $(\mathbf{p}^{(m)} \mid \mathrm{q}^{(m)})$ and $(\mu^{\#(m)} \mid \lambda^{\#(m)})$ denote the market and shadow price vectors that support the market equilibrium for any two vectors of design parameters $\theta^{(m)} \in \Theta$ and $\varepsilon^{(m)} \in \mathrm{E}$ (where $m = 1,2$)

Then
$$\mu_i^{\#(1)} = \mu_i^{\#(2)} = \mu_i^{\#} \qquad \text{for all } i \in \mathrm{I}$$

$$\lambda_i^{\#(1)} = \lambda_i^{\#(2)} = \lambda_i^{\#} \qquad \text{for all } i \in \mathrm{I}$$

$$\mathrm{q}^{(1)} = \mathrm{q}^{(2)} = \mathrm{q}$$

$$p_j^{(1)} \quad = \quad p_j^{(2)} \quad = \quad p_j \qquad \text{for all } j \in \mathrm{J}$$

$$p_l^{(1)} \quad = \quad p_l^{(2)} \quad = \quad 1 \qquad \text{for all } l \in \mathrm{L}$$

$$p_k^{(1)} - p_k^{(2)} \quad = \quad q\,(\alpha_k + \beta_k)\,(\theta_k^{(1)} - \theta_k^{(2)}) \qquad \text{for all } k \in \mathrm{K}$$

Proof: *see Appendix A3.6*

The impact of a change to the value of the assignment parameter θ_k for a particular commodity $k \in \mathrm{K}$ depends on the values of rule parameters for that commodity. For example, if $\alpha_k > 0$ (and hence $\beta_k = 0$) then an increase in the value of θ_k increases the price of the commodity by an amount equal to $q\ \alpha_k\ \Delta\theta_k$. Hence, *ceteris paribus* there is an increase in the gross economic benefit of all those agents that produce the commodity, and a reduction in the benefit of those agents that use it as an input.

However, as was noted in Chapter 2, for $\alpha_k > 0$ an increase in the value of θ_k reduces the producer's share of the endogenous property rights that result from any transactions involving that commodity.[20] Thus, for each unit of the commodity that an agent produces, the revenue that it receives from the sale of performance credits is reduced by an amount equal to $q\ \alpha_k\ \Delta\theta_k$ – i.e. exactly the same amount as the increase in the price of the commodity. Consequently, while the nominal price of the commodity is affected by changes to the value of the assignment parameter, the *total unit benefit* received by the producers of the commodity is not. That is:

$$p_k^{(1)} + q\,\alpha_k(1 - \theta_k^{(1)}) \quad = \quad p_k^{(2)} + q\,\alpha_k(1 - \theta_k^{(2)}) \qquad \text{for all } k \in \mathrm{K}$$

Similarly, the *total unit cost* of the commodity to those agents that use it as an input is also unaffected by the value of the assignment parameter. That is

[20] See Chapter 2 for a discussion of obligations and property rights,

$$p_k^{(1)} - q\,\alpha_k\,\theta_k^{(1)} \quad = \quad p_k^{(2)} - q\,\alpha_k\,\theta_k^{(2)} \qquad \text{for all } k \in K$$

It follows directly that the choice of values for the assignment parameters has no impact on the distribution of the total gross economic benefit between agents – i.e. after taking into account the financial values of the transfers of performance credits. That is:

Proposition 3.7

For any $(\mathbf{A}^{\#} \mid \mathbf{B}^{\#} \mid \tilde{\mathbf{a}}^{\#} \mid \tilde{\mathbf{b}}^{\#} \mid \tilde{\mathbf{w}}^{\#} \mid \tilde{\mathbf{y}}^{\#}) \in \Psi^{\#}$

Let $\Pi_i^{\#(m,n)}$ denote the gross economic benefit for agent $i \in I$ in the market equilibrium, for given vectors of design parameters $\theta^{(m)} \in \Theta$ ($m = 1, 2$) and $\varepsilon^{(n)} \in E$ ($n = 1, 2$)

Then $\Pi_i^{\#(1,n)} - \Pi_i^{\#(2,n)} = 0$

$$\Pi_i^{\#(m,1)} - \Pi_i^{\#(m,2)} = q\,(\varepsilon_i^{(1)} - \varepsilon_i^{(2)})$$

for all $i \in I$

Proof: see Appendix A3.7

In contrast, the choice of values for the performance adjustment factors (ε) does affect the distribution of the aggregate cost between the agents. Increasing the value of the factor for a particular agent $i \in I$ increases its total gross economic benefit by an amount equal to the magnitude of the change, multiplied by the market price of performance credits. This of course is not surprising. As was discussed in Chapter 2, changing the values of these adjustment factors results in a reallocation of property rights between agents; changing the number of performance credits that each agent can sell, or that it must acquire. Since this has no impact on the market price of credits, this has the effect of making lump-sum transfers between the agents. As such, varying the values of the factors is equivalent to varying the initial allocation of permits in a traditional "cap and trade" scheme for an absolute limit.

Interpretation of market price of performance credits (q)

The necessary condition ME7 for a market equilibrium requires that $q = \lambda_i^{\#}$ for all $i \in$ I. Thus, the price of performance credits is equal to shadow price of the individual performance rule for each agent. However, in the proof of proposition 3.5 it is demonstrated that $\lambda_i^{\#} = \lambda^*$ for all $i \in$ I. Consequently, the price also provides a measure of the increase in aggregate benefit that would arise from a relaxation of the aggregate performance rule – via marginal changes to the target value parameters L and r. This, of course, is directly analogous to the interpretation of the price of permits in a "cap & trade" scheme (where r = 0).

It is interesting to explore the first of these two interpretations further. This is facilitated by decomposing the regulated benefit maximization problem (RB_i) of each agent $i \in$ I into two steps, i.e.

Maximize $V^i(v_i) + q\, v_i$

where $\quad V^i(v_i) \quad = \quad$ Maximize $\displaystyle\sum_{j \in J^{i-}} p_j\, a_{ji} + \sum_{l \in L^{i+}} p_l\, b_{li} + \sum_{k \in K^{i-}} p_k\, w_{ki} + \sum_{k \in K^{i+}} p_k\, y_{ki}$

subject to $(\boldsymbol{a}_i \mid \boldsymbol{b}_i \mid \boldsymbol{w}_i \mid \boldsymbol{y}_i \mid v_i) \in G^i \cap R^i$

In the first step, the agent chooses the production plan that maximizes its gross economic benefit for a given quantity of performance credits transferred. In the second step, it chooses the transfer quantity that maximizes its total benefit – including the financial value of the transfer. It follows that $\lambda_i^{\#} = -V^{i\prime}(v_i^{\#}) = q$. That is, for each agent, the Lagrange multiplier associated with their individual performance constraint is equal to the reduction in gross economic benefit arising from an incremental increase in the quantity of performance credits transferred (i.e. an increase in the number sold, or a decrease in the number purchased). This in turn is equal to the price of performance credits.

As with the interpretation of the Lagrange multiplier (λ^*) in the aggregate cost minimization problem, in order to interpret the price of performance credits in terms of a change in the quantity of some commodity (or commodities) used by the agent, it is necessary to make some assumptions regarding the form of the individual performance rule. Using the same aggregate performance rule that was adopted for the interpretation of λ^*, let us assume that:

$I^{k+} \subset I^{M-}$ for all $k \in K$ with $a_k > 0$

$\theta_k = 0$ for all $k \in K$

$\varepsilon_i = -L / I$ for all $i \in I \backslash I^{M-}$

That is, all agents that produce secondary commodities that are included in the performance rule also emit at least one of the pollutants; all of the endogenous property rights arising from the sale of a secondary commodity are assigned to the producer; and the constant terms in the individual performance rules of non-emitting agents are all set to zero. Under these assumptions, all non-emitting agents $i \in I \backslash I^{M-}$ are inactive, with $v_i^{\#} = 0$ in the equilibrium. The individual performance rules for the active agents $i \in I^{M-}$ are:

$$r\, Y_i + Z_i + L_i \geq 0$$

where:

$Z_i = \sum_{m \in M^{i-}} c_m h^{mi}(\boldsymbol{w}_i) < 0$ denotes total emissions of agent $i \in I^{M-}$

$$Y_i \quad = \quad \sum_{k \in K^{i+}} a_k\, y_{ki} \quad > \quad 0 \qquad \text{denotes total output of agent } i \in I^{M-}$$

$$L_i \quad = \quad \delta_i - v_i^{\#}$$

Thus, an increase in the value of $v_i^{\#}$ is equivalent to a decrease in the value of the "individual" regulatory parameter L_i. The analysis of the impact of a change in the value of $v_i^{\#}$ is therefore completely equivalent to the first case considered in the interpretation of λ^* (i.e. keeping the value of r fixed, and changing the value of L), with the addition of agent subscripts. Therefore, it follows that:

$$q \quad = \quad \lambda^{\#}_i \quad = \quad - \frac{1}{(1+\xi_i^{\#})} \frac{d\,\Pi_i^{\#}}{d\,Z_i^{\#}} \qquad \text{where} \quad \xi_i^{\#} = r\,\frac{d\,Y_i^{\#}}{d\,Z_i^{\#}}$$

Two points are clear from this. First, the price of performance credits is not equal to the individual marginal cost of abatement of any agent, unless the regulatory parameter $r = 0$. This of course is analogous to the divergence between λ^* and the aggregate marginal cost of abatement. Second, for the marginal cost of abatement to be equalised across agents, it is necessary that the value of $\xi_i^{\#}$ is the same for all agents. However, since the relative change in output and emissions will depend on the specification of the agent's production function and its emission function, there is no reason to expect that this will be the case. Consequently, the marginal cost of abatement will not, in general, be equalised across agents. That is:

$$\frac{d\,\Pi_i^{\#}}{d\,Z_i^{\#}} \quad \neq \quad \frac{d\,\Pi_{i'}^{\#}}{d\,Z_{i'}^{\#}} \qquad \text{for all } i, i' \in I^{M-}$$

In particular, if the regulatory target takes the form of a standard for specific emissions (i.e. $r > 0$ and $L = 0$)[21], then while the marginal cost of achieving the (relative) standard is equalised across agents, the marginal cost of (absolute) emissions reduction is not.

[21] Note that this does not necessarily imply that $L_i = 0$ for any $i \in I^{M-}$.

Equivalence to a tax-subsidy scheme

As was noted in Chapter 1, there are a number of different types of mechanism that can be used to implement a regulatory target. Of course, not all of these will necessarily minimize the cost of meeting the target. However, Baumol & Oates (1988) have shown that in the case of an absolute limit for aggregate emissions, a (market-based) tradable permit system and a (price-based) emissions tax are equivalent, at least in terms of cost efficiency. If the tax is set at a value that is the same as the market price of emission permits (i.e. $\tau = q$), then the real outcomes under the two mechanisms will be identical. This conclusion can be extended to any regulatory target that can be represented by an aggregate performance rule.

It is clear from the marginal product conditions (3.10) that the market equilibrium – and hence (by Proposition 3.5) the aggregate cost minimum – could also be induced by the following tax-subsidy scheme for market and non-market commodities:[22]

$\tau_k^+ = q\,(-\rho_k)$	a tax / subsidy applied to the output price of endogenous secondary commodity $k \in K$
$\tau_k^- = q\,\sigma_k$	a tax / subsidy applied to the input price of endogenous secondary commodity $k \in K$
$\tau_m^- = q\,\chi_m$	a tax applied to the (zero) input price of exogenous non-market commodity $m \in M$

While the value of τ_m^- is non-negative by definition, the values of τ_k^+ and τ_k^- may both be positive or negative.[23] A positive value indicates that a tax is imposed, while a

[22] If the aggregate performance rule contains no secondary commodities and only one non-market environmental service, then this reduces to the simple tax scheme analysed by Baumol & Oates (1988).

[23] The "+" and "–" superscripts indicate whether the tax / subsidy is applied to an output price or an input price.

negative value indicates that a subsidy is paid. It follows that a subsidy must be paid on any secondary commodity that has a positive output parameter (i.e. $\alpha_k > 0$) or a negative input parameter (i.e. $\beta_k < 0$). For example, in the case of a relative standard for specific emissions, a subsidy must be paid on sales of all output commodities.[24]

Since the values of ρ_k and σ_k both depend on the value of the design parameter θ_k, there are a range of different tax-subsidy schemes that can support the aggregate cost minimum. In particular, it is possible for both the input price and the output price of a secondary commodity to be taxed, or for both to be subsidized. However, the sum of the two taxes / subsidies is constant, i.e.

$$\tau_k^+ + \tau_k^- \quad = \quad q\,(\sigma_k - \rho_k) \quad = \quad q\,(\beta_k - \alpha_k) \qquad \text{for all } k \in \mathrm{K}$$

Consequently, the different schemes merely reflect different assignments (or shares) of a fixed set of taxes and subsidies between the buyers and sellers of the secondary commodities. This of course is completely analogous to the effect of the parameters on the assignment of property rights and obligations in the trading scheme.

The net tax revenue received by the government under the scheme is:

$$\begin{aligned}
\mathrm{T} \quad &= \quad \sum_{k\in\mathrm{K}} \tau_k^+ \sum_{i\in \mathrm{I}^{k+}} y_{ki}^{\#} \;+\; \sum_{k\in\mathrm{K}} \tau_k^- \sum_{i\in \mathrm{I}^{k-}} (-\,w_{ki}^{\#}) \;+\; \sum_{m\in\mathrm{M}} \tau_m^- \sum_{i\in \mathrm{I}^{m-}} (-\,h^{mi}(\mathbf{w}_i^{\#})) \\
&= \quad -\;q \sum_{k\in\mathrm{K}} \rho_k \sum_{i\in \mathrm{I}^{k+}} y_{ki}^{\#} \;-\; q \sum_{k\in\mathrm{K}} \sigma_k \sum_{i\in \mathrm{I}^{k-}} w_{ki}^{\#} \;-\; q \sum_{m\in\mathrm{M}} \chi_m \sum_{i\in \mathrm{I}^{m-}} h^{mi}(\mathbf{w}_i^{\#}) \\
&= \quad -\;q \left[\sum_{k\in\mathrm{K}} \alpha_k \sum_{i\in \mathrm{I}^{k+}} y_{ki}^{\#} \;+\; \sum_{k\in\mathrm{K}} \beta_k \sum_{i\in \mathrm{I}^{k-}} w_{ki}^{\#} \;+\; \sum_{m\in\mathrm{M}} \chi_m \sum_{i\in \mathrm{I}^{m-}} h^{mi}(w_{ki}^{\#}) \right]
\end{aligned}$$

[24] This is consistent with the observation of Fischer (2001) and Gielen *et al* (2002) that setting a relative standard for specific emissions has the effect of subsidizing output (see Chapter 1, section 1.2)

$$+ \quad q \left[\sum_{k \in K} \theta_k (\alpha_k + \beta_k) \left[\sum_{i \in I^{k+}} y^{\#}_{ki} + \sum_{i \in I^{k-}} w^{\#}_{ki} \right] \right]$$

$$= \quad q\,\delta$$

Thus, the net revenue received by the government is equal to the market equilibrium price of performance credits (q), multiplied by the value of the constant term in the aggregate performance rule (δ). This implies that for any "pure" relative performance standard (i.e. where the constant term $\delta = 0$), the supporting tax-subsidy scheme is revenue neutral.[25]

It should be noted that the subsidies play a different role to that proposed by Pezzey (1992). In his tax-subsidy scheme – which applies to the traditional case of an absolute limit, the subsidies are included in order to change the distributional impacts of the mechanism, to reflect alternative assumptions regarding property rights over the environment. The efficiency of the mechanism does not rely on the inclusion of any subsidies, and the real outcome is completely independent of the values of the subsidies. In contrast, in the case of a relative standard (when either $\alpha_k > 0$ or $\beta_k < 0$ for some $k \in K$), a subsidy is required in order to achieve an efficient outcome. Of course, when $\delta > 0$ (and hence the scheme is not revenue neutral), there is nothing to preclude the inclusion of other "distributional" subsidies as envisaged by Pezzey.

3.4 Summary

The cost efficiency and the distributional flexibility of the performance-based credit trading mechanism have been analysed using a short-run, static, general equilibrium

[25] This was noted by Gielen *et al* (2002) for the particular case of a relative standard for the specific emissions rate of greenhouse gases.

model. While the analysis rests on a number of explicit assumptions regarding the structure of the production system; the technical properties of the production and emission functions; the values of the performance rule parameters; none of these are particularly restrictive. Under these assumptions, the existence of a finite solution to the regulated aggregate cost minimization problem is assured for any aggregate performance rule that is technically feasible.

Cost efficiency requires that a wedge be driven between the input price and output price of each secondary commodity that is included (either directly or indirectly) in the aggregate performance rule. Depending on the specific values of the performance rule parameters the input price of a particular commodity may be higher than the output price, or it may be lower. For example, if the aggregate performance rule represents a relative standard for vehicle fuel efficiency then the output price is higher than the input price for models with better fuel efficiency than the standard, and lower models with worse fuel efficiency.

The shadow value of the aggregate performance rule provides a measure of the increase in aggregate gross economic benefit that would arise from a relaxation of the rule – via marginal changes to the regulatory control variables L and r. For the case of a relative performance standard for emissions (where L = 0) this is equal to the marginal cost of (absolute) emissions reduction multiplied by a factor that reflects the relative responsiveness of aggregate output and emissions to the changes in the target rate (r).

Any regulated aggregate cost minimum can be attained as a market equilibrium with performance credits, for any choice of values for assignment parameters (θ) and the performance adjustment factors (ε). Furthermore, for any choice of values for the regulatory design parameters, any resultant market equilibrium is a regulated aggregate cost minimum. Consequently, a market equilibrium is guaranteed to exist for any values of the design parameters, and the equilibrium quantities of commodities produced and used by individual agents are all independent of parameter values.

The market prices of all primary and utility commodities are all independent of the values of the regulatory design parameters, as is the price of performance credits. However, the price of any secondary commodity that has a non-zero performance rule parameter depends on the value of the assignment parameter (θ_k) for that commodity.

The impact of a change to the value of the assignment parameter for a particular commodity depends on the values of rule parameters for that commodity. If it causes the price to rise, then there is an increase in the gross economic benefit of any agent that produces the commodity, and a reduction in the benefit of any agent that uses it as an input. If it causes the price to fall, then the impacts are reversed. However, in each case the change is offset by an equal and opposite change in the total value of the performance credits that the agent must acquire, or that it can sell. Consequently, while the nominal price of the commodity is affected by changes to the value of the assignment parameter, the *total unit benefit* received by the producers, and the *total unit cost* paid by the users, are both unaffected.

It follows directly that the choice of values for the assignment parameters has no impact on the distribution of the total gross economic benefit between agents – i.e. after taking into account the financial values of the transfers of performance credits. In contrast, the choice of values for the performance adjustment factors (ε) does affect the distribution of the aggregate cost between the agents. As was discussed in Chapter 2, changing the values of these adjustment factors results in a reallocation of property rights between agents; changing the number of performance credits that each agent can sell, or that it must acquire. Since this has no impact on the market price of performance credits, this has the effect of making lump-sum transfers between the agents.

The shadow values of the agents' individual performance rule constraints represent their respective marginal WTP / WTA values for performance credits; with these all being equalised at market price in the equilibrium. Again for the case of a relative performance standard for emissions, this is equal to each agent's marginal cost of abatement, multiplied by a factor that reflects the relative responsiveness of its individual output and emissions. Thus, while the marginal cost of achieving the (relative) performance standard is equalised across agents, the marginal cost of (absolute) emissions reduction is only equalised if the value of this factor is the same for all agents – and there is no *a priori* reason why this should be the case.

This analysis has focussed on the use of a market-based mechanism to implement any regulatory target that can be represented as an aggregate performance rule. However, it is clear form the analysis that the market equilibrium – and hence the aggregate cost minimum – could also be induced by a tax-subsidy scheme for market and non-market commodities. Under such a scheme, the net revenue received by the government is equal to the market equilibrium price of performance credits multiplied by the value of the constant term in the aggregate performance rule. This implies that for any "pure" relative performance standard, where the constant term is set equal to zero, the supporting tax-subsidy scheme is revenue neutral.

While all of these results have been derived in the context of a general equilibrium, they can be transferred directly into a partial equilibrium setting. The shadow / market prices of all utility commodities are constant by construction. Furthermore, if the total endowment of each primary resource is sufficiently large, then the prices of all primary resources are also constant. In this special case, the model can interpreted as a partial equilibrium model for the basket of all secondary commodities. By direct analogy, the results also apply to a production system for a subset of "endogenous" secondary commodities, where the prices of all other "exogenous" commodities (which may include primary and utility commodities) are constant.[26] It is this partial equilibrium framework that is used in chapters 5 and 6 to analyse the impacts of performance-based credit trading for two specific policy applications.

[26] In this case, the assumption of quasi-linearity in direct consumption of primary resources is replaced by two (implicit) *ad hoc* assumptions. First, that the production system for the subset of "endogenous" secondary commodities under consideration accounts for only a small proportion of the total supply of, or the demand for, each of the remaining "exogenous" market commodities. Second, that consumer expenditure on any endogenous commodity – or on all exogenous commodities derived from it – represents only a small proportion of total consumer expenditure.

Appendix A3.1 Kuhn-Tucker conditions

Table A3.1.1: Kuhn-Tucker conditions for a regulated aggregate cost minimum

RC1	$A_j \le 0$	$\phi_j - 1$	≥ 0	$A_j(\phi_j - 1)$	$= 0$	$\forall j \in J$
RC2	$B_l \le 0$	$\gamma_l - 1$	≥ 0	$B_l(\gamma_l - 1)$	$= 0$	$\forall l \in L$
RC3	$a_{ji} \le 0$	$\phi_j - \mu_i f^i_j$	≥ 0	$a_{ji}(\phi_j - \mu_i f^i_j)$	$= 0$	$\forall j \in J,\ i \in I^{j-}$
RC4	$b_{li} \ge 0$	$\gamma_l - \mu_i$	≤ 0	$b_{li}(\gamma_l - \mu_i)$	$= 0$	$\forall l \in L,\ i \in I^{l-}$
RC5	$w_{ki} \le 0$	$\eta_k + \lambda(\beta_k + \sum_{m \in M^{i-}} \chi_m h_k^{mi}) - \mu_i f^i_k$	≥ 0	$w_{ki}(\eta_k + \lambda(\beta_k + \sum_{m \in M^{i-}} \chi_m h_k^{mi}) - \mu_i f^i_k)$	$= 0$	$\forall k \in K,\ i \in I^{k-}$
RC6	$y_{ki} \ge 0$	$\eta_k + \lambda \alpha_k - \mu_i$	≤ 0	$y_{ki}(\eta_k + \lambda \alpha_k - \mu_i)$	$= 0$	$\forall k \in K,\ i \in I^{k+}$
RC7	$\mu_i \ge 0$	$b_{li} + f^i(\boldsymbol{a}_i \mid \boldsymbol{w}_i)$	≤ 0	$\mu_i(b_{li} + f^i(\boldsymbol{a}_i \mid \boldsymbol{w}_i))$	$= 0$	$\forall l \in L,\ i \in I^{l+}$
RC8	$\mu_i \ge 0$	$y_{ki} + f^i(\boldsymbol{a}_i \mid \boldsymbol{w}_i)$	≤ 0	$\mu_i(y_{ki} + f^i(\boldsymbol{a}_i \mid \boldsymbol{w}_i))$	$= 0$	$\forall k \in K,\ i \in I^{k+}$
RC9	$\phi_j \ge 0$	$\sum_{i \in I^{j-}} a_{ji} + A_j + A_j^0$	≥ 0	$\phi_j(\sum_{i \in I^{j-}} a_{ji} + A_j + A_j^0)$	$= 0$	$\forall j \in J$
RC10	$\gamma_l \ge 0$	$\sum_{i \in I^{l+}} b_{li} + B_l$	≥ 0	$\gamma_l(\sum_{i \in I^{l+}} b_{li} + B_l)$	$= 0$	$\forall l \in L$
RC11	$\eta_k \ge 0$	$\sum_{i \in I^{k+}} y_{ki} + \sum_{i \in I^{k-}} w_{ki}$	≥ 0	$\eta_k(\sum_{i \in I^{k+}} y_{ki} + \sum_{i \in I^{k-}} w_{ki})$	$= 0$	$\forall k \in K$

Table A3.1.1: Kuhn-Tucker conditions for a regulated aggregate cost minimum (cont'd)

RC12	$\lambda \geq 0$	$\sum_{k\in K}\alpha_k \sum_{i\in I^{k+}} y_{ki} + \sum_{k\in K}\beta_k \sum_{i\in I^{k-}} w_{ki} + \sum_{m\in M}\chi_m \sum_{i\in I^{m-}} h^{mi}(w_i) + \delta$	≥ 0	$\lambda\left(\sum_{k\in K}\alpha_k \sum_{i\in I^{k+}} y_{ki} + \sum_{k\in K}\beta_k \sum_{i\in I^{k-}} w_{ki} + \sum_{m\in M}\chi_m \sum_{i\in I^{m-}} h^{mi}(w_i) + \delta \right)$	$= 0$	

Table A3.1.2: Kuhn-Tucker conditions for a market equilibrium

ME1	$A_j \leq 0$	$\varpi p_j - 1$	≥ 0	$A_j(\varpi p_j - 1)$	$= 0$	$\forall j \in J$
ME2	$B_l \leq 0$	$\varpi p_l - 1$	≥ 0	$B_l(\varpi p_l - 1)$	$= 0$	$\forall l \in L$
ME3	$a_{ji} \leq 0$	$p_j - \mu_i f^i_j$	≥ 0	$a_{ji}(p_j - \mu_i g^i_j)$	$= 0$	$\forall i \in I,\ j \in J^{i-}$
ME4	$b_{li} \geq 0$	$p_l - \mu_i$	≤ 0	$b_{li}(p_l - \mu_i g^i_j)$	$= 0$	$\forall i \in I,\ l \in L^{i+}$
ME5	$w_{ki} \leq 0$	$p_k + \lambda_i(\sigma_k + \sum_{m\in M^{i-}} \chi_m h^{mi}_k) - \mu_i f^i_k$	≥ 0	$w_{ki}(p_k + \lambda_i(\sigma_k + \sum_{m\in M^{i-}} \chi_m h^{mi}_k) - \mu_i f^i_k)$	$= 0$	$\forall i \in I,\ k \in K^{i+}$
ME6	$y_{ki} \geq 0$	$p_k + \lambda_i \rho_k - \mu_i$	≤ 0	$y_{ki}(p_k + \lambda_i \rho_k - \mu_i)$	$= 0$	$\forall i \in I, k \in K^{i-}$
ME7	v_i	$q - \lambda_i$	$= 0$			$\forall i \in I$

Table A3.1.2: Kuhn-Tucker conditions for a market equilibrium (cont'd)

ME8	$\mu_i \geq 0$	$b_{li} + f^{l}(\boldsymbol{a}_i \mid \boldsymbol{w}_i) \leq 0$	$\mu_i \left(b_{li} + f^{l}(\boldsymbol{a}_i \mid \boldsymbol{w}_i) \right) = 0$	$\forall\, l \in L,\ i \in I^{l+}$
ME9	$\mu_i \geq 0$	$y_{ki} + f^{i}(\boldsymbol{a}_i \mid \boldsymbol{w}_i) \leq 0$	$\mu_i \left(y_{ki} + f^{i}(\boldsymbol{a}_i \mid \boldsymbol{w}_i) \right) = 0$	$\forall\, k \in K,\ i \in I^{k+}$
ME10	$\varpi \geq 0$	$\Pi + \sum_{j \in J} p_j (A_j^0 + A_j) + \sum_{l \in L} p_j B_l \geq 0$	$\varpi \left(\Pi + \sum_{j \in J} p_j (A_j^0 + A_j) + \sum_{l \in L} p_j B_l \right) = 0$	
ME11	$p_j \geq 0$	$\sum_{i \in I^{j-}} a_{ji} + A_j + A_j^0 \geq 0$	$p_j \left(\sum_{i \in I^{j-}} a_{ji} + A_j + A_j^0 \right) = 0$	$\forall\, j \in J$
ME12	$p_l \geq 0$	$\sum_{i \in I^{l+}} b_{li} + B_l \geq 0$	$p_l \left(\sum_{i \in I^{l+}} b_{li} + B_l \right) = 0$	$\forall\, l \in L$
ME13	$p_k \geq 0$	$\sum_{i \in I^{k+}} y_{ki} + \sum_{i \in I^{k-}} w_{ki} \geq 0$	$p_k \left(\sum_{i \in I^{k+}} y_{ki} + \sum_{i \in I^{k-}} w_{ki} \right) = 0$	$\forall\, k \in K$
ME14	$q \geq 0$	$\sum_{i \in I} v_i \geq 0$	$q \left(\sum_{i \in I} v_i \right) = 0$	
ME15	$\lambda_i \geq 0$	$\sum_{k \in K^{i+}} \rho_k y_{ki} + \sum_{k \in K^{i-}} \sigma_k w_{ki} + \sum_{m \in M^{i-}} \chi_m h^{mi}(\boldsymbol{w}_i) + \delta_i - v_i \geq 0$	$\lambda_i \left(\sum_{k \in K^{i+}} \rho_k y_{ki} + \sum_{k \in K^{i-}} \sigma_k w_{ki} + \sum_{m \in M^{i-}} \chi_m h^{mi}(\boldsymbol{w}_i) + \delta_i - v_i \right) = 0$	$\forall\, i \in I$

where $\rho_k = (1-\theta_k)\,\alpha_k - \theta_k\,\beta_k$ and $\sigma_k = (1-\theta_k)\,\beta_k - \theta_k\,\alpha_k$

Appendix A3.2 Proof of proposition 3.2

If the allocation $(\mathbf{A}^* \mid \mathbf{B}^* \mid \tilde{\mathbf{a}}^* \mid \tilde{\mathbf{b}}^* \mid \tilde{\mathbf{w}}^* \mid \tilde{\mathbf{y}}^*)$ and the shadow price vector $(\mu^* \mid \phi^* \mid \gamma^* \mid \eta^* \mid \lambda^*)$ satisfy conditions RC1-RC12, and assumptions A1-A18 are satisfied,

then (i) $(\mathbf{B}^* \mid \tilde{\mathbf{a}}^* \mid \tilde{\mathbf{w}}^*) \ll \mathbf{0}$ and $(\tilde{\mathbf{b}}^* \mid \tilde{\mathbf{y}}^*) \gg \mathbf{0}$

(ii) $\mu_i^* > 0 \qquad b_{li}^* + f^i(\mathbf{a}_i^* \mid \mathbf{w}_i^*) = 0 \qquad$ for all $l \in$ L and $i \in \mathrm{I}^{l+}$

$\mu_i^* > 0 \qquad y_{ki}^* + f^i(\mathbf{a}_i^* \mid \mathbf{w}_i^*) = 0 \qquad$ for all $k \in$ K and $i \in \mathrm{I}^{k+}$

(iii) $\phi_j^* > 0 \qquad \sum_{i \in \mathrm{I}^{j-}} a_{ji}^* + A_j^0 + A_j^* = 0 \qquad$ for all $j \in$ J

$\gamma_l^* > 0 \qquad \sum_{i \in \mathrm{I}^{l+}} b_{li}^* + B_l^* = 0 \qquad$ for all $l \in$ L

$\eta_k^* > 0 \qquad \sum_{i \in \mathrm{I}^{k+}} y_{ki}^* + \sum_{i \in \mathrm{I}^{k-}} w_{ki}^* = 0 \qquad$ for all $k \in$ K

Proof

Together, conditions RC1, RC2, RC4 and RC7 in Table A3.1.1 (Appendix 3.1) imply that:

$\phi_j^* > 0 \qquad \sum_{i \in \mathrm{I}^{j-}} a_{ji}^* + A_j^0 + A_j^* = 0 \qquad$ for all $j \in$ J

$\gamma_l^* > 0 \qquad \sum_{i \in \mathrm{I}^{l+}} b_{li}^* + B_l^* = 0 \qquad$ for all $l \in$ L

$\mu_i^* > 0 \qquad b_{li}^* + f^i(\mathbf{a}_i^* \mid \mathbf{w}_i^*) = 0 \qquad$ for all $l \in$ L and $i \in \mathrm{I}^{l+}$

Furthermore, it must be the case that $f^i(\mathbf{a}_i^* \mid \mathbf{w}_i^*) < 0$ for all $i \in$ I. Otherwise, under assumptions A7 and A10 it would be possible to increase aggregate welfare by diverting marginal quantities of secondary and primary inputs to any agent whose output is equal to zero. Consequently, it follows that $b_{li}^* > 0$ for all $l \in L$ and $i \in I^{l+}$, and hence that $B_l^* < 0$ for all $l \in L$.

The proof of the parts of the proposition relating the quantities and shadow prices of the secondary commodities, and the agents that produce these, is by induction. Define the following sequences of subsets of secondary commodities, and subsets of agents:

$$K^0 = \varnothing \qquad I^0 = I^{L+}$$

$$K^r = \Big[\bigcup_{i \in I^{r-1}} K^{i-} \Big] \cup K^{r-1} \qquad I^r = \bigcup_{k \in K^r} I^{k+} \qquad r = 1, 2, 3 \ldots$$

Thus, K^1 is the subset of secondary commodities used as inputs by the producers of utility commodities; I^1 is the subset of agents that that produce the secondary commodities used as inputs by the producers of utility commodities; K^2 is the set of secondary commodities used as inputs by the producers of the secondary commodities that are used as inputs by the producers of utility commodities; and so on.

By construction, $K^{r-1} \subset K^r$ and $I^{r-1} \subset I^r$ (i.e. the number of commodities / agents never declines at any iteration). Consequently, since there are a finite number of secondary commodities, the sequence $\{K^r\}$ must converge in a finite number of steps. That is, there exists a finite integer R such that $K^r = K^{r-1}$ for all $r \geq R$. Furthermore, it must be the case that $K^R = K$ (i.e. the algorithm identifies all secondary commodities). Otherwise, it follows that for all $k \in K \backslash K^R$ (i.e. all commodities not identified by the algorithm), $I^{k-} \subset \tilde{I} = I^{K+} \backslash I^R$, which contradicts assumption A4. It follows directly that the sequence $\{I^r\}$ also converges, with $I^R = I^{K+}$.

- <u>Step 1</u>: $\eta_k^* > 0$ for all $k \in K^1$, and $\mu_i^* > 0$ for all $i \in I^0$

For <u>any</u> $l \in L$, the necessary first order conditions for <u>any</u> agent $i \in I^{l+} \subset I^0$ (i.e. for any producer of that utility commodity), evaluated at the solution, are:

$$a_{ji}^* \leq 0 \qquad \phi_j^* - \mu_i^* f_j^{i*} \geq 0 \qquad a_{ji}^* [\phi_j^* - \mu_i^* f_j^{i*}] = 0 \qquad \dots (RC3)$$

$$b_{li}^* \geq 0 \qquad \gamma_l^* - \mu_i^* \leq 0 \qquad b_{li}^* [\gamma_l^* - \mu_i^*] = 0 \qquad \dots (RC4)$$

$$w_{ki}^* \leq 0 \qquad \eta_k^* + \lambda^* [\beta_k + \sum_{m \in M^{i-}} \chi_m h_k^{mi*}] - \mu_i^* f_k^{i*} \geq 0$$

$$w_{ki}^* [\eta_k^* + \lambda^* [\beta_k + \sum_{m \in M^{i-}} \chi_m h_k^{mi*}] - \mu_i^* f_k^{i*}] = 0 \qquad \dots (RC5)$$

Since $b_{li}^* > 0$, condition RC4 implies that $\mu_i^* = 1$. Therefore, since η_k^* and λ^* are finite, and the partial derivatives of the emissions function $h^{mi}{}_k$ are all finite (by assumption A14), it follows from conditions RC3 and RC5, and assumption A9, that $a_{ji}^* < 0$ for all $j \in J^{i-}$ and $w_{ki}^* < 0$ for all $k \in K^{i-}$. Therefore

$$\eta_k^* + \lambda^* [\beta_k + \sum_{m \in M^{i-}} \chi_m h_k^{mi*}] = \mu_i^* f_k^{i*} > 0 \quad \text{for all } k \in K^{i-}.$$

Given assumptions A16-A18, there are six possible cases that can arise:

(i) $\alpha_k = 0 \qquad \beta_k = 0 \qquad \sum_{m \in M^{i-}} \chi_m h_k^{mi*} = 0$

(ii) $\alpha_k \neq 0 \qquad \beta_k = 0 \qquad \sum_{m \in M^{i-}} \chi_m h_k^{mi*} = 0$

(iii) $\alpha_k = 0 \qquad \beta_k < 0 \qquad \sum_{m \in M^{i-}} \chi_m h_k^{mi*} = 0,$

(iv) $\alpha_k = 0$ $\beta_k = 0$ $\sum_{m \in M^{i-}} \chi_m h_k^{mi\,*} < 0,$

(v) $\alpha_k = 0$ $\beta_k > 0$ $\sum_{m \in M^{i-}} \chi_m h_k^{mi\,*} = 0,$

(vi) $\alpha_k = 0$ $\beta_k = 0$ $\sum_{m \in M^{i-}} \chi_m h_k^{mi\,*} > 0,$

In the first four cases, it follows directly that $\eta_k^* > 0$. In the last two cases however, one cannot rule out the possibility that $\eta_k^* = 0$. However, this would imply that for all $i \in I^{k+}$ (i.e. for all producers of that commodity), either $y_{ki}^* = 0$ or $\mu_i^* = 0$. But if $\mu_i^* = 0$ then it follows that $a_{ji}^* = 0$ for some essential input $j \in J^{i-}$. Hence, if $\eta_k^* = 0$, then $y_{ki}^* = 0$ for all $i \in I^{k+}$. However, this cannot be so since $w_{ki}^* < 0$ for the agent under consideration, and hence the total input quantity is strictly negative. Consequently, it must be the case that $\eta_k^* > 0$. Thus, $\eta_k^* > 0$ for all $k \in K^{i-}$.

Since this conclusion is independent of the choice of agent, it follows that $\mathbf{a}_i^* << \mathbf{0}$ and $\mathbf{w}_i^* << \mathbf{0}$ for all $i \in I^0$, and that $\eta_k^* > 0$ for all $k \in K^1$.

- Step 2: If $\eta_k^* > 0$ for all $k \in K^r$, then $\eta_k^* > 0$ for all $k \in K^{r+1}$, and $\mu_i^* > 0$ for all $I \in I^r$.

For any $k \in K^r$, the necessary first order conditions for any agent $i \in I^{k+} \subset I^r$ (i.e. for any producer of that secondary commodity), evaluated at the solution, are:

$$a_{ji}^* \leq 0 \qquad \phi_j^* - \mu_i^* f_j^{i\,*} \geq 0 \qquad a_{ji}^* [\, \phi_j^* - \mu_i^* f_j^{i\,*} \,] = 0 \qquad \ldots \text{(RC3)}$$

$$w_{k'i}^{*} \le 0 \qquad \eta_{k'}^{*} + \lambda^{*}[\,\beta_{k'} + \sum_{m\in M^{i-}} \chi_m h_{k'}^{mi\,*}\,] - \mu_i^{*} f_{k'}^{i\,*} \ge 0$$

$$w_{k'i}^{*}[\,\eta_{k'}^{*} + \lambda^{*}[\,\beta_{k'} + \sum_{m\in M^{i-}} \chi_m h_{k'}^{mi\,*}\,] - \mu_i^{*} f_{k'}^{i\,*}\,] = 0 \qquad \ldots\text{(RC5)}$$

$$y_{ki}^{*} \ge 0 \qquad \eta_k^{*} + \lambda^{*}\alpha_k - \mu_i^{*} \le 0 \qquad y_{ki}^{*}[\eta_k^{*} + \lambda^{*}\alpha_k - \mu_i^{*}] = 0 \qquad \ldots\text{(RC6)}$$

Condition RC6 implies that $\mu_i^{*} \ge \eta_k^{*} + \lambda^{*}\alpha_k$. However, since $\eta_k^{*} + \lambda^{*}\alpha_k > 0$, it follows that $\mu_i^{*} > 0$, and hence that $y_{ki}^{*} = -f^{i}(\mathbf{a}_i^{*} \mid \mathbf{w}_i^{*}) > 0$.[5] Then by the same argument that was employed in step 1, it must be the case that $\eta_{k'}^{*} > 0$ for all $k' \in K^{i-}$. Again, since the conclusion is independent of the choice of commodity and the choice of agent, it follows that $\mu_i^{*} > 0$, $y_{ki}^{*} = -f^{i}(\mathbf{a}_i^{*} \mid \mathbf{w}_i^{*}) > 0$, $\mathbf{a}_i^{*} \ll \mathbf{0}$ and $\mathbf{w}_i^{*} \ll \mathbf{0}$ for all $i \in I^{r}$, and that $\eta_k^{*} > 0$ for all $k \in K^{r+1}$.

By induction, $\mu_i^{*} > 0$, $y_{ki}^{*} = -f^{i}(\mathbf{a}_i^{*} \mid \mathbf{w}_i^{*}) > 0$, $\mathbf{a}_i^{*} \ll \mathbf{0}$ and $\mathbf{w}_i^{*} \ll \mathbf{0}$ for all $i \in I^{r}$, and $\eta_k^{*} > 0$ for all $k \in K^{r}$, for all values of $r \ge 1$(including $r = R$). Hence:

$$(\mathbf{B}^{*} \mid \tilde{\mathbf{a}}^{*} \mid \tilde{\mathbf{w}}^{*}) \ll \mathbf{0} \qquad \text{and} \qquad (\tilde{\mathbf{b}}^{*} \mid \tilde{\mathbf{y}}^{*}) \gg \mathbf{0}$$

$$\mu_i^{*} > 0 \qquad y_{ki}^{*} + f^{i}(\mathbf{a}_i^{*} \mid \mathbf{w}_i^{*}) = 0 \qquad \text{for all } k \in K \text{ and } i \in I^{k+}$$

$$\eta_k^{*} > 0 \qquad \sum_{i\in I^{k+}} y_{ki}^{*} + \sum_{i\in I^{k-}} w_{ki}^{*} = 0 \qquad \text{for all } k \in K$$

(QED)

---

[5] If $\eta_k^{*} + \lambda^{*}\alpha_k < 0$, then it follows directly from RC6 that $y_{ki}^{*} = 0$ for all $i \in I^{k+}$. If $\eta_k^{*} + \lambda^{*}\alpha_k = 0$, then either $y_{ki}^{*} = 0$ or $\mu_i^{*} = 0$ for all $i \in I^{k+}$. But if $\mu_i^{*} = 0$ then $a_{ji}^{*} = 0$ for all exogenous inputs $j \in J^{i-}$, at least one of which is essential. Hence, if $\eta_k^{*} + \lambda^{*}\alpha_k \le 0$ then $y_{ki}^{*} = 0$ for all $i \in I^{k+}$. But this cannot be so as $w_{ki}^{*} < 0$ for all $i \in I^{k-}$. Thus, it must be the case that $\eta_k^{*} + \lambda^{*}\alpha_k > 0$.

Appendix A3.3 Proof of proposition 3.3

If the allocation $(\mathbf{A}^{\#} \mid \mathbf{B}^{\#} \mid \tilde{\mathbf{a}}^{\#} \mid \tilde{\mathbf{b}}^{\#} \mid \tilde{\mathbf{w}}^{\#} \mid \tilde{\mathbf{y}}^{\#})$, and the price vectors $(\boldsymbol{\mu}^{\#} \mid \boldsymbol{\lambda}^{\#} \mid \boldsymbol{\varpi}^{\#})$ and $(\mathbf{p} \mid \mathrm{q})$ satisfy conditions ME1-ME15, and assumptions A1-A18 are satisfied,

then (i) $(\mathbf{B}^{\#} \mid \tilde{\mathbf{a}}^{\#} \mid \tilde{\mathbf{w}}^{\#}) << \mathbf{0}$ and $(\tilde{\mathbf{b}}^{\#} \mid \tilde{\mathbf{y}}^{\#}) >> \mathbf{0}$

(ii) $\mu_i^{\#} > 0 \qquad b_{li}^{\#} + f^i(\mathbf{a}_i^{\#} \mid \mathbf{w}_i^{\#}) = 0 \qquad$ for all $l \in L$ and $i \in I^{l+}$

$\mu_i^{\#} > 0 \qquad y_{ki}^{\#} + f^i(\mathbf{a}_i^{\#} \mid \mathbf{w}_i^{\#}) = 0 \qquad$ for all $k \in K$ and $i \in I^{k+}$

(iii) $\varpi^{\#} > 0 \qquad \Pi^{\#} + \sum_{j \in J} p_j (A_j + A_j^0) + \sum_{l \in L} p_l B_l = 0$

(iv) $p_j^{\#} > 0 \qquad \sum_{i \in I^{j-}} a_{ji}^{\#} + A_j^0 + A_j^{\#} = 0 \qquad$ for all $j \in J$

$p_l^{\#} > 0 \qquad \sum_{i \in I^{l+}} b_{li}^{\#} + B_l^{\#} = 0 \qquad$ for all $l \in L$

$p_k^{\#} > 0 \qquad \sum_{i \in I^{k+}} y_{ki}^{\#} + \sum_{i \in I^{k-}} w_{ki}^{\#} = 0 \qquad$ for all $k \in K$

Proof

Together, conditions ME1, ME2, ME4 and ME8 in Table A3.1.1 (Appendix 3.1) imply that:

$\varpi^{\#} > 0 \qquad \Pi^{\#} + \sum_{j \in J} p_j (A_j + A_j^0) + \sum_{l \in L} p_l B_l = 0$

$p_j^{\#} > 0 \qquad \sum_{i \in I^{j-}} a_{ji}^{\#} + A_j^0 + A_j^{\#} = 0 \qquad$ for all $j \in J$

$p_l^{\#} > 0 \qquad \sum_{i \in I^{l+}} b_{li}^{\#} + B_l^{\#} = 0 \qquad$ for all $l \in L$

$\mu_i^{\#} > 0 \qquad b_{li}^{\#} + f^i(\mathbf{a}_i^{\#} \mid \mathbf{w}_i^{\#}) = 0 \qquad$ for all $l \in L$ and $i \in I^{l+}$

Furthermore, it must be the case that:

- $b_{li}^{\#} \;=\; -\,f^{i}(\mathbf{a}_i^{\#} \mid \mathbf{w}_i^{\#}) > 0$ for all $l \in L$ with $p_l > 0$, $i \in I^{l+}$

- $y_{ki}^{\#} \;=\; -\,f^{i}(\mathbf{a}_i^{\#} \mid \mathbf{w}_i^{\#}) > 0$ for all $k \in K$ with $p_k + q\sigma_k > 0$, $i \in I^{k+}$

since – under assumption A10 – any agent with zero production of a commodity that has a strictly positive (effective) price, can increase its gross economic benefit by producing a small (marginal) quantity of output. Consequently, it follows that $b_{li}^{\#} > 0$ for all $l \in L$ and $i \in I^{l+}$, and hence that $B_l^{\#} < 0$ for all $l \in L$.

The remainder of the proof is directly analogous to that used for Proposition 3.2 – with the appropriate minor changes to notation.[6] The only significant difference arises when one is considering the expression:

$$p_k \;+\; \lambda_i^{\#}\,[\,\sigma_k \;+\; \sum_{m\in M^{i-}} \chi_m h_k^{mi\,\#}\,] \quad = \quad \mu_i^{\#} f_k^{i\,\#} \quad > \quad 0 \qquad \text{for all } k \in K^{i-}.$$

Noting that, by construction, the parameters ρ_k and σ_k cannot have the same sign (unless they are both equal to zero), there are four possible cases that can arise:

(i) $\quad \rho_k = 0 \qquad \sigma_k = 0 \qquad \sum_{m\in M^{i-}} \chi_m h_k^{mi\,*} = 0$

(ii) $\quad \rho_k \geq 0 \qquad \sigma_k \leq 0 \qquad \sum_{m\in M^{i-}} \chi_m h_k^{mi\,*} = 0$ $\quad$ (ρ_k and σ_k not both zero)

(iii) $\quad \rho_k \leq 0 \qquad \sigma_k \geq 0 \qquad \sum_{m\in M^{i-}} \chi_m h_k^{mi\,*} = 0$ $\quad$ (ρ_k and σ_k not both zero)

[6] That is, the superscripts * are replaced by $^\#$; α_k and β_k are replaced by ρ_k and σ_k; η_k is replaced by p_k; and λ is replaced by q.

(iv) $\rho_k = 0$ $\sigma_k = 0$ $\sum_{m \in M^{i-}} \chi_m h_k^{mi\,*} \neq 0,$

In the first two cases, it follows directly that $p_k > 0$. In the last two cases however, one cannot rule out the possibility that $p_k = 0$. However, this would imply either that $y_{ki}^{\#} = 0$ (if $\rho_k \leq 0$), or that $\mu_i^{\#} = 0$ (if $\rho_k = 0$), for all $i \in I^{k+}$ (i.e. for all producers of that commodity). But if $\mu_i^{\#} = 0$ then it follows that $a_{ji}^{\#} = 0$ for some essential input $j \in J^{i-}$. Hence, if $p_k = 0$, then $y_{ki}^{\#} = 0$ for all $i \in I^{k+}$. However, this cannot be so as $w_{ki}^{\#} < 0$ for the agent under consideration, and hence aggregate input is strictly less than zero. Consequently, it must be the case that $p_k > 0$.

(QED)

Appendix A3.4 Proof of proposition 3.4

$\Psi^* \subset \Psi^\#(\theta \mid \varepsilon)$ for any $\theta \in \Theta$ and $\varepsilon \in E$.

Proof

Consider <u>any</u> allocation $(\mathbf{A}^* \mid \mathbf{B}^* \mid \tilde{\mathbf{a}}^* \mid \tilde{\mathbf{b}}^* \mid \tilde{\mathbf{w}}^* \mid \tilde{\mathbf{y}}^*) \in \Psi^*$. Then there exists a vector of shadow prices $(\mu^* \mid \phi^* \mid \gamma^* \mid \eta^* \mid \lambda^*)$, such that $(\mathbf{A}^* \mid \mathbf{B}^* \mid \tilde{\mathbf{a}}^* \mid \tilde{\mathbf{b}}^* \mid \tilde{\mathbf{w}}^* \mid \tilde{\mathbf{y}}^*)$ and $(\mu^* \mid \phi^* \mid \gamma^* \mid \eta^* \mid \lambda^*)$ satisfy the <u>necessary</u> and sufficient conditions RC1-RC12 for a regulated aggregate cost minimum (see Table A3.1.1 in Appendix 3.1).

Noting that $\rho_k \equiv (1-\theta_k)\alpha_k - \theta_k\beta_k$, $\sigma_k \equiv (1-\theta_k)\beta_k - \theta_k\alpha_k$, and $\delta_i \equiv \delta/\mathrm{I} + \varepsilon_i$, condition RC12 implies that:

$$\sum_{i\in \mathrm{I}} \Big[\sum_{k\in \mathrm{K}^{i+}} \rho_k y^*_{ki} + \sum_{k\in \mathrm{K}^{i-}} \sigma_k w^*_{ki} + \sum_{m\in \mathrm{M}^{i-}} \chi_m h^{mi}(\mathbf{w}^*_i) + \delta_i \Big]$$

$$\geq \quad \sum_{i\in \mathrm{I}} \varepsilon_i - \sum_{k\in \mathrm{K}} \theta_k(\alpha_k + \beta_k) \Big[\sum_{k\in \mathrm{K}^{i+}} y^*_{ki} + \sum_{k\in \mathrm{K}^{i-}} w^*_{ki} \Big]$$

$$\lambda^* \Big[\sum_{i\in \mathrm{I}} \Big[\sum_{k\in \mathrm{K}^{i+}} \rho_k y^*_{ki} + \sum_{k\in \mathrm{K}^{i-}} \sigma_k w^*_{ki} + \sum_{m\in \mathrm{M}^{i-}} \chi_m h^{mi}(\mathbf{w}^*_i) + \delta_i \Big]\Big]$$

$$= \quad \lambda^* \Big[\sum_{i\in \mathrm{I}} \varepsilon_i - \sum_{k\in \mathrm{K}} \vartheta_k(\alpha_k + \beta_k) \Big[\sum_{k\in \mathrm{K}^{i+}} y^*_{ki} + \sum_{k\in \mathrm{K}^{i-}} w^*_{ki} \Big]\Big]$$

But $\sum_{i\in \mathrm{I}} \varepsilon_i = 0$, and by Proposition 3.2, $\sum_{i\in \mathrm{I}^{k+}} y^*_{ki} + \sum_{i\in \mathrm{I}^{k-}} w^*_{ki} = 0$ for all $k \in \mathrm{K}$.

Therefore:

$$\sum_{i\in I} [\sum_{k\in K^{i+}} \rho_k y^*_{ki} + \sum_{k\in K^{i-}} \sigma_k w^*_{ki} + \sum_{m\in M^{i-}} \chi_m h^{mi}(\mathbf{w}^*_i) + \delta_i] \geq 0$$

$$\lambda^* [\sum_{i\in I} [\sum_{k\in K^{i+}} \rho_k y^*_{ki} + \sum_{k\in K^{i-}} \sigma_k w^*_{ki} + \sum_{m\in M^{i-}} \chi_m h^{mi}(\mathbf{w}^*_i) + \delta_i]] = 0$$

If $\lambda^* > 0$, one can construct a (unique) vector $\tilde{\mathbf{v}}^*$, such that $\sum_{i\in I} \tilde{\mathbf{v}}^*_i = 0$ and

$$\sum_{k\in K^{i+}} \rho_k y^*_{ki} + \sum_{k\in K^{i-}} \sigma_k w^*_{ki} + \sum_{m\in M^{i-}} \chi_m h^{mi}(\mathbf{w}^*_i) + \delta_i = v^*_i \qquad \text{for all } i \in I$$

Similarly, if $\lambda^* = 0$, one can construct a vector $\tilde{\mathbf{v}}^*$, such that $\sum_{i\in I} \tilde{\mathbf{v}}^*_i \geq 0$ and

$$\sum_{k\in K^{i+}} \rho_k y^*_{ki} + \sum_{k\in K^{i-}} \sigma_k w^*_{ki} + \sum_{m\in M^{i-}} \chi_m h^{mi}(\mathbf{w}^*_i) + \delta_i \geq v^*_i \qquad \text{for all } i \in I$$

Thus, there exists a vector $\tilde{\mathbf{v}}^*$ such that for any $\boldsymbol{\theta} \in \Theta$ and $\boldsymbol{\varepsilon} \in E$, the following two complementary slackness conditions are necessarily satisfied:

(RC13) $\quad \lambda^* \geq 0 \qquad \sum_{i\in I} \mathbf{v}^*_i \geq 0 \qquad \lambda^* [\sum_{i\in I} \mathbf{v}^*_i] = 0$

(RC14) $\quad \lambda^* \geq 0$

$$\sum_{k\in K^{i+}} \rho_k y^*_{ki} + \sum_{k\in K^{i-}} \sigma_k w^*_{ki} + \sum_{m\in M^{i-}} \chi_m h^{mi}(\mathbf{w}^*_i) + \delta_i - v^*_i \geq 0$$

$$\lambda^* [\sum_{k\in K^{i+}} \rho_k y^*_{ki} + \sum_{k\in K^{i-}} \sigma_k w^*_{ki} + \sum_{m\in M^{i-}} \chi_m h^{mi}(\mathbf{w}^*_i) + \delta_i - v^*_i] = 0$$

for all $i \in I$

Together, conditions RC9-RC11 and RC13 imply that:

$$\sum_{j\in J} \varphi_j^* (A_j^0 + A_j^*) + \sum_{l\in L} \gamma_l^* B_l^* + \Pi^* = 0$$

where $\Pi^* = \sum_{i\in I} \left[\sum_{j\in J^{i-}} \varphi_j^* a_{ji}^* + \sum_{l\in L^{i+}} \gamma_l^* b_{li}^* + \sum_{k\in K^{i-}} \eta_k^* w_{ki}^* + \sum_{k\in K^{i+}} \eta_k^* y_{ki}^* + \lambda^* v_i^* \right]$

Thus, for <u>any</u> arbitrary constant $\varpi^* > 0$, the following complementary slackness condition is necessarily satisfied:

(RC15) $\varpi^* \geq 0$ $\quad \sum_{j\in J} \varphi_j^* (A_j^0 + A_j^*) + \sum_{l\in L} \gamma_l^* B_l^* + \Pi^* \geq 0$

$$\varpi^* \left[\sum_{j\in J} \varphi_j^* (A_j^0 + A_j^*) + \sum_{l\in L} \gamma_l^* B_l^* + \Pi^* \right] = 0$$

Now set

$$(\mathbf{A}^\# \mid \mathbf{B}^\# \mid \tilde{\mathbf{a}}^\# \mid \tilde{\mathbf{b}}^\# \mid \tilde{\mathbf{w}}^\# \mid \tilde{\mathbf{y}}^\#) \equiv (\mathbf{A}^* \mid \mathbf{B}^* \mid \tilde{\mathbf{a}}^* \mid \tilde{\mathbf{b}}^* \mid \tilde{\mathbf{w}}^* \mid \tilde{\mathbf{y}}^*)$$

$\tilde{\mathbf{v}}^\# \equiv \tilde{\mathbf{v}}^*$

$\mu_i^\# \equiv \mu_i^* / \varpi^*$ for all $i \in I$

$\lambda_i^\# \equiv \lambda^* / \varpi^*$ for all $i \in I$

$\varpi^\# \equiv \varpi^*$

$\Pi^\# \equiv \Pi^*$

$p_j \equiv \phi_j^* / \varpi^*$ for all $j \in J$

$$p_l \equiv \gamma_l^* / \varpi^* \qquad \text{for all } l \in L$$

$$p_k \equiv [\, \eta_k^* + \theta_k (\alpha_k + \beta_k) \lambda^* \,] / \varpi^* \qquad \text{for all } k \in K$$

$$q \equiv \lambda^* / \varpi^*$$

From which it follows that:

$$p_k + \lambda_i^\# \rho_k \equiv [\, \eta_k^* + \lambda^* \alpha_k \,] / \varpi^* \qquad \text{for all } i \in I,\ k \in K^{i+}$$

$$p_k + \lambda_i^\# \sigma_k \equiv [\, \eta_k^* + \lambda^* \beta_k \,] / \varpi^* \qquad \text{for all } i \in I,\ k \in K^{i-}$$

In conjunction with the <u>necessary</u> conditions for a regulated aggregate cost minimum RC1-RC11, and the derived conditions RC13-RC15,[7] these identities imply that the allocation $(\mathbf{A}^\# \mid \mathbf{B}^\# \mid \tilde{\mathbf{a}}^\# \mid \tilde{\mathbf{b}}^\# \mid \tilde{\mathbf{w}}^\# \mid \tilde{\mathbf{y}}^\#)$, together with the performance credit plan $\tilde{\mathbf{v}}^\#$, shadow prices $(\boldsymbol{\mu}^\# \mid \boldsymbol{\lambda}^\# \mid \varpi^\#)$, and market prices $(\mathbf{p} \mid \mathrm{q})$ satisfy the necessary and <u>sufficient</u> conditions for a market equilibrium ME1-ME15. Thus, this particular regulated cost minimum allocation is an element of $\Psi^\#(\boldsymbol{\theta} \mid \boldsymbol{\varepsilon})$ for any $\boldsymbol{\theta} \in \Theta$ and $\boldsymbol{\varepsilon} \in \mathrm{E}$.

Since the argument can be applied to <u>any</u> allocation $(\mathbf{A}^* \mid \mathbf{B}^* \mid \tilde{\mathbf{a}}^* \mid \tilde{\mathbf{b}}^* \mid \tilde{\mathbf{w}}^* \mid \tilde{\mathbf{y}}^*) \in \Psi^*$, it follows that $\Psi^* \subset \Psi^\#(\boldsymbol{\theta} \mid \boldsymbol{\varepsilon})$ for any $\boldsymbol{\theta} \in \Theta$ and $\boldsymbol{\varepsilon} \in \mathrm{E}$.

(QED)

[7] For conditions RC3-RC11 and RC13-RC15, the shadow prices are all scaled by the arbitrary positive factor $1/\varpi^*$. This does not affect the values of any of the real variables in the solution, or any of the shadow prices.

Appendix A3.5 Proof of proposition 3.5

For any $\theta \in \Theta$ and $\varepsilon \in E$, $\Psi^{\#}(\theta \mid \varepsilon) \subset \Psi^{*}$

Proof

For <u>any</u> $\theta \in \Theta$ and $\varepsilon \in E$, take <u>any</u> allocation $(\mathbf{A}^{\#} \mid \mathbf{B}^{\#} \mid \tilde{\mathbf{a}}^{\#} \mid \tilde{\mathbf{b}}^{\#} \mid \tilde{\mathbf{w}}^{\#} \mid \tilde{\mathbf{y}}^{\#}) \in \Psi^{\#}(\theta \mid \varepsilon)$.

Then there exist a performance credit plan $\tilde{\mathbf{v}}^{\#}$, shadow prices $(\mu^{\#} \mid \lambda^{\#} \mid \varpi^{\#})$, and market prices $(\mathbf{p} \mid \mathrm{q})$, such that $(\mathbf{A}^{\#} \mid \mathbf{B}^{\#} \mid \tilde{\mathbf{a}}^{\#} \mid \tilde{\mathbf{b}}^{\#} \mid \tilde{\mathbf{w}}^{\#} \mid \tilde{\mathbf{y}}^{\#})$, $\tilde{\mathbf{v}}^{\#}$, $(\mu^{\#} \mid \lambda^{\#} \mid \varpi^{\#})$ and $(\mathbf{p} \mid \mathrm{q})$ satisfy the <u>necessary</u> and sufficient conditions ME1-ME15 for a market equilibrium (see Table A3.1.2 in Appendix 3.1).

Noting that $\rho_k \equiv (1-\theta_k)\alpha_k - \theta_k\beta_k$, $\sigma_k \equiv (1-\theta_k)\beta_k - \theta_k\alpha_k$, and $\delta_i \equiv \delta/\mathrm{I} + \varepsilon_i$, conditions ME7 and ME15 imply that:

$$\sum_{k\in K}\alpha_k \sum_{i\in I^{k+}} y^{\#}_{ki} + \sum_{k\in K}\beta_k \sum_{i\in I^{k-}} w^{\#}_{ki} + \sum_{m\in M}\chi_m \sum_{i\in I^{m-}} h^{mi}(\mathbf{w}^{\#}_i) + \delta$$

$$\geq \sum_{k\in K} \theta_k(\alpha_k + \beta_k) \,[\, \sum_{k\in K^{i+}} y^{\#}_{ki} + \sum_{k\in K^{i-}} w^{\#}_{ki} \,] - \sum_{i\in I} \varepsilon_i$$

$$\mathrm{q}\,[\, \sum_{k\in K}\alpha_k \sum_{i\in I^{k+}} y^{\#}_{ki} + \sum_{k\in K}\beta_k \sum_{i\in I^{k-}} w^{\#}_{ki} + \sum_{m\in M}\chi_m \sum_{i\in I^{m-}} h^{mi}(\mathbf{w}^{\#}_i) + \delta \,]$$

$$= \mathrm{q}\,[\, \sum_{k\in K} \theta_k(\alpha_k + \beta_k) \,[\, \sum_{k\in K^{i+}} y^{\#}_{ki} + \sum_{k\in K^{i-}} w^{\#}_{ki} \,] - \sum_{i\in I} \varepsilon_i \,]$$

But $\sum_{i\in I} \varepsilon_i = 0$, and by Proposition 3.3, $\sum_{k\in K^{i+}} y^{\#}_{ki} + \sum_{k\in K^{i-}} w^{\#}_{ki} = 0$ for all $k \in K$.

Therefore, the following complementary slackness condition is necessarily satisfied:

ME16: $q \;\geq\; 0$

$$\sum_{k\in K}\alpha_k \sum_{i\in I^{k+}} y^{\#}_{ki} \;+\; \sum_{k\in K}\beta_k \sum_{i\in I^{k-}} w^{\#}_{ki} \;+\; \sum_{m\in M}\chi_m \sum_{i\in I^{m-}} h^{mi}(\mathbf{w}^{\#}_i) \;+\; \delta \;\geq\; 0$$

$$q\,[\,\sum_{k\in K}\alpha_k \sum_{i\in I^{k+}} y^{\#}_{ki} \;+\; \sum_{k\in K}\beta_k \sum_{i\in I^{k-}} w^{\#}_{ki} \;+\; \sum_{m\in M}\chi_m \sum_{i\in I^{m-}} h^{mi}(\mathbf{w}^{\#}_i) \;+\; \delta\;] \;=\; 0$$

Now set:

$$(\mathbf{A}^* \mid \mathbf{B}^* \mid \tilde{\mathbf{a}}^* \mid \tilde{\mathbf{b}}^* \mid \tilde{\mathbf{w}}^* \mid \tilde{\mathbf{y}}^*) \;\equiv\; (\mathbf{A}^{\#} \mid \mathbf{B}^{\#} \mid \tilde{\mathbf{a}}^{\#} \mid \tilde{\mathbf{b}}^{\#} \mid \tilde{\mathbf{w}}^{\#} \mid \tilde{\mathbf{y}}^{\#})$$

$\mu_i^* \;\equiv\; \varpi^{\#}\,\mu_i^{\#}$ for all $i \in I$

$\phi_j^* \;\equiv\; \varpi^{\#}\,p_j$ for all $j \in J$

$\gamma_l^* \;\equiv\; \varpi^{\#}\,p_l$ for all $l \in L$

$\eta_k^* \;\equiv\; \varpi^{\#}\,[\,p_k \;-\; \theta_k\,(\alpha_k + \beta_k)\,q\,]$ for all $k \in K$[8]

$\lambda^* \;\equiv\; \varpi^{\#}\,q$

From which it follows that:

$\eta_k^* \;+\; \lambda^*\,\alpha_k \;\equiv\; \varpi^{\#}\,[\,p_k \;+\; \lambda_i^{\#}\,\rho_k\,]$ for all $k \in K,\ i \in I^{k+}$

$\eta_k^* \;+\; \lambda^*\,\beta_k \;\equiv\; \varpi^{\#}\,[\,p_k \;+\; \lambda_i^{\#}\,\sigma_k\,]$ for all $k \in K,\ i \in I^{k-}$

---

[8] In the proof of proposition 3.3 it is demonstrated that $p_k + \rho_k q > 0$ and $p_k + \sigma_k q > 0$. Together, these strict inequalities imply that $p_k - \theta_k\,(\alpha_k + \beta_k)\,q > 0$.

In conjunction with the necessary conditions for a market equilibrium ME1-ME6; ME8-ME9. and the derived condition ME16,[9] these identities imply that the production plan $(\mathbf{A}^* \mid \mathbf{B}^* \mid \tilde{\mathbf{a}}^* \mid \tilde{\mathbf{b}}^* \mid \tilde{\mathbf{w}}^* \mid \tilde{\mathbf{y}}^*)$, together with the shadow prices $(\mu^* \mid \phi^* \mid \gamma^* \mid \eta^* \mid \lambda^*)$ satisfy the necessary and sufficient conditions for a regulated cost minimum RC1- RC12. Thus, this particular market equilibrium allocation is an element of Ψ^*.

Since the argument can be applied to any values of $\theta \in \Theta$ and $\varepsilon \in E$, and to any market equilibrium allocation $(\mathbf{A}^\# \mid \mathbf{B}^\# \mid \tilde{\mathbf{a}}^\# \mid \tilde{\mathbf{b}}^\# \mid \tilde{\mathbf{w}}^\# \mid \tilde{\mathbf{y}}^\#) \in \Psi^\#(\theta \mid \varepsilon)$, it follows that for any $\theta \in \Theta$ and $\varepsilon \in E$, $\Psi^\#(\theta \mid \varepsilon) \subset \Psi^*$.

(QED)

[9] For conditions ME3-ME15, the shadow prices and market prices can all be scaled by the strictly positive value of the shadow price of the representative consumer's budget constraint $\varpi^\#$. This does not affect the values of any of the real variables in the market equilibrium, or any of the shadow and market prices.

Appendix A3.6 Proof of proposition 3.6

For any $(\mathbf{A}^{\#} \mid \mathbf{B}^{\#} \mid \tilde{\mathbf{a}}^{\#} \mid \tilde{\mathbf{b}}^{\#} \mid \tilde{\mathbf{w}}^{\#} \mid \tilde{\mathbf{y}}^{\#}) \in \Psi^{\#}$

Let $(\mathbf{p}^{(m)} \mid q^{(m)})$ and $(\boldsymbol{\mu}^{\#(m)} \mid \boldsymbol{\lambda}^{\#(m)} \mid 1)$ denote the market and shadow price vectors that support the market equilibrium for any two vectors of design parameters $\theta^{(m)} \in \Theta$ and $\varepsilon^{(m)} \in \mathrm{E}$ (where $m = 1,2$)

Then
$$\mu_i^{\#(1)} = \mu_i^{\#(2)} = \mu_i^{\#} \qquad \text{for all } i \in \mathrm{I}$$

$$\lambda_i^{\#(1)} = \lambda_i^{\#(2)} = \lambda_i^{\#} \qquad \text{for all } i \in \mathrm{I}$$

$$q^{(1)} = q^{(2)} = q$$

$$p_j^{(1)} = p_j^{(2)} = p_j \qquad \text{for all } j \in \mathrm{J}$$

$$p_l^{(1)} = p_l^{(2)} = 1 \qquad \text{for all } l \in \mathrm{L}$$

$$p_k^{(1)} - p_k^{(2)} = q\,(\alpha_k + \beta_k)\,(\theta_k^{(1)} - \theta_k^{(2)}) \qquad \text{for all } k \in \mathrm{K}$$

Proof

Consider any allocation $(\mathbf{A}^{\#} \mid \mathbf{B}^{\#} \mid \tilde{\mathbf{a}}^{\#} \mid \tilde{\mathbf{b}}^{\#} \mid \tilde{\mathbf{w}}^{\#} \mid \tilde{\mathbf{y}}^{\#}) \in \Psi^{\#}$. If the price vectors $(\mathbf{p}^{(m)} \mid q^{(m)})$ and $(\boldsymbol{\mu}^{\#(m)} \mid \boldsymbol{\lambda}^{\#(m)} \mid 1)$ support the market equilibrium for $\theta^{(m)} \in \Theta$ and $\varepsilon^{(m)} \in \mathrm{E}$ ($m = 1, 2$) respectively, then (from the proof of proposition 3.5):

$$(\mathbf{A}^{*} \mid \mathbf{B}^{*} \mid \tilde{\mathbf{a}}^{*} \mid \tilde{\mathbf{b}}^{*} \mid \tilde{\mathbf{w}}^{*} \mid \tilde{\mathbf{y}}^{*}) \equiv (\mathbf{A}^{\#} \mid \mathbf{B}^{\#} \mid \tilde{\mathbf{a}}^{\#} \mid \tilde{\mathbf{b}}^{\#} \mid \tilde{\mathbf{w}}^{\#} \mid \tilde{\mathbf{y}}^{\#})$$

$$\mu_i^{*} \equiv \mu_i^{\#(m)} \qquad \text{for all } i \in \mathrm{I}$$

$$\phi_j^{*} \equiv p_j^{(m)} \qquad \text{for all } j \in \mathrm{J}$$

$$\gamma_l^* \equiv p_l^{(m)} \qquad \text{for all } l \in L$$

$$\eta_k^* \equiv p_k^{(m)} - \theta_k (\alpha_k + \beta_k)\, q^{(m)} \qquad \text{for all } k \in K$$

$$\lambda^* \equiv q^{(m)}$$

satisfy the necessary and sufficient conditions for a regulated cost minimum RC1-RC12.

But, for any $(\mathbf{A}^* \mid \mathbf{B}^* \mid \tilde{\mathbf{a}}^* \mid \tilde{\mathbf{b}}^* \mid \tilde{\mathbf{w}}^* \mid \tilde{\mathbf{y}}^*) \in \Psi^*$, the value of the shadow price vector $(\boldsymbol{\mu}^* \mid \boldsymbol{\phi}^* \mid \boldsymbol{\gamma}^* \mid \boldsymbol{\eta}^* \mid \lambda^*)$ is unique. Therefore:

$$\mu_i^{\#(1)} = \mu_i^* = \mu_i^{\#(2)} \qquad \text{for all } i \in I$$

$$p_j^{(1)} = \phi_j^* = p_j^{(2)}$$

$$p_l^{(1)} = \gamma_l^* = p_l^{(2)}$$

$$q^{(1)} = \lambda^* = q^{(2)}$$

$$p_k^{(1)} - \theta_k (\alpha_k + \beta_k)\, q^{(1)} = \eta_k^* = p_k^{(2)} - \theta_k (\alpha_k + \beta_k)\, q^{(2)} \qquad \text{for all } k \in K$$

Noting that $\lambda_i^{\#(m)} = q^{(m)}$ for all $i \in I$ (condition ME5), it follows directly that:

$$\lambda_i^{\#(1)} = \lambda^* = \lambda_i^{\#(2)} \qquad \text{for all } i \in I$$

$$q^{(1)} = q^{(2)} = q$$

$$p_j^{(1)} = p_j^{(2)} = p_j \qquad \text{for all } j \in J$$

$$p_l^{(1)} = p_l^{(2)} = p_l = 1 \qquad \text{for all } l \in L$$

$$p_k^{(1)} - p_k^{(2)} = q\,(\alpha_k + \beta_k)\,(\theta_k^{(1)} - \theta_k^{(2)}) \qquad \text{for all } k \in K$$

(QED)

Appendix A3.7 Proof of proposition 3.7

For any allocation $(\mathbf{A}^{\#} \mid \mathbf{B}^{\#} \mid \tilde{\mathbf{a}}^{\#} \mid \tilde{\mathbf{b}}^{\#} \mid \tilde{\mathbf{w}}^{\#} \mid \tilde{\mathbf{y}}^{\#}) \in \Psi^{\#}$

Let $\Pi_i^{\#(a,b)}$ denote the gross economic benefit for agent $i \in \mathrm{I}$ in the market equilibrium for given vectors of design parameters $\boldsymbol{\theta}^{(m)} \in \Theta$ $(m = 1, 2)$ and $\boldsymbol{\varepsilon}^{(n)} \in \mathrm{E}$ $(n = 1, 2)$

Then
$$\Pi_i^{\#(1,n)} - \Pi_i^{\#(2,n)} = 0$$

$$\Pi_i^{\#(m,1)} - \Pi_i^{\#(m,2)} = \mathrm{q}\,(\,\varepsilon_i^{(1)} - \varepsilon_i^{(2)}\,)$$

for all $i \in \mathrm{I}$

Proof

Take any allocation $(\mathbf{A}^{\#} \mid \mathbf{B}^{\#} \mid \tilde{\mathbf{a}}^{\#} \mid \tilde{\mathbf{b}}^{\#} \mid \tilde{\mathbf{w}}^{\#} \mid \tilde{\mathbf{y}}^{\#}) \in \Psi^{\#}$. Then by definition (3.7), Proposition 3.6, and market equilibrium conditions (ME7) and (ME15), it follows that:

$$\Pi_i^{\#(m,n)}$$

$$= \Big[\sum_{j \in \mathrm{J}^{i+}} \mathrm{b}_{ji}^{\#} + \sum_{j \in \mathrm{J}^{i-}} \mathrm{a}_{ki}^{\#} + \sum_{k \in \mathrm{K}^{i+}} \mathrm{p}_k^{(m,n)} \mathrm{y}_{ki}^{\#} + \sum_{k \in \mathrm{K}^{i-}} \mathrm{p}_k^{(m,n)} \mathrm{w}_{ki}^{\#} \Big] + \mathrm{q}\,\mathrm{v}_i^{\#(m,n)}$$

$$= \Big[\sum_{j \in \mathrm{J}^{i+}} \mathrm{b}_{ji}^{\#} + \sum_{j \in \mathrm{J}^{i-}} \mathrm{a}_{ki}^{\#} + \sum_{k \in \mathrm{K}^{i+}} \mathrm{p}_k^{(m,n)} \mathrm{y}_{ki}^{\#} + \sum_{k \in \mathrm{K}^{i-}} \mathrm{p}_k^{(m,n)} \mathrm{w}_{ki}^{\#} \Big]$$

$$+ \mathrm{q}\Big[\sum_{k \in \mathrm{K}^{i+}} \rho_k \mathrm{y}_{ki}^{\#} + \sum_{k \in \mathrm{K}^{i-}} \sigma_k \mathrm{w}_{ki}^{\#} + \sum_{m \in \mathrm{M}^{i-}} \chi_m \mathrm{h}^{mi}(\mathbf{w}_i^{\#}) + \delta_i \Big]$$

$$= \Big[\sum_{j \in \mathrm{J}^{i+}} \mathrm{b}_{ji}^{\#} + \sum_{j \in \mathrm{J}^{i-}} \mathrm{a}_{ki}^{\#} \Big]$$

$$+ \quad \sum_{k \in K^{i+}} (p_k^{(m,n)} - \theta_k^{(m)}(\alpha_k + \beta_k)\, q)\, y_{ki}^{\#} \quad + \quad \sum_{k \in K^{i-}} (p_k^{(m,n)} - \theta_k^{(m)}(\alpha_k + \beta_k)\, q)\, w_{ki}^{\#}$$

$$+ \quad q\, [\sum_{k \in K^{i+}} \alpha_k\, y_{ki}^{\#} \quad + \quad \sum_{k \in K^{i-}} \beta_k\, w_{ki}^{\#} \quad + \quad \sum_{m \in M^{i-}} \chi_m\, h^{mi}(\mathbf{w}_i^{\#}) \quad + \quad \delta / I \quad + \quad \varepsilon_i^{(n)} \ ]$$

But, by Proposition 3.6

$$p_k^{(1,n)} \; - \; q\,(\alpha_k + \beta_k)\, \theta_k^{(1)} \quad = \quad p_k^{(2,n)} \; - \; q\,(\alpha_k + \beta_k)\, \theta_k^{(2)}$$

$$p_k^{(m,1)} \; - \; q\,(\alpha_k + \beta_k)\, \theta_k^{(m)} \quad = \quad p_k^{(m,2)} \; - \; q\,(\alpha_k + \beta_k)\, \theta_k^{(m)}$$

Therefore, for all $i \in$ I:

$$\Pi_i^{\#(1,n)} \; - \; \Pi_i^{\#(2,n)} \quad = \quad 0$$

$$\Pi_i^{\#(m,1)} \; - \; \Pi_i^{\#(m,2)} \quad = \quad q\,(\varepsilon_i^{(1)} \; - \; \varepsilon_i^{(2)})$$

(QED)

Chapter 4 Comparative statics analysis for performance rule parameters

Having established the environmental effectiveness, cost efficiency and distributional flexibility of performance-based credit trading, attention is now turned to the impact of changes in the values of performance rule parameters. Since these changes reflect changes in the regulatory control variables (i.e. L and r), this is equivalent to considering how the outcome responds to changes in the stringency of the regulation.

While it is possible to conduct comparative statics analyses for any of the real variables (or aggregations thereof) or prices in the production system, the analyses undertaken in this chapter focus on the impact of the changes on the market equilibrium price of performance credits.[1] However, since it is easier to work with the regulated aggregate cost minimization problem, the issue will be addressed indirectly, by considering the impact on the shadow value of the aggregate performance rule.[2]

In order to facilitate the analyses, a number of assumptions are made regarding the structure of the production system; and the properties of the production and emission functions of the various agents. These are set out in the first section of the chapter.

[1] The impacts on a number of real variables and prices are investigated in chapters 5 and 6, in the context of two specific applications of performance-based credit trading.

[2] As was noted in Chapter 3, the market equilibrium price of performance credits is equal to the shadow price of the aggregate performance rule in the regulated aggregate cost minimization problem (i.e. $q = \lambda^*$)

While the resultant model still satisfies the general assumptions A1-A18 that are set out in Chapter 3, it has a much more restrictive structure. Nonetheless it can still accommodate a number of different forms of regulatory intervention. In the succeeding three sections, the model is used to investigate respectively the impacts of marginal increases in the values of an output parameter, an input parameter, and the constant term in the aggregate performance rule.

4.1 Simplified model

In order to facilitate the analysis, the following assumptions are made regarding the structure of the economy, and the properties of the production and emission functions of the various agents.

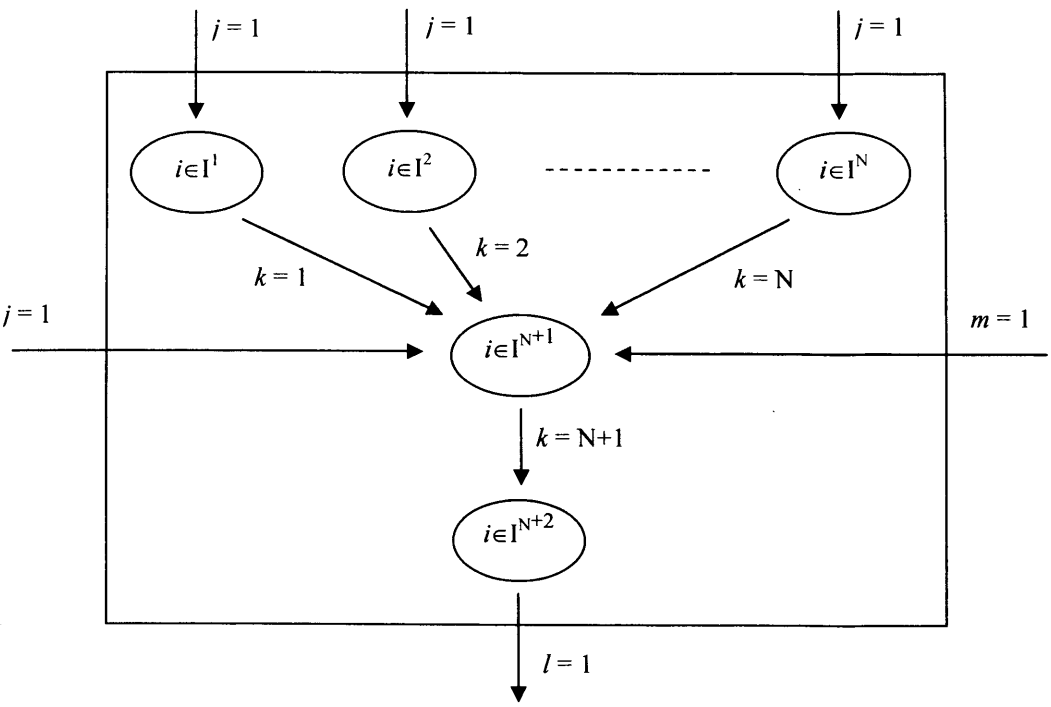
- There is a single primary resource; N+1 secondary commodities; a single utility commodity; and a single non-market (environmental) commodity. Agents are partitioned into N+2 subsets, or sectors – denoted by I^n ($n = 1, \ldots, N+2$). The agents in the first N+1 sectors are all firms, while those in the final sector are all individuals.

- The producer and user subsets for the various commodities are:

I^{j-}	$=$	$I \setminus I^{N+2}$	for primary commodity $j = 1$
I^{l+}	$=$	I^{N+2}	for utility commodity $l = 1$
I^{k+}	$=$	$I^n \quad (n = k)$	for secondary commodity $k = 1, \ldots, N+1$
I^{k-}	$=$	I^{N+1}	for secondary commodity $k = 1, \ldots, N$
	$=$	I^{N+2}	for secondary commodity $k = N+1$
I^{m-}	$=$	I^{N+1}	for non-market commodity $m = 1$

Thus each secondary commodity is produced by firms in a particular sector. The first N commodities are used as inputs by firms in sector N+1 to produce commodity N+1; which is then used by individuals in sector N+2 to produce consumption utility. Finally, only firms in sector N+1 generate emissions during production (i.e. use the non-market commodity as an input). Figure 4.1 provides a schematic representation of the production system.

Figure 4.1 Simplified production system



- In order to simplify notation, the analysis is conducted at the sector level.[3] The aggregate production plan for sector I^n is denoted by $(\boldsymbol{a}_n \mid \boldsymbol{b}_n \mid \boldsymbol{w}_n \mid y_n)$, and the aggregate production function for sector I^n is denoted by $F^n(\boldsymbol{a}_n \mid \boldsymbol{w}_n)$. In order to

[3] This does not affect the generality of the results. The same simplification could have been achieved by assuming that there is only a single agent in each sector.

facilitate the interpretation of the results, all inputs – market and non-market – are represented by positive values (i.e. $(\boldsymbol{a}_n \mid \boldsymbol{b}_n \mid \boldsymbol{w}_n \mid \boldsymbol{y}_n \mid \boldsymbol{z}_n) \geq \mathbf{0}$).

- The aggregate emissions function for sector N+1 is linear, with

$$h(\boldsymbol{w}_{N+1}) \quad = \quad \boldsymbol{\zeta}' \boldsymbol{w}_{N+1}$$

where $\boldsymbol{\zeta}$ is a N-dimension vector of emission coefficients.

- The performance rule parameters are $\boldsymbol{\alpha} = (0, \dots, 0, \alpha)$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_N, 0)$, and $\boldsymbol{\chi} = (\chi_1)$. Consequently, the aggregate performance rule can be re-stated as:

$$\alpha\, y_{N+1} \;-\; \boldsymbol{\gamma}' \boldsymbol{w}_{N+1} \;+\; \delta \;\geq\; 0$$

where $\boldsymbol{\gamma}$ is a N-dimension vector, with $\gamma_k \;=\; \beta_k \;+\; \chi_1\, \zeta_k$. Thus, a change to the value of γ_k may reflect either a change to β_k, or a change to χ_1. Furthermore, it is assumed that $\beta_k = \zeta_k = 0$ for at least one of the commodities $k = 1, \dots, N$.

- The aggregate production functions for sectors $n = 1, \dots, N$, and N+2 are all linear, taking the forms:[4]

$$F^n(\cdot) \quad = \quad \phi_n\, a_n \qquad \text{for all } n = 1, \dots, N$$

$$F^{N+2}(\cdot) \quad = \quad \phi_{N+2}\, w_{N+2}$$

where ϕ_n is a sector specific constant. It is assumed that the shadow price of the primary commodity is constant (equal to one).[5] Consequently, the shadow prices

[4] Strictly speaking the production functions should be defined as composite functions in order to comply with assumptions A10-A11 (see Figure 3.3). However, given the assumption that is made regarding the production function of sector N+1, the quantities of all secondary commodities are strictly positive and finite in the solution, and hence all of the sectors operate on the linear part of their respective production functions.

of all secondary commodities are also constant. In order to clarify notation, the prices of commodities $k = 1, \ldots, N$ are denoted by the vector $\mathbf{c} = (c_1, \ldots, c_N)$, and the price of commodity N+1 by the scalar p, where $c_n = 1 / \phi_n$ and $p = \phi_{N+2}$.

- The aggregate production function for sector N+1 is homogeneous of degree $\psi < 1$. Since the shadow prices of all secondary commodities are constant, and $\gamma_k = 0$ for some $k = 1, \ldots, N$, the requirement that all firms must use the primary input (i.e. assumption A5) can be discarded for this sector. Consequently:[6]

$$F^{N+1}(w_{N+1}) = \frac{1}{\psi}\left[\sum_{k=1}^{N} w_{k\,N+1}\, F^{N+1}_{k}\right] \quad \text{and}$$

$$F^{N+1}_{k}(w_{N+1}) = \frac{1}{\psi - 1}\left[\sum_{l=1}^{N} w_{l\,N+1}\, F^{N+1}_{kl}\right]$$

By Proposition 3.2 of Chapter 3, the production constraint for each sector is binding in the regulated aggregate cost minimum; as is the resource constraint for each commodity. Consequently, under the above assumptions, the Lagrangian for the regulated aggregate cost minimization problem can be redefined solely in terms of the aggregate input vector of sector I^{N+1}. That is:

$$\mathcal{L} = p\,F(\boldsymbol{w}) - \mathbf{c}\,\boldsymbol{w} + \lambda\,[\,\alpha\,F(\boldsymbol{w}) - \boldsymbol{\gamma}\,\boldsymbol{w} + \delta\,]$$

where the sector index N+1 has been omitted for simplicity. Assuming that the aggregate performance rule is binding, then by Propositions 3.1 and 3.2 of Chapter 3, there exists a vector of inputs $\mathbf{w}^* \gg \mathbf{0}$ and a shadow price $\lambda^* > 0$ that satisfy the following necessary and sufficient first order conditions:

[5] This implies that some of the primary resource is consumed directly. For example, if the resource is time, then this direct consumption could be in the form of sleep.

[6] In this chapter all index labels (i.e. *j*, *k*, *l*) relate to endogenous market commodities.

$$\mathcal{L}_k \;=\; (p + \lambda\alpha)\, F_k(\boldsymbol{w}) \;-\; (c_k + \lambda\gamma_k) \;=\; 0 \qquad \text{for } k = 1, \ldots, N \qquad \ldots (4.1)$$

$$\mathcal{L}_\lambda \;=\; \alpha\, F(\boldsymbol{w}) \;-\; \gamma\, \boldsymbol{w} \;+\; \delta \;=\; 0 \qquad \ldots (4.2)$$

Furthermore, provided that $(p, \mathbf{c}) \neq \varpi\,(\alpha, \gamma)$, where ϖ is some constant, the solution is unique. Since the first order conditions (4.1) and (4.2) are necessary and sufficient, the second order sufficiency conditions for a maximum are satisfied. Consequently, the *Implicit Function Theorem* is applicable, and the conditions can be solved explicitly to yield $w_k^* = w_k(\alpha, \gamma, \delta)$ and $\lambda^* = \lambda(\alpha, \gamma, \delta)$.[7]

The impact on the shadow price λ^* of changes to the values of the performance rule parameters is given by the total differential:

$$d\lambda^* \;=\; \lambda_\alpha\, d\alpha \;+\; \sum_{k=1}^{N} \lambda_{\gamma_k}\, d\gamma_k \;+\; \lambda_\delta\, d\delta \qquad \ldots (4.3)$$

The three components of the total differential will be considered in turn; with the values of the other rule parameters being held constant in each case.

4.2 Impact of changes to the output parameter α

If the values of γ and δ are held constant, then the total differential (4.3) simplifies to:

$$d\lambda^* \;=\; \lambda_\alpha\, d\alpha.$$

Substituting the explicit solutions $w_k^* = w_k(\alpha, \gamma, \delta)$ and $\lambda^* = \lambda(\alpha, \gamma, \delta)$ back into the first order conditions (4.1) and (4.2), yields the following identities:

[7] Again, in the spirit of simplification, the other arguments in the functions for the solution variables are suppressed.

$$(p + \alpha \lambda(\cdot)) F_k(\mathbf{w}(\cdot)) - (c_k + \gamma_k \lambda(\cdot)) \equiv 0 \qquad \text{for all } k \in K \qquad \ldots (4.4)$$

$$\alpha F(\mathbf{w}(\cdot)) - \gamma \mathbf{w}(\cdot) + \delta \equiv 0 \qquad \ldots (4.5)$$

Since these are identities, they can be partially differentiated with respect to α; to yield the following system of equations:

$$\left[\begin{array}{c|ccc} 0 & \mathcal{L}_{\lambda 1} & \cdots & \mathcal{L}_{\lambda N} \\ \hline \mathcal{L}_{1\lambda} & \mathcal{L}_{11} & \cdots & \mathcal{L}_{1N} \\ \vdots & \vdots & & \\ \mathcal{L}_{N\lambda} & \mathcal{L}_{N1} & & \mathcal{L}_{NN} \end{array}\right] \begin{bmatrix} \dfrac{\partial \lambda}{\partial \alpha} \\ \dfrac{\partial w_1}{\partial \alpha} \\ \vdots \\ \dfrac{\partial w_N}{\partial \alpha} \end{bmatrix} = \begin{bmatrix} -\mathcal{L}_{\lambda\alpha} \\ -\mathcal{L}_{1\alpha} \\ \vdots \\ -\mathcal{L}_{N\alpha} \end{bmatrix}$$

where

$$\mathcal{L}_{kl} = (p + \lambda^* \alpha) F_{kl}(\mathbf{w}^*) \qquad \text{for } k,l = 1, \ldots, N$$

$$\mathcal{L}_{k\lambda} = \alpha F_k(\mathbf{w}^*) - \gamma_k \qquad \text{for } k = 1, \ldots, N$$

$$\mathcal{L}_{k\alpha} = \lambda^* F_k(\mathbf{w}^*) \qquad \text{for } k = 1, \ldots, N$$

$$\mathcal{L}_{\lambda\alpha} = F(\mathbf{w}^*)$$

By Cramer's rule:

$$\frac{\partial \lambda}{\partial \alpha} \equiv \lambda_\alpha = \frac{A}{G} \qquad \ldots (4.6)$$

where

$$A = \left|\begin{array}{c|ccc} -\mathcal{L}_{\lambda\alpha} & \mathcal{L}_{\lambda 1} & \cdots & \mathcal{L}_{\lambda N} \\ \hline -\mathcal{L}_{1\alpha} & \mathcal{L}_{11} & \cdots & \mathcal{L}_{1N} \\ \vdots & \vdots & & \\ -\mathcal{L}_{N\alpha} & \mathcal{L}_{N1} & & \mathcal{L}_{NN} \end{array}\right| \qquad G = \left|\begin{array}{c|ccc} 0 & \mathcal{L}_{\lambda 1} & \cdots & \mathcal{L}_{\lambda N} \\ \hline \mathcal{L}_{1\lambda} & \mathcal{L}_{11} & \cdots & \mathcal{L}_{1N} \\ \vdots & \vdots & & \\ \mathcal{L}_{N\lambda} & \mathcal{L}_{N1} & & \mathcal{L}_{NN} \end{array}\right|$$

In appendix A4.1, the following expression is derived for the determinant of matrix **A**:

$$A \quad = \quad \left(\frac{-1}{\psi - 1}\right) (p + \lambda^{*}\alpha)^{N-1} H [(\psi - 1) p F(\mathbf{w}^{*}) + \lambda^{*}\delta] \qquad \ldots (4.7)$$

where H is the determinant of the Hessian matrix of $F(\boldsymbol{w})$.

Together, (4.6) and (4.7) imply that:

$$\lambda_{\alpha} \quad = \quad \left(\frac{1}{1-\psi}\right) \left(\frac{H}{G}\right) (p + \lambda^{*}\alpha)^{N-1} [\lambda^{*}\delta - (1-\psi) p F(\mathbf{w}^{*})] \qquad \ldots (4.8)$$

There are several points to note about expression (4.8). First, its derivation does not rely on any assumptions about the functional form of the production function – other than homogeneity. Second, although the performance rule parameters γ do not appear explicitly in (4.8), this does not imply that the impact on the shadow price of a change in the parameter α is independent of the values of γ. Different values of these parameters will result in different solution values for $F(\mathbf{w}^{*})$, and for the values of the determinants H and G. Third, since $(p + \lambda\alpha) > 0$, and H and G always have the same sign, the first three terms are all positive, and hence the overall sign depends on the sign of the final term (in square brackets). [8]

When $\delta = 0$, the final term is unambiguously negative (since $\psi < 1$), and hence $\lambda_{\alpha} < 0$. However, when $\delta > 0$, the sign of this term is ambiguous; depending on the relative magnitudes of the two component terms, which are both positive. However, if $\alpha > 0$ it

[8] It is shown in the proof of Proposition 3.1 (see Appendix 3.1) that $p + \lambda\alpha > 0$ for all values of α. Since $F(\boldsymbol{w})$ is homogeneous of degree $\psi < 1$, the Hessian matrix is negative definite. Consequently, $H < 0$ if N is odd and $H > 0$ if N is even. From the sufficient conditions for a maximum, $G < 0$ if N is odd and $G > 0$ if N is even. Hence, H and G always have the same sign.

can be shown that the second term dominates for values of $\psi \leq 0.5$, and hence the overall sign is negative.

$$\lambda^*\delta - (1-\psi)\,p\,F(\mathbf{w}^*) = \lambda^*[\gamma\mathbf{w}^* - \alpha F(\mathbf{w}^*)] - (1-\psi)\,p\,F(\mathbf{w}^*)$$

$$= \lambda^*\gamma\mathbf{w}^* - [(1-\psi)\,p + \alpha\lambda^*]\,F(\mathbf{w}^*)$$

$$= \lambda^*\gamma\mathbf{w}^* - \left[\left(\frac{1-\psi}{\psi}\right)p + \left(\frac{1}{\psi}\right)\alpha\lambda^*\right]\sum_{k=1}^{N} w_k^* F_k$$

$$= \sum_{k=1}^{N} w_k^*\left[\lambda^*\gamma_k - \left[\left(\frac{1-\psi}{\psi}\right)p + \left(\frac{1}{\psi}\right)\alpha\lambda^*\right]F_k\right]$$

But if $\psi \leq 0.5$, $\left(\frac{1-\psi}{\psi}\right) \geq 1$ and $\left(\frac{1}{\psi}\right) \geq 1$. Noting that $c_k > 0$, it follows that for $\alpha > 0$

$$\lambda^*\gamma_k - \left[\left(\frac{1-\psi}{\psi}\right)p + \left(\frac{1}{\psi}\right)\alpha\lambda^*\right]F_k < (\lambda^*\gamma_k + c_k) - (p + \lambda^*\alpha)\,F_k$$

for all $k = 1, \ldots, N$

and hence that:

$$(\psi - 1)\,p\,F(\mathbf{w}^*) + \lambda^*\delta < \sum_{k=1}^{N} w_k^*[(\lambda^*\gamma_k + c_k) - (p + \alpha\lambda^*)\,F_k]$$

But from the first order conditions (4.1), for each $k = 1, \ldots, N$, the term inside the square brackets in the summation is equal to zero, and hence so too is the sum. Consequently, $(\psi - 1)\,p\,F(\mathbf{w}^*) + \lambda^*\delta < 0$.

Thus, when $\psi \leq 0.5$ (i.e. the production function is highly concave), $\lambda_\alpha < 0$ and hence an increase in α will lead to a reduction in the shadow price. Of course, this is a sufficient – rather than a necessary – condition. Consequently, the expression may still be negative for values of $\psi > 0.5$. Indeed, by continuity, this will definitely be the case for values close to 0.5. However, when the value of ψ is close to one, it is possible that the expression may be positive, and hence that an increase in a may cause the price to rise. This possibility is illustrated in the following numerical example.

Numerical example 1

For the purposes of this example it is assumed that there are only two inputs, and that the aggregate production function is Cobb Douglas – taking the form $F(\boldsymbol{w}) = w_1^{0.5} w_2^{0.3}$ (i.e. $\psi = 0.8$). The exogenous output price and input prices are respectively $p = 10$ and $\mathbf{c} = (1, 3)$.

The values of input parameters and the constant term in the aggregate performance rule parameters are respectively $\gamma = (1, 0)$ and $\delta = 10$. This yields the following hybrid regulatory target:

$$\frac{w_1}{y} \leq \alpha + \frac{10}{y} \equiv \tau(y, \alpha)$$

For example, if the first input commodity is energy, then $\tau(y, \alpha)$ represents a hybrid target for energy use, with α representing the "base" rate of specific energy consumption.[9] With the chosen parameter values, the pre-regulation values of the variables are $y^u = 55.9$, $w_1^u = 279.5$ and $w_2^u = 55.9$, and the performance rule is binding for all values of $\alpha < 4.82$.

Ceteris paribus, the regulatory target becomes less stringent as the value of α is increased. This is illustrated in Figure 4.2, which shows the regulated production possibilities set (shaded grey) when the quantity of the second input is held equal to its pre-regulation optimal level (i.e. $w_2 = w_2^u$). An increase in the value of α causes the regulated production possibilities set to expand. In particular, those points which lie on the segment of the transformation frontier (B, B$'$] become allowable under the aggregate performance rule. This allows the regulated aggregate cost minimum to move closer to the pre-regulation solution (denoted by the point X on the frontier); with a resultant increase in aggregate welfare.[10]

Figure 4.2 Impact on production set of changes to α

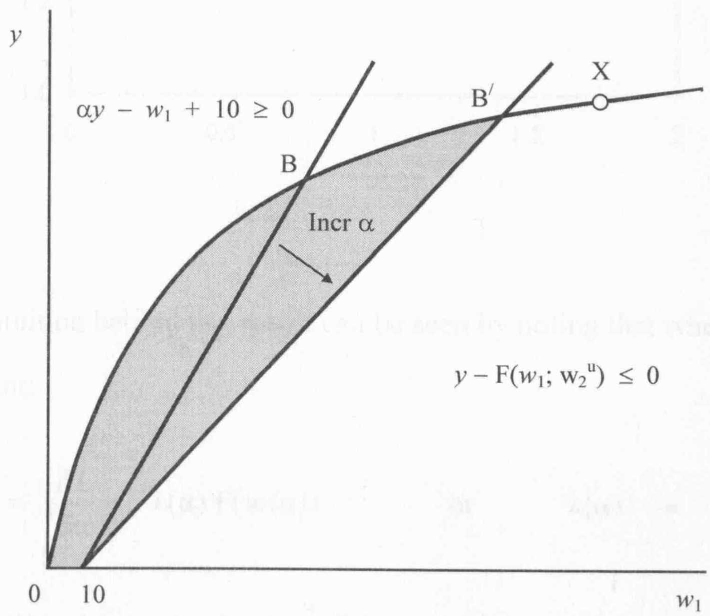
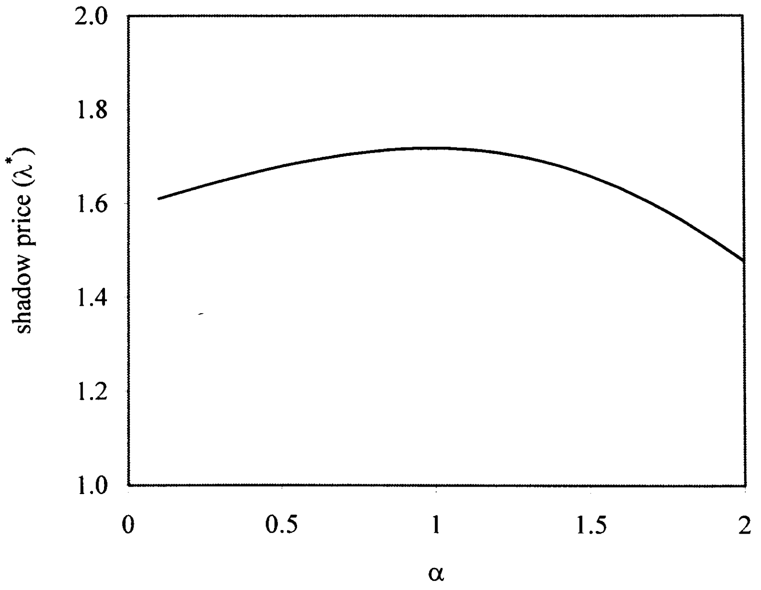


Figure 4.3 shows the relationship between the shadow price λ^* and the output parameter α, for values of $\alpha \in (0, 2]$. As can be seen, the relationship is not monotonic. For values of $\alpha > 1$, the shadow price falls as the value of α increases (as one might expect). However, for values of α below one, an increase in the value of the parameter causes the shadow price to rise.

[9] A hybrid policy target of this form is discussed in Chapter 2.

[10] Of course the quantity of the second input does not remain constant at its pre-regulation value. Consequently, the points B and B$'$ do not represent the regulated solution for the two values of r. However, the general point made by Figure 4.2 remains valid.

Figure 4.3 Relationship between shadow price and the value of α



The intuition behind this result can be seen by noting that when γ and δ are held constant:

$$\frac{dW}{d\alpha} = \frac{\partial \mathcal{L}}{\partial \alpha} = \lambda(\alpha)\, F(w(\alpha)) \qquad \text{or} \qquad \lambda(\alpha) = \frac{dW}{d\alpha}\frac{1}{F(w(\alpha))}$$

where W is the maximal value of the aggregate welfare function . Hence λ^* is equal to the change in aggregate welfare resulting form a marginal increase in α, divided by the output level. Taking the partial derivative with respect to α yields

$$\frac{d\lambda}{d\alpha} = \frac{1}{F(\mathbf{w}(\alpha))}\left[\frac{d^2W}{d\alpha^2} - \frac{dW}{d\alpha}\frac{1}{F(\mathbf{w}(\alpha))}\sum_{k=1}^{K} F_k \frac{\partial w_k}{\partial \alpha}\right]$$

which can be rearranged to give:

$$\frac{1}{\lambda}\frac{d\lambda}{d\alpha} = \frac{d^2W}{d\alpha^2}\Big/\frac{dW}{d\alpha} - \frac{1}{F(\mathbf{w}(\alpha))}\sum_{k=1}^{K} F_k \frac{\partial w_k}{\partial \alpha}$$

Thus the percentage change in the shadow value is equal to the difference between the percentage change in the rate of change of aggregate welfare and the percentage change in output. As can be seen in Table 4.1, for values of r < 1, the acceleration in welfare exceeds the growth in output, and hence the shadow value rises. However, for values of r > 1, the situation is reversed, and hence the shadow value falls.

Table 4.1 Simulation results when $\gamma = (1, 0)$ and $\delta = 10$

α	λ^*	τ^*	w_1^*	w_2^*	y^*	$\Delta y^*/y^*$	W	$\Delta^2 W/\Delta W$
0.1	1.610	1.947	10.5	5.5	5.4		27.1	
0.2	1.629	1.964	11.1	5.9	5.7	4.6%	28.0	
0.3	1.647	1.982	11.8	6.2	5.9	4.7%	29.0	5.8%
0.4	1.664	2.002	12.5	6.7	6.2	4.9%	30.0	5.9%
0.5	1.679	2.023	13.3	7.1	6.6	5.0%	31.0	5.9%
0.6	1.693	2.045	14.2	7.6	6.9	5.2%	32.2	6.0%
0.7	1.704	2.070	15.1	8.2	7.3	5.4%	33.4	6.0%
0.8	1.712	2.096	16.2	8.8	7.7	5.5%	34.7	6.0%
0.9	1.717	2.125	17.3	9.4	8.2	5.7%	36.0	6.0%
1	1.718	2.156	18.7	10.1	8.7	5.8%	37.5	5.9%
1.1	1.716	2.189	20.1	10.9	9.2	6.0%	39.0	5.9%
1.2	1.709	2.224	21.7	11.8	9.8	6.1%	40.6	5.8%
1.3	1.697	2.263	23.5	12.7	10.4	6.2%	42.3	5.6%
1.4	1.681	2.304	25.5	13.7	11.1	6.3%	44.1	5.4%
1.5	1.659	2.348	27.7	14.7	11.8	6.4%	46.0	5.2%
1.6	1.633	2.395	30.1	15.9	12.6	6.4%	48.1	5.0%
1.7	1.601	2.445	32.8	17.1	13.4	6.5%	50.2	4.7%
1.8	1.565	2.498	35.8	18.4	14.3	6.5%	52.4	4.4%
1.9	1.524	2.555	39.0	19.7	15.3	6.5%	54.6	4.0%
2	1.479	2.614	42.6	21.1	16.3	6.4%	57.0	3.6%

4.3 Changes in input parameter γ_j

If the values of α and δ, and all γ_k with the exception of $k = \mathrm{j}$, are held constant, then the total differential (4.3) simplifies to:

$$d\lambda^* = \lambda_{\gamma_j} d\gamma_j.$$

Partially differentiating the identities (4.4) and (4.5) with respect to β_j yields the following system of equations:

$$\left[\begin{array}{c|ccc} 0 & \mathcal{L}_{\lambda 1} & \cdots & \mathcal{L}_{\lambda N} \\ \hline \mathcal{L}_{1\lambda} & \mathcal{L}_{11} & \cdots & \mathcal{L}_{1N} \\ \vdots & \vdots & & \\ \mathcal{L}_{N\lambda} & \mathcal{L}_{N1} & & \mathcal{L}_{NN} \end{array}\right] \begin{bmatrix} \dfrac{\partial \lambda}{\partial \gamma_j} \\ \dfrac{\partial w_1}{\partial \gamma_j} \\ \vdots \\ \dfrac{\partial w_N}{\partial \gamma_j} \end{bmatrix} = \begin{bmatrix} -\mathcal{L}_{\lambda\gamma_j} \\ -\mathcal{L}_{1\gamma_j} \\ \vdots \\ -\mathcal{L}_{N\gamma_j} \end{bmatrix}$$

where	$\mathcal{L}_{kl}$	=	$(p + \lambda^*\alpha)\, F_{kl}(\boldsymbol{w}^*)$	for $k,l = 1, \ldots, N$
	$\mathcal{L}_{k\lambda}$	=	$\alpha\, F_k(\boldsymbol{w}^*) - \gamma_k$	for $k = 1, \ldots, N$
	$\mathcal{L}_{j\gamma_j}$	=	$-\lambda^*$	
	$\mathcal{L}_{k\gamma_j}$	=	0	for $k = 1, \ldots, N$ and $k \neq j$
	$\mathcal{L}_{\lambda\gamma_j}$	=	$-w_j^*$	

By Cramer's rule:

$$\frac{\partial \lambda}{\partial \gamma_j} \equiv \lambda_{\gamma_j} = \frac{A}{G} \qquad \ldots (4.9)$$

where the determinant G is defined in section 4.2, and

$$A = \begin{vmatrix} -\mathcal{L}_{\lambda\beta_j} & \mathcal{L}_{\lambda 1} & \cdots & \mathcal{L}_{\lambda N} \\ \hline -\mathcal{L}_{1\beta_j} & \mathcal{L}_{11} & \cdots & \mathcal{L}_{1N} \\ \vdots & \vdots & & \\ -\mathcal{L}_{N\beta_j} & \mathcal{L}_{N1} & & \mathcal{L}_{NN} \end{vmatrix} = \begin{vmatrix} w_j^* & \mathcal{L}_{\lambda 1} & \cdots & \mathcal{L}_{\lambda N} \\ \hline 0 & \mathcal{L}_{11} & \cdots & \mathcal{L}_{1N} \\ \vdots & \vdots & & \\ \lambda^* & \mathcal{L}_{j1} & & \mathcal{L}_{jN} \\ \vdots & \vdots & & \\ 0 & \mathcal{L}_{N1} & & \mathcal{L}_{NN} \end{vmatrix}$$

In appendix A4.2, the following expression is derived for the determinant of matrix **A**:

$$A = (p + \lambda^*\alpha)^{N-1} \left[w_j^* H \left(p + \left(1 + \frac{1}{1-\psi} \right) \lambda^* \alpha \right) + \lambda^* \left(\sum_{k=1}^{N} (-1)^{j+k} \gamma_j H_{jk} \right) \right] \quad \dots (4.10)$$

where H is the determinant of the Hessian matrix for $F(\boldsymbol{w})$, and H_{jk} are the (un-signed) cofactors of H.

Together, (4.9) and (4.10) imply that:

$$\lambda_{\gamma_j} = \frac{(p + \lambda^*\alpha)^{N-1}}{G} \left[w_j^* H \left(p + \left(1 + \frac{1}{1-\psi} \right) \lambda^* \alpha \right) + \lambda^* \left(\sum_{k=1}^{N} (-1)^{j+k} \gamma_j H_{jk} \right) \right]$$

Unfortunately, there is not much that can be said at this level of generality. If $\alpha \geq 0$, then the sign of the first term in square brackets is known (i.e. it has the same sign as H). However, if $\alpha < 0$ then the sign depends on the magnitude of λ^*. Furthermore, it is not possible to determine the sign of the second term *a priori*.

Therefore, in order to proceed, it will be assumed that there are only two input commodities (i.e. N = 2), and that the output parameter and the constant term in the performance rule are both equal to zero (i.e. $\alpha = 0$ and $\delta = 0$). Under these assumptions the performance rule simplifies to $\gamma_1 w_1 + \gamma_2 w_2 = 0$, and hence:

$$A = p\,[\,p\,w_2^*\,H - \lambda^*(\gamma_1 H_{21} - \gamma_2 H_{22})\,]$$

$$= p\,[\,p\,w_2^*\,(F_{11}F_{22} - F_{12}F_{21}) - \lambda^*(\gamma_1 F_{12} - \gamma_2 F_{11})\,]$$

$$= p\,[\,p\,w_2^*\,(F_{11}F_{22} - F_{12}F_{21}) + \lambda^*\gamma_2((-\gamma_1/\gamma_2)F_{12} + F_{11})\,]$$

$$= p^2 w_2^*\,(F_{11}F_{22} - F_{12}F_{21}) + \frac{\gamma_2\lambda^* p}{w_1^*}\left(w_1^* F_{11} + w_2^* F_{12}\right)$$

$$= p^2 w_2^*\,(F_{11}F_{22} - F_{12}F_{21}) + \frac{\gamma_2\lambda^* p(\psi-1)F_1}{w_1^*}$$

$$= \frac{p\,(1-\psi)\,F_1}{w_1^*}\left[\,pF_2\left(\frac{w_1^* w_2^*(F_{11}F_{22} - F_{12}F_{21})}{(1-\psi)F_1F_2}\right) - \gamma_2\lambda^*\right]$$

From the first order conditions (4.1), it follows that $pF_2 - \gamma_2\lambda^* = c_2 > 0$. Therefore, a sufficient condition for the expression inside the square brackets to be positive is that:

$$w_1^* w_2^*\,(F_{11}F_{22} - F_{12}F_{21}) \geq (1-\psi)\,F_1F_2 \qquad \ldots (4.11)$$

However, since $F(\boldsymbol{w})$ is homogeneous of degree ψ, it follows that

$(\psi - 1)\,F_k = w_1 F_{k1} + w_2 F_{k2}$ (for $k = 1,2$),

and hence that:

$$(\psi-1)^2 F_1F_2 = w_1w_2\,(F_{11}F_{22} - F_{12}F_{21}) + w_1 F_{21}(\psi-1)F_1 + w_2 F_{12}(\psi-1)F_2$$

Consequently, the sufficient condition (4.11) can be re-stated as:

$$(\psi-1)^2 F_1F_2 - w_1^* F_{21}(\psi-1)F_1 - w_2^* F_{12}(\psi-1)F_2 \geq (1-\psi)F_1F_2$$

Dividing by $(\psi - 1)$, canceling terms and re-arranging yields:

$$w_1^* F_{21} F_1 + w_2^* F_{12} F_2 \geq \psi F_1 F_2 \qquad \ldots (4.12)$$

Thus, if condition (4.12) is satisfied then the determinant A is positive, and since G is positive when K = 2, so too is λ_{γ_1}. Unfortunately however, this condition is not guaranteed to be satisfied for all homogeneous functions of degree $\psi < 1$. For example, if $F(w) = (w_1^\rho + w_2^\rho)^{\psi/\rho}$ (with $\rho \leq 1$), then $w_1 F_{21} F_1 + w_2 F_{12} F_2 = (\psi - \rho) F_1 F_2$, and hence the sufficient condition is only satisfied for values of $\rho \leq 0$. Of course, this is not a necessary condition, so A may still be positive for values of $\rho > 0$. However, as the following example illustrates, when the value of ρ is close to one, the sign of A, and hence that of λ_{γ_1} may change as the value of γ_2 increases.

Numerical Example 2

For the purposes of this example, it is assumed that the CES production function parameters take the values $\psi = 0.2$ and $\rho = 0.9$, and the exogenous output price and input prices are respectively p = 10 and $\mathbf{c} = (1, 2)$. The performance rule parameters are $\alpha = 0$, $\gamma_2 = -1$ and $\delta = 0$, which yields the following regulatory target:

$$\frac{w_2}{w_1 + w_2} \geq \frac{\gamma_1}{1 + \gamma_1} \equiv \tau(\gamma_1)$$

For example, if the two input commodities are virgin content newsprint and recycled content newsprint, then $\tau(\gamma_1)$ represents a standard for the proportion of newsprint made from recycled pulp. With the chosen parameter values, the pre-regulation share of recycled newsprint is 0.1%. Hence, $\lambda^* > 0$ for any target $\tau(\gamma_1)$ that exceeds this value.

Since $\tau'(\gamma_1) > 0$, the target becomes more stringent as the value of γ_1 is increased. This is illustrated in Figure 4.4, which shows the input requirement set when output is held constant at the pre-regulation level (i.e. $y = y^u$). An increase in the value of γ_1 causes the constraint set to contract, so that points on the segment of the input requirement frontier [B, B/) are no longer allowable under the performance rule. This forces the regulated solution to move further away from the pre-regulation solution (denoted by the point X on the frontier); with a resultant decrease in aggregate welfare.[11]

Figure 4.4 Impact on input requirement set of changes to γ

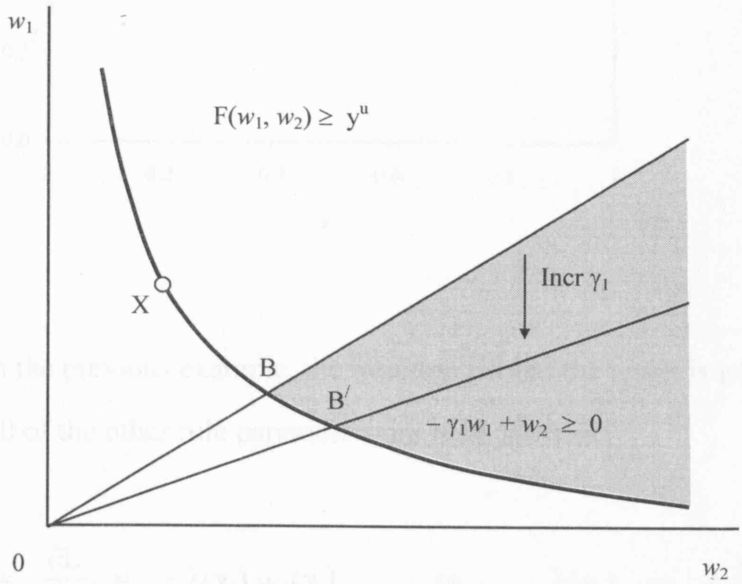
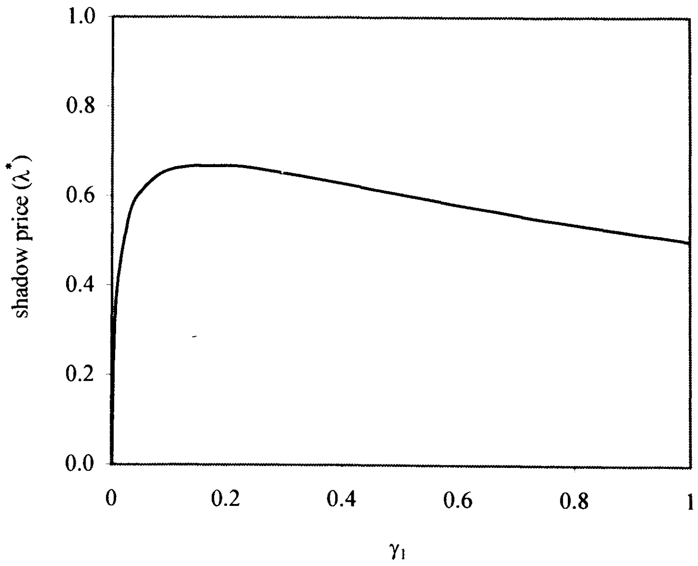


Figure 4.5 shows the relationship between the shadow price λ^* and γ_1, for values of $\gamma_1 \in (0, 1]$. As with the previous example, the relationship is not monotonic. For values of $\gamma_1 < 0.2$, the shadow price rises as the parameter value increases. However, for values of $\gamma_1 > 0.2$, an increase in the value of γ_1 causes the price to decline. Thus, the shadow price initially rises as the target proportion $\tau(\gamma_1)$ increases. But once the target reaches c.16%, further increases cause the price to fall.

[11] Of course, the output quantity does not remain constant at the pre-regulation level as the value of r changes. Consequently, the points B and $B^{/}$ do not represent the actual solutions under the performance rule. However, as with the first numerical example, the point made by Figure 4.4 remains valid.

Figure 4.5 Relationship between shadow price and γ_1



As with the previous example, the intuition behind the result is provided by noting that when all of the other rule parameters are held constant:

$$\frac{dW}{d\gamma_1} = \frac{\partial L}{\partial \gamma_1} = -\lambda(\gamma_1)\, w_1(\gamma_1) \qquad \text{or} \qquad \lambda(\gamma_1) = -\frac{dW}{d\gamma_1}\frac{1}{w_1(\gamma_1)}$$

Hence λ^* is equal to the change in aggregate welfare resulting form a marginal increase in γ_1, divided by the input quantity of commodity 1. Taking the derivative with respect to γ_1 yields:

$$\frac{d\lambda}{d\gamma_1} = -\frac{1}{w_1(\gamma_1)}\left[\frac{d^2W}{d\gamma_1^2} - \frac{dW}{d\gamma_1}\frac{1}{w_1(\gamma_1)}\frac{dw_1}{d\gamma_1}\right]$$

which can be rearranged to give:

$$\frac{1}{\lambda}\frac{d\lambda}{d\gamma_1} = \frac{d^2W}{d\gamma_1^2}\Big/\frac{dW}{d\gamma_1} - \frac{1}{w_1(\gamma_1)}\frac{dw_1}{d\gamma_1}$$

Thus the percentage change in the shadow price is equal to the difference between the percentage change in the rate of change of profit and the percentage change in quantity of commodity 1 that is used in production. The second term is negative for all values of γ_1. However, as can be seen in Table 4.2, the first term changes sign as the value of γ_1 increases. For values of $\gamma_1 \leq 0.2$, it is either positive, or smaller in magnitude than the second term. Consequently, the overall expression is positive. However, for values of $\gamma_1 > 0.2$ the first term is dominant, and hence the shadow price declines as the value of the parameter increases.

Table 4.2 Simulation results when $\alpha = 0$, $\gamma_2 = -1$ and $\delta = 0$

γ_1	λ^*	τ^*	w_1^*	w_2^*	y^*	$\Delta w_1^*/w_1^*$	W	$\Delta^2W/\Delta W$
0.001	0.000	0.1%	2.37	0.00	1.19		9.514	
0.002	0.138	0.2%	2.37	0.00	1.19	-201%	9.514	
0.003	0.211	0.3%	2.36	0.01	1.19	-204%	9.514	31684%
0.004	0.261	0.4%	2.36	0.01	1.19	-206%	9.513	26531%
0.005	0.299	0.5%	2.35	0.01	1.19	-205%	9.512	9573%
0.01	0.409	1.0%	2.33	0.02	1.19	-202%	9.508	1471%
0.03	0.556	2.9%	2.24	0.07	1.19	-202%	9.485	462%
0.05	0.610	4.8%	2.15	0.11	1.18	-186%	9.460	28%
0.1	0.658	9.1%	1.96	0.20	1.17	-168%	9.394	-105%
0.2	0.668	16.7%	1.66	0.33	1.16	-158%	9.274	-134%
0.3	0.652	23.1%	1.43	0.43	1.15	-137%	9.172	-166%
0.4	0.629	28.6%	1.26	0.50	1.14	-121%	9.086	-160%
0.5	0.604	33.3%	1.13	0.56	1.13	-109%	9.013	-149%
0.6	0.581	37.5%	1.02	0.61	1.12	-98%	8.949	-139%
0.7	0.559	41.2%	0.93	0.65	1.11	-90%	8.894	-129%
0.8	0.538	44.4%	0.85	0.68	1.11	-83%	8.845	-121%
0.9	0.518	47.4%	0.79	0.71	1.10	-76%	8.802	
1	0.500	50.0%	0.73	0.73	1.10		8.764	

4.4 Impact of changes to the constant term δ

If α and γ are held constant then the total differential (4.3) simplifies to:

$$d\lambda^* = \lambda_\delta \, d\delta.$$

Partially differentiating the identities (4.4) and (4.5) with respect to δ yields the following system of equations:

$$\left[\begin{array}{c|cccc} 0 & \mathcal{L}_{\lambda 1} & \cdots & & \mathcal{L}_{\lambda N} \\ \hline \mathcal{L}_{1\lambda} & \mathcal{L}_{11} & \cdots & & \mathcal{L}_{1N} \\ \vdots & \vdots & & & \\ & & & & \\ \mathcal{L}_{N\lambda} & \mathcal{L}_{N1} & & & \mathcal{L}_{NN} \end{array}\right] \begin{bmatrix} \dfrac{\partial\lambda}{\partial\delta} \\ \dfrac{\partial w_1}{\partial\delta} \\ \vdots \\ \dfrac{\partial w_N}{\partial\delta} \end{bmatrix} = \begin{bmatrix} -\mathcal{L}_{\lambda\delta} \\ -\mathcal{L}_{1\delta} \\ \vdots \\ \\ -\mathcal{L}_{N\delta} \end{bmatrix}$$

where	L_{kl}	$=$	$(p + \lambda^*\alpha)\, F_{kl}(\mathbf{w}^*)$	for $k,l = 1, \ldots, N$
	$L_{k\lambda}$		$\alpha\, F_k(\mathbf{w}^*) - \gamma_k$	for $k = 1, \ldots, N$
	$L_{k\delta}$	$=$	0	for $k = 1, \ldots, N$
	$L_{\lambda\delta}$	$=$	1	

By Cramer's rule:

$$\frac{\partial\lambda}{\partial\delta} \equiv \lambda_\delta = \frac{A}{G} \qquad \ldots (4.13)$$

where G is defined in section 4.2, and:

$$A = \left|\begin{array}{c|cccc} -\mathcal{L}_{\lambda\delta} & \mathcal{L}_{\lambda 1} & \cdots & & \mathcal{L}_{\lambda N} \\ \hline -\mathcal{L}_{1\delta} & \mathcal{L}_{11} & \cdots & & \mathcal{L}_{1N} \\ \vdots & \vdots & & & \\ & & & & \\ -\mathcal{L}_{N\delta} & \mathcal{L}_{N1} & & & \mathcal{L}_{NN} \end{array}\right| = \left|\begin{array}{c|cccc} -1 & \mathcal{L}_{\lambda 1} & \cdots & & \mathcal{L}_{\lambda N} \\ \hline 0 & \mathcal{L}_{11} & \cdots & & \mathcal{L}_{1N} \\ \vdots & \vdots & & & \\ & & & & \\ 0 & \mathcal{L}_{N1} & & & \mathcal{L}_{NN} \end{array}\right|$$

It follows directly that:

$$A \quad = \quad -(p + \lambda^{*}\alpha)^{N} H \qquad \ldots (4.14)$$

where H is the determinant of the Hessian matrix for $F(\boldsymbol{w})$.

Together, (4.13) and (4.14) imply that:

$$\lambda_{\delta} = \quad -\left(\frac{H}{G}\right)(\acute{p} + \lambda^{*}\alpha)^{N} \quad < \quad 0$$

Thus an increase in the value of the constant δ will always lead to a reduction in the shadow price λ^{*}, irrespective of the values of the other performance rule parameters.

4.5 Summary

This chapter has explored the relationship between the shadow value of the aggregate performance rule (λ^{*}) and the values of the rule parameters ($\alpha \mid \beta \mid \chi \mid \delta$). However, given the direct relationships between the shadow value and the equilibrium price of performance credits, and between the performance rule parameters and the values of the regulatory control variables (r and L), the findings can be interpreted in terms of the impact of these variables on the equilibrium price of performance credits.

The model that underpins the analysis has a more restrictive structure than the general model used in Chapter 3 to establish the cost efficiency and distributional flexibility of performance-based credit trading. As a result, it is not applicable to all forms of regulatory target. In particular, it cannot accommodate aggregate performance rules in which the output parameters (α), or the input parameters (β) represent linear combinations of a common regulatory control variable and commodity-specific scaling parameters, as is the case – for example – with the "weighted average" performance standard for vehicle fuel efficiency (where $\alpha_k = r - a_k$).[12] However, the model can accommodate most other common forms of regulatory target, including the following applications that were highlighted in Chapter 2:

(i) an absolute emissions limit $\quad c\,\zeta' \boldsymbol{w} \leq L$

(ii) a relative specific energy standard $\quad \dfrac{\mathbf{b}'\boldsymbol{w}}{y} \leq r$

(iii) a hybrid emissions target $\quad \dfrac{c\,\zeta' \boldsymbol{w}}{y} \leq r + \dfrac{L}{y}$

(iv) a relative recycled newsprint standard $\quad \dfrac{w_1}{w_1 + w_2} \geq r$

The analyses in the preceding sections have considered, in turn, the impacts of small changes to the output parameter (α), to an input parameter (β_k), and to the constant term (δ) in the aggregate performance rule. However, rather than reprise the results of these parameter-related analyses, it is more interesting – and more relevant – to use these four applications to illustrate the insights that the analyses provide regarding the relationship between the price of performance credits and the stringency of the regulatory target.

In the first two applications, the implications of the analysis are unambiguous. An increase in the stringency of the regulatory target – via a reduction in the value of the respective control variable – leads to a rise in the equilibrium price performance credits. This is in line with intuition, and (in the first case) is consistent with the impact of a reduction in the total number of permits that are issued in a "cap and trade" permit system.

[12] An exception is when the weighted average has only two components. However, as was noted in Chapter 1, in this case the regulatory target can always be reformulated in terms of a proportion.

In the other two applications however, the relationship is not so clear-cut. For example, in the case of the hybrid emissions target, the price of performance credits unambiguously increases if the value of the "limit control variable" (L) is reduced. However, if the value of the "rate control variable" (r) is reduced, then the impact depends the on the shape of the aggregate production function, and the current value of the variable. In particular, if returns to scale for the aggregate production are close to being constant, and the current value of the variable is relatively low (i.e. the target is already relatively stringent), then it is possible that a further decrease may actually cause the price of performance credits to fall.

A similar response can also occur in the final application – the recycled content standard for newsprint. In this case, the impact of an increase in the value of the target rate (r) depends on the elasticity of substitution between the two input commodities. If this is less than or equal to one, then the price of performance credits unambiguously increases. However, if it significantly greater than one (i.e. the commodities are very close substitutes), and the target rate is relatively high, then it is possible that the price of credits may fall.

Appendix A4.1 Derivation of expression (4.7) for determinant of A

The derivation makes use of the following properties of the determinants of the Hessian and bordered Hessian of the aggregate production function F(*w*) – denoted respectively by H and B; and of the determinant A. Property P1 is a standard property for any square matrix.[13] The derivations of Properties P2 and P3 are provided in Appendix A4.3, and the derivation of P4(a) is provided in Appendix A4.4.

- *Property P1*

$$\mathrm{H} = \sum_{k=1}^{\mathrm{N}} \mathrm{F}_{kl}\,(-1)^{k+l}\,\mathrm{H}_{kl} = \sum_{l=1}^{\mathrm{N}} \mathrm{F}_{kl}\,(-1)^{k+l}\,\mathrm{H}_{kl}$$

$$= (-1)^{l+1} \sum_{k=1}^{\mathrm{N}} (-1)^{k-1}\,\mathrm{F}_{kl}\,\mathrm{H}_{kl} = (-1)^{k+1} \sum_{l=1}^{\mathrm{N}} (-1)^{l-1}\,\mathrm{F}_{kl}\,\mathrm{H}_{kl}$$

where H_{kl} are the (unsigned) cofactors of **H**

- *Property P2*

$$\mathrm{B}_{11} = \mathrm{H}$$

$$\mathrm{B}_{1l+1} = \sum_{k=1}^{\mathrm{N}} (-1)^{k-1}\,\mathrm{F}_k\,\mathrm{H}_{kl} \qquad \text{for all } l = 1, \dots, \mathrm{N}$$

where B_{11}, B_{1l+1} are the (unsigned) cofactors of the first row of **B**.

- *Property P3*

$$\sum_{l=1}^{\mathrm{N}} (-1)^{k+l}\,\mathrm{F}_l\,\mathrm{H}_{kl} = \frac{w_k\,\mathrm{H}}{\psi - 1} \qquad \text{for all } k = 1, \dots, \mathrm{N}$$

$$\sum_{k=1}^{\mathrm{N}} (-1)^{k+l}\,\mathrm{F}_k\,\mathrm{H}_{kl} = \frac{w_l\,\mathrm{H}}{\psi - 1} \qquad \text{for all } l = 1, \dots, \mathrm{N}$$

[13] See for example Theorem 1, page 102 of *Silberberg & Suen* (2001).

- *Property P4(a)*

$$A_{11} = (p + \lambda^* \alpha)^N B_{11},$$

$$A_{1k} = -\lambda^* (p + \lambda^* \alpha)^{N-1} B_{1k} \qquad \text{for } k = 2, \dots, N$$

where A_{11}, A_{1k} are the (unsigned) cofactors of the first row of matrix **A**

Then:

$$A = \left| \begin{array}{c|ccc} -\mathcal{L}_{\lambda\alpha} & \mathcal{L}_{\lambda 1} & \cdots & \mathcal{L}_{\lambda N} \\ \hline -\mathcal{L}_{1\alpha} & \mathcal{L}_{11} & \cdots & \mathcal{L}_{1N} \\ \vdots & \vdots & & \\ & & & \\ -\mathcal{L}_{N\alpha} & \mathcal{L}_{N1} & & \mathcal{L}_{NN} \end{array} \right|$$

$$= -\mathcal{L}_{\lambda\alpha} A_{11} - \sum_{k=1}^{N} (-1)^{k-1} \mathcal{L}_{\lambda k} A_{1k+1}$$

$$= \lambda^*(p + \lambda^* \alpha)^{N-1} \sum_{k=1}^{N} (-1)^{k-1} \left(\alpha F_k(\mathbf{w}^*) - \gamma_k\right) B_{1k+1} - F(\mathbf{w}^*) (p + \lambda^* \alpha)^N B_{11}$$

(by property P4(a))

$$= \lambda^*(p + \lambda^* \alpha)^{N-1} \sum_{k=1}^{N} (-1)^{k-1} (\alpha F_k - \gamma_k) \left[\sum_{j=1}^{N} (-1)^{j-1} F_j H_{jk} \right] - F(\mathbf{w}^*) (p + \lambda^* \alpha)^N B_{11}$$

(by property P2)

$$= \frac{\lambda^*(p + \lambda^* \alpha)^{N-1}}{\psi - 1} \left[\sum_{k=1}^{N} \left(\alpha F_k - \gamma_k\right) w_k^* H \right] - F(\mathbf{w}^*) (p + \lambda^* \alpha)^N B_{11}$$

(by property P3)

$$= \frac{\lambda^*(p+\lambda^*\alpha)^{N-1} H}{\psi - 1}\left[\psi\,\alpha\, F(\mathbf{w}^*) - \sum_{k=1}^{N} \gamma_k w_k^*\right] - F(\mathbf{w}^*)\,(p+\lambda^*\alpha)^N B_{11}$$

(since F(w) is homogeneous)

$$= \frac{\lambda^*(p+\lambda^*\alpha)^{N-1} H}{\psi - 1}\left[(\psi - 1)\,\alpha\, F(\mathbf{w}^*) - \delta\right] - F(\mathbf{w}^*)\,(p+\lambda^*\alpha)^N B_{11}$$

(since $\boldsymbol{\gamma}\mathbf{w}^* = \alpha\, F(\mathbf{w}^*) + \delta$)

$$= \left(\frac{-1}{\psi-1}\right)(p+\lambda^*\alpha)^N H\,[\,(\psi-1)\,(p+\lambda^*\alpha)\,F(\mathbf{w}^*) - \lambda^*(\psi-1)\,\alpha\,F(\mathbf{w}^*) + \lambda^*\delta\,]$$

(by property P2)

$$= \left(\frac{-1}{\psi-1}\right)(p+\lambda^*\alpha)^{N-1} H\,[\,(\psi-1)\,p\,F(\mathbf{w}^*) + \lambda^*\delta\,]$$

(QED)

Appendix A4.2 Derivation of expression (4.10) for determinant A

The derivation makes use properties P1-P3 of the determinants of the Hessian and bordered Hessian of the aggregate production function F(*w*) – denoted respectively by H and B (see Appendix A4.1); and of the following property of the determinant A, which is derived in Appendix A4.4.

- *Property P4(b)*

$$A_{11} = (p + \lambda^*\alpha)^N B_{11}$$

$$A_{1k} = A_{1k} = (-1)^{j+1} \lambda^* (p + \lambda^*\alpha)^{N-1} H_{jk-1} \qquad \text{for } k = 2, \dots, N$$

where A_{11}, A_{1k} are the (unsigned) cofactors of the first row of matrix **A**

Then

$$A = \begin{vmatrix} w_j^* & \mathcal{L}_{\lambda 1} & \cdots & \mathcal{L}_{\lambda N} \\ 0 & \mathcal{L}_{11} & \cdots & \mathcal{L}_{1N} \\ \vdots & \vdots & & \\ \lambda^* & \mathcal{L}_{j1} & & \mathcal{L}_{jN} \\ \vdots & \vdots & & \\ 0 & \mathcal{L}_{N1} & & \mathcal{L}_{NN} \end{vmatrix}$$

$$= w_j^* A_{11} - \sum_{i=1}^{N} (-1)^{i-1} \mathcal{L}_{\lambda i} A_{1i+1}$$

$$= w_j^* (p + \lambda^*\alpha)^N H - (-1)^{j+1} \lambda^* (p + \lambda^*\alpha)^{N-1} \sum_{k=1}^{N} (-1)^{k-1} \left(\alpha F_k(\mathbf{w}^*) - \gamma_k\right) H_{jk}$$

(by property P4(b))

$$= \quad w_j^* (p+\lambda^*\alpha)^N H \quad - \quad \lambda^*\alpha\,(p+\lambda^*\alpha)^{N-1} \sum_{k=1}^{N} (-1)^{j+k}\, F_k(\mathbf{w}^*)\; H_{jk}$$

$$+ \quad \lambda^* (p+\lambda^*\alpha)^{N-1} \sum_{k=1}^{N} (-1)^{j+k}\, \gamma_k\; H_{jk}$$

$$= \quad w_j^* (p+\lambda^*\alpha)^N H \quad - \quad \lambda^*\alpha\,(p+\lambda^*\alpha)^{N-1} \frac{w_j^* H}{\psi - 1}$$

$$+ \quad \lambda^* (p+\lambda^*\alpha)^{N-1} \sum_{k=1}^{N} (-1)^{j+k}\, \gamma_k\; H_{jk}$$

(by property P3)

$$= \quad \frac{w_j^* H\,(p+\lambda^*\alpha)^{N-1}}{\psi - 1}\left[(\psi-1)(p+\lambda^*\alpha) \quad - \quad \lambda^*\alpha\right]$$

$$+ \quad \lambda^* (p+\lambda^*\alpha)^{N-1} \sum_{k=1}^{N} (-1)^{j+k}\, \gamma_k\; H_{jk}$$

$$= \quad (p+\lambda^*\alpha)^{N-1}\left[\; w_j^* H\left(p + \left(1 + \frac{1}{1-r}\right)\lambda^*\alpha\right) + \lambda^*\left(\sum_{k=1}^{N} (-1)^{j+k}\, \gamma_k\; H_{jk}\right)\right]$$

(QED)

Appendix A4.3 Derivation of properties P2 – P3

Noting that:

$$
H = \begin{vmatrix} F_{11} & \cdots & F_{1N} \\ \vdots & & \\ & & \\ F_{N1} & & F_{NN} \end{vmatrix}
\qquad
B = \left|\begin{array}{c|ccc} 0 & F_1 & \cdots & F_N \\ \hline F_1 & F_{11} & \cdots & F_{1N} \\ \vdots & \vdots & & \\ & & & \\ F_N & F_{N1} & & F_{NN} \end{array}\right|
$$

Property P2

$$
B_{11} = \begin{vmatrix} F_{11} & \cdots & F_{1N} \\ \vdots & & \\ & & \\ F_{N1} & & F_{NN} \end{vmatrix}
$$

$$
= H
$$

$$
B_{1l+1} = \begin{vmatrix} F_1 & F_{11} & \cdots & F_{1l-1} & F_{1l+1} & \cdots & F_{1N} \\ F_2 & F_{21} & \cdots & F_{2l-1} & F_{2l+1} & & F_{2N} \\ \vdots & \vdots & & & & & \\ & & & & & & \\ F_N & F_{N1} & & F_{Nl-1} & F_{Nl+1} & & F_{NN} \end{vmatrix} \qquad \text{for all } l = 1, \ldots, N
$$

$$
= \sum_{k=1}^{N} (-1)^{k-1} F_k H_{kl} \qquad \text{for all } l = 1, \ldots, N
$$

where B_{11}, B_{1l+1} are the (unsigned) cofactors of the first row of **B**

(QED)

Property P3

If $F(\boldsymbol{w})$ is homogeneous of degree ψ, then

$$\sum_{k=1}^{N} (-1)^{k+l} F_k H_{kl} = \frac{1}{\psi-1} \sum_{k=1}^{N} (-1)^{k+l} H_{kl} \left[\sum_{m=1}^{N} w_m F_{km} \right]$$

$$= \frac{1}{\psi-1} \sum_{m=1}^{N} w_m \left[\sum_{k=1}^{N} (-1)^{k+l} F_{km} H_{kl} \right]$$

$$= \frac{w_l H}{\psi-1}$$

since $\sum_{k=1}^{N} (-1)^{k+l} F_{km} H_{kl} = 0$ if $l \neq m$ (expansion by alien co-factors)

$= H$ if $l = m$ (by Property P1)

where H_{kl} are the (unsigned) cofactors of $\mathbf{H}$.

Similarly

$$\sum_{l=1}^{N} (-1)^{k+l} F_l H_{kl} = \frac{1}{\psi-1} \sum_{l=1}^{N} (-1)^{k+l} H_{kl} \left[\sum_{m=1}^{N} w_m F_{lm} \right]$$

$$= \frac{1}{\psi-1} \sum_{m=1}^{N} w_m \left[\sum_{l=1}^{N} (-1)^{k+l} F_{lm} H_{kl} \right]$$

$$= \frac{w_k H}{\psi-1}$$

(QED)

Appendix A4.4 Derivation of properties P4(a) and P4(b)

Property 4(a)

Noting that

$$A = \left| \begin{array}{c|ccc} -L_{\lambda\alpha} & L_{\lambda 1} & \cdots & L_{\lambda N} \\ \hline -L_{1\alpha} & L_{11} & \cdots & L_{1N} \\ \vdots & \vdots & & \\ -L_{N\alpha} & L_{N1} & & L_{NN} \end{array} \right|$$

Then

$$A_{11} = \begin{vmatrix} L_{11} & \cdots & L_{1N} \\ \vdots & & \\ L_{N1} & & L_{NN} \end{vmatrix}$$

$$= (p + \lambda^{*}\alpha)^{N} \begin{vmatrix} F_{11} & \cdots & F_{1N} \\ \vdots & & \\ F_{N1} & & F_{NN} \end{vmatrix} \qquad \text{since} \quad L_{kl} = (p + \lambda^{*}\alpha)\, F_{kl}$$

$$= (p + \lambda^{*}\alpha)^{N}\, B_{11}$$

$$A_{12} = \begin{vmatrix} -L_{1\alpha} & L_{12} & \cdots & L_{1N} \\ \vdots & & & \\ -L_{N\alpha} & L_{N2} & & L_{NN} \end{vmatrix}$$

$$= -\lambda^{*} (p + \lambda^{*}\alpha)^{N-1} \begin{vmatrix} F_{1} & F_{12} & \cdots & F_{1N} \\ \vdots & & & \\ F_{N} & F_{N2} & & F_{NN} \end{vmatrix}$$

$$= \quad - \lambda^* (p + \lambda^*\alpha)^{N-1} B_{12}$$

$\vdots$

$$A_{1i+1} = \quad - \lambda^* (p + \lambda^*\alpha)^{N-1} B_{1i+1}$$

where A_{11}, A_{12}, … etc are the (unsigned) cofactors of the first row of **A**.

(QED)

Property 4(b)

Noting that

$$A = \left| \begin{array}{c|ccc} -L_{\lambda\beta_i} & L_{\lambda 1} & \cdots & L_{\lambda N} \\ \hline -L_{1\beta_i} & L_{11} & \cdots & L_{1N} \\ \vdots & \vdots & & \\ & & & \\ -L_{N\beta_i} & L_{N1} & & L_{NN} \end{array} \right|$$

Then

$$A_{11} = \begin{vmatrix} L_{11} & \cdots & L_{1N} \\ \vdots & & \\ & & \\ L_{N1} & & L_{NN} \end{vmatrix}$$

$$= (p + \lambda^*\alpha)^N \begin{vmatrix} F_{11} & \cdots & F_{1N} \\ \vdots & & \\ & & \\ F_{N1} & & F_{NN} \end{vmatrix}$$

$$= (p + \lambda^*\alpha)^N B_{11}$$

$$A_{12} = \begin{vmatrix} 0 & L_{12} & \cdots & L_{1N} \\ \vdots & & & \\ \lambda^* & L_{j2} & & L_{jN} \\ \vdots & & & \\ 0 & L_{N2} & & L_{NN} \end{vmatrix}$$

$$= (p + \lambda^*\alpha)^{N-1} \begin{vmatrix} 0 & F_{12} & \cdots & F_{1N} \\ \vdots & & & \\ \lambda^* & F_{j2} & & F_{jN} \\ \vdots & & & \\ 0 & F_{N2} & & F_{NN} \end{vmatrix}$$

$$= (-1)^{j+1}\ \lambda^*\ (p + \lambda^*\alpha)^{N-1}\ H_{j1}$$

$\vdots$

$$A_{1i+1} = (-1)^{j+1}\ \lambda^*\ (p + \lambda^*\alpha)^{N-1}\ H_{ji}$$

where A_{11}, A_{12}, ... etc. are the (unsigned) cofactors of the first row of **A.**

(QED)

Chapter 5 Implementation of an industrial energy efficiency standard

In the previous two chapters, the core theoretical properties of performance-based credit trading have been established within a relatively generic framework. In this chapter, and the next, the properties of the mechanism are explored further in the context of two specific applications. Not only does this allow some of the general properties to be illustrated, it also enables some additional insights to be gained. In this chapter the policy context is industrial energy efficiency; with performance-based credit trading being used to implement a target rate is set for aggregate specific energy consumption (i.e. energy consumption per unit of output) for a particular sector, or a group of sectors.

In order to make the application more concrete, it will be assumed that the regulatory intervention – and hence the trading scheme – applies to the chemical sector. However, it should be stressed that this is done only to elucidate the discussion of the impacts of the mechanism, and some of the issues that can arise. It is not an objective of this chapter to provide a detailed analysis of the chemicals sector. A corollary of this is that the various results that are derived are not specific to the chemicals sector. In particular, they would be equally applicable in the (more likely) case that the intervention was applied to a group of sectors.

The chapter starts by providing some background on the use of targets for industrial energy efficiency, and describes one particular case in the United Kingdom where the implementation mechanism includes the possibility of affected firms participating in a trading scheme. It also provides a brief overview of the UK chemicals sector, on which the model that is used in the analysis is (loosely) based. This model is described in detail in the second section of the chapter. Following on from the earlier comment, it should be noted that while it captures many of the salient features of the UK chemicals sector, the model is not intended to provide a detailed and faithful representation of the sector.

The analysis presented in the third section is divided into three parts. The first part, which considers the regulated aggregate cost minimum, assesses the impact of the regulatory intervention on the "effective" prices of inputs and outputs, and on the relative mix of inputs used by chemical producers. It also provides an alternative interpretation for the shadow price of the aggregate performance rule. The other two parts relate to the trading mechanism. The first of these assesses the impact of an increase in the market price of performance credits on the behaviour of an individual producer – in particular the number of credits that it chooses to transfer, the quantity of output that its produces, and the amount of energy that it uses. The second considers the market equilibrium for performance credits, and assesses the impact of the introduction of the trading mechanism on aggregate levels of output and energy consumption, compared with a system of common individual performance standards.

The impact of the introduction of the trading scheme is explored further in the fourth section, using a simple numerical simulation. In particular, the market equilibrium outcome is compared with that under a system of common, individual performance standards, both in terms of the aggregate cost of the intervention and the distribution of the cost between different types of chemical producer. The fifth section of the chapter considers how performance adjustment factors can be used to alter the distributional impact of the regulatory intervention. In particular, it considers one approach that might be adopted that has the affect of translating the aggregate target rate into a series of differentiated targets for constituent sub-sectors, where these represent equal percentage improvements over their respective pre-regulation rates.

5.1 Background

5.1.1 Industrial energy efficiency targets

Industrial energy efficiency has been a significant policy issue since the late 1970's. Originally this was for reasons of industrial competitiveness. While this remains an important motivation, over the last decade it has been reinforced by concerns about the environmental impacts arising from the production and use of energy products. This has led (*inter alia*) to the introduction of a range of regulatory interventions to improve the energy efficiency of industry. These have taken a variety of different forms, including reporting requirements; best practice dissemination programmes; support for particular technologies such as combined heat and power (CHP); and a range of other subsides and tax incentives.

While the underlying objective of these interventions has been to improve the energy efficiency of industry, there have been few instances of quantified targets for specific energy consumption. That is, targets for the quantity of energy used per unit of output produced (typically measured in tonnes). Where they have arisen, targets have largely been set in the context of non-binding voluntary agreements between governments and industry. For example:

- In Germany, eighteen industry sectors / associations – accounting for around two thirds of total industrial energy consumption – set quantified targets for reducing specific energy consumption (or specific CO_2 emissions) under the 1996

Declaration of German Industry on Global Warming Prevention. The overall target is for a 20% reduction between 1990 and 2005, with individual sector targets varying around this figure to reflect their respective energy saving potentials.

- In Japan, eighteen industries set targets for specific energy consumption (or specific CO_2 emissions) under the *Voluntary Action Plan on the Environment* that was set up by the Japan Federation of Economic Organisations. For example, the Japan Chemicals Association has set a target of a 10% reduction in specific energy consumption by 2010, compared to a 1990 baseline.

In both of the above examples, the targets are collective (i.e. they are set for a sector as a whole, rather than for individual firms), and are not legally binding. However, there are two countries – the Netherlands and the United Kingdom – that have introduced legally binding energy efficiency targets for individual firms.

The Netherlands was the first country to introduce quantified energy efficiency targets for industry, when the 1990 *Energy Efficiency Policy Document* set a target of a 20% reduction in specific energy consumption by the year 2000 compared with 1989. The policy was implemented through a series of legally binding *Long Term Agreements* (LTAs) with forty one sectors (or branches), which together accounted for 90% of total industrial energy consumption.[1] These agreements covered 1200 firms, each of which was required to produce an *Energy Efficiency Plan* which included (*inter alia*) a quantified improvement target. Individual targets were allowed to differ (to reflect different starting points, etc.), provided that in aggregate they were consistent with the sectoral target.

[1] A small number of these agreements were with individual companies rather than sector associations. For example, agreements were concluded with Phillips, KLM and Schiphol Airport.

5.1.2 UK Climate Change Agreements

In the United Kingdom, forty-four energy intensive sectors have been granted an 80% exemption from the *Climate Change Levy* (and industrial energy tax) since its introduction in April 2001, in return for entering into legally binding negotiated agreements with the government. In recognition of the variation in circumstances between different sectors, the government allowed sectors to choose between absolute and relative targets, for either energy consumption or carbon emissions. However, all but four of the sectors elected for relative energy targets – with thirty-seven agreeing explicit targets for specific energy consumption (SEC), and three agreeing targets that are expressed in the form of an energy efficiency ratio (EER). This ratio – which is a refinement of the concept developed by NOVEM in the Netherlands – allows for variation in the volume of throughput and changes in the product mix that would distort trends in aggregate specific energy consumption.

The *Climate Change Agreements* have a two-tier structure. At the top level there is an "umbrella" agreement with the relevant sector association, which sets an aggregate target for the sector. Underneath this, there are a series of "underlying agreements" between the government and the constituent target units, which stipulate individual performance targets. In all, more around 5,700 underlying agreements were signed by the time that the Levy came into force in April 2001, covering more than 12,000 individual facilities (or sites).[2] The number of underlying agreements varies considerably between sectors – ranging from one in the case of environmental services, to almost one thousand in the case of the food and drink sector.

The agreements last until the year 2010, with interim targets defined at two-yearly intervals, starting in 2002. Performance is reviewed against each of the interim targets,

[2] A target unit is an industrial facility, or group of facilities, that share a single target. A facility is an IPPC eligible activity, plus any permitted associated activities. For six of the sectors, the underlying agreements are between the target units and the respective sector association.

and if the sector target is met, then all constituent target units are automatically "re-certified" for the next two years. If it is not, then the discount is only extended for those units that have met their individual targets. At the end of the first target period, twenty-four of the sectors had met their respective aggregate targets, and 88% of the target units (75% of facilities) were re-certified. Around 200 target units failed to achieve their individual targets and were not re-certified for 2002-2004, although they may regain the exemption in future periods. Finally, around 500 target units either left the agreements, or failed to submit any data, and have therefore forfeited their exemption for the entire period of the agreements.

The agreements include a number of mechanisms that allow firms flexibility in meeting their individual targets. In particular, they are eligible to participate in the UK Emissions Trading Scheme (ETS), which was launched in April 2002.[3] Unlike a traditional emission trading scheme, the ETS comprises two sectors – an *absolute sector*, in which firms with absolute targets participate on a "cap and trade" basis; and a *relative sector*, in which firms with relative targets participate on a "baseline and credit" basis. Agreement participants that choose to join the scheme are placed into the relative sector (apart from those in the two sectors that adopted absolute targets), and are free to trade allowances with any other firm in that sector – not just those in their own industry. They can also trade allowances with firms in the absolute sector. However, these transactions are subject to the restriction that they must not result in a net transfer of allowances from the relative to the absolute sector (the so-called "gateway" restriction).

If an agreement participant achieves a specific energy consumption rate that is lower than its target value, then it is allowed to convert this over-achievement into allowances, subject to verification of its performance by an accredited verifier. The maximum

[3] The target units are not automatically included in the trading scheme. If a unit wishes to participate, it must register with the Secretary of State for Environment, Food and Rural Affairs. A comprehensive overview of the scheme is provided in DEFRA (2001).

number of allowances that it can create is equal to the difference between its target and actual rates, multiplied by its output level for the year, and by a carbon-energy conversion factor to reflect the fact that allowances are denominated in tonnes of carbon dioxide equivalent (tCO_2e). These allowances can either be sold to other participants in the ETS, or banked for use in future target years. Conversely, if an agreement participant fails to meet its target, it must purchase sufficient allowances to make up the shortfall (using an equivalent calculation).

During the first target year, one-fifth of all agreement participants took advantage of the trading scheme; with 565,000 tCO_2e of allowances being purchased by 743 participants in 27 different sectors, for use towards meeting their targets; and over 1.3 million tCO_2e of allowances being allocated to 123 participants from 17 sectors.[4] Since the gateway remained open throughout the year, this implies that at least 800,000 tCO_2e of the allowances were banked. In addition, approximately 2.7 million tCO_2e of over-achievement was ring-fenced for future conversion into allowances (i.e. effectively banked).[5]

5.1.3 UK chemicals sector

As can be seen in Table 5.1, the chemicals sector is a significant part of the UK economy, contributing one tenth of total industry gross value added (GVA) at current basic prices in the year 2000. While it is less significant in terms of employment, it uses more energy than any other sector, accounting for more than 20% of total industrial energy consumption (measured in tonnes of oil equivalent).

---

[4] The 866 participants in the trading scheme represented 1208 target units, reflecting the fact that some firms have several target units. The various figures for the CCA participants are taken from a review of the first year of operation of the ETS undertaken for DEFRA by Future Energy Solutions (FES, 2003).

Table 5.1 Comparison of major industrial sectors - 2000

SIC (92) Classification		Share of total industry			Relative	
		GVA (1)	Energy (2)	Emp's (3)	GVA per TOE	GVA per FTE
15-16	Food, beverages and tobacco	13.3%	10.5%	12.5%	1.3	1.1
21-22	Pulp, paper, publishing and printing	13.5%	7.0%	11.7%	1.9	1.2
24	Chemicals	10.0%	21.2%	6.0%	0.5	1.7
27	Basic metals	2.6%	14.4%	3.0%	0.2	0.9
28-29	Machinery ,equipment and fabricated metal products	16.0%	5.3%	19.2%	3.0	0.8
30-33	Electrical and optical equipment	13.7%	2.8%	12.5%	5.0	1.1
34-35	Transport equipment	10.4%	4.7%	10.2%	2.2	1.0
	Other sectors	20.4%	34.3%	24.9%	0.6	0.8
	Total industry	**100.0%**	**100.0%**	**100.0%**	**1.0**	**1.0**

Sources:

(1) Gross value added at current basic price: *Table 16.4, Annual Abstract of Statistics*

(2) Total energy consumption in tonnes of oil equivalent: *Table 1.3, Digest of UK Energy Statistics*

(3) Number of employees in full time equivalents: *Table 7.5, Annual Abstract of Statistics*

The chemicals sector has a relatively high rate of labour productivity; with gross value added per full time employee being seventy percent higher than the average for industry as a whole. However, in terms of its energy inputs, the sector has a relatively low rate of productivity. Gross value added per tonne of oil equivalent is half the average for total industry; with only the basic metals sector being less efficient.

[5] In order to reduce the relative scale of the transaction cost involved in creating allowances (particularly for smaller participants), any over-achievement can also be "ring-fenced", to be combined with over-achievements in future years and verified as a larger block.

One of the striking features of the chemicals sector is its heterogeneity. It comprises a number of distinct sub-sectors, with products ranging from industrial gases and chemicals, to perfumes and pharmaceuticals. However, essentially chemicals producers can be divided into two broad groups – *basic chemicals* producers (corresponding to SIC(92) subclass DG / 24.1) and *speciality chemicals* producers (corresponding to subclasses DG / 24.2 -24.7).[6]

Table 5.2 Profile of the chemicals sector – 2000

	Share of total chemicals sector				Relative		
	Output (1)	GVA (2)	Tonnes (3)	Energy (4)	Price per tonne	GVA per TOE	SEC
Basic chemicals	36%	25%	85%	79%	0.45	0.32	0.93
Specialty chemicals	64%	75%	15%	21%	3.20	3.57	1.40
Total chemicals	100%	100%	100%	100%	1.00	1.00	1.00

Sources:

(1) Output value at current basic prices: *UK Supply-Use Tables*

(2) Gross value added at current basic prices: *UK Supply-Use Tables*

(3) Tonnage of output: *Chemical Industries Association estimate*

(4) Energy consumption in tonnes of oil equivalent: *Energy Consumption in the UK*

In financial terms, specialty chemicals is the more significant, accounting for almost two-thirds of the sales value for the sector, and three-quarters of the gross value added. However, in physical terms the situation is reversed, with basic chemicals accounting for approximately 85% of the tonnage produced by the sector, and almost 80% of the total energy consumption. This is reflected in the relative prices, and the relative energy efficiencies, of the two sub-sectors. It is interesting to compare the two measures of

[6] Basic chemicals comprise: industrial gases; dyes and pigments; inorganic base chemicals; organic base chemicals; fertilizers and nitrogen compounds; plastics in primary form. Speciality chemicals comprise:

energy efficiency that are provided in Table 5.2. While specialty chemicals production is eleven times more efficient than basic chemicals production in terms of gross value added per tonne of oil equivalent (i.e. a financial measure of efficiency), it actually has the higher rate of specific energy consumption (i.e. a physical measure of inefficiency)!

Table 5.3 Demand for chemicals products - 2000

	Basic chemicals	Specialty chemicals	Total chemicals
Basic chemicals	20.6%	0.0%	7.2%
Specialty chemicals	10.4%	6.7%	8.0%
Other intermediate	32.7%	30.3%	31.2%
Final	1.6%	30.6%	20.6%
Gross capital formation	0.1%	0.7%	0.5%
Export	34.5%	31.6%	32.6%
Total	**100.0%**	**100.0%**	**100.0%**

Source: *UK Supply-Use Tables*

The markets for chemicals products are highly international. This is reflected in the fact that exports account for approximately one-third of total demand for chemicals products, while imports account for over one-quarter of total supply (not shown in Figure 5.3). While, there is little difference between basic and specialty chemicals in terms of this element of demand, there are considerable differences in relation to some of the others. In particular, final demand by households accounts for almost one-third of the total for specialty chemicals, but is insignificant for basic chemicals. Chemicals producers themselves account for around 15% of total demand. Half of this represents "internal demand" (e.g. use of industrial gases in the production of industrial gases,

pesticides and agro-chemicals; paints, varnishes, & inks; pharmaceuticals; soaps, detergents, & perfumes; other chemical products; man-made fibres.

etc.). The majority of the other half comprises demand for industrial gases, and organic and inorganic chemicals by other chemical producers – principally producers of plastics, pesticides, paints and varnishes, and man-made fibres.

The chemicals sector was one of the earliest to address the issue of energy efficiency. The Chemical Industry Association (CIA) first introduced a voluntary target for improvements in the aggregate energy efficiency of its members in 1991/92, as part of its Responsible Care Programme. In 1997, the CIA entered into a formal (but non-binding) voluntary agreement with the UK government, in which it agreed to reduce aggregate specific energy consumption by 20% between 1990 and 2005.[7] This agreement ran until April 2001, when it was superseded by a Climate Change Agreement that set an improvement target for the sector's energy efficiency ratio (EER) of 18.3% between 1998 and 2010. This corresponds to an improvement of 34% between 1990 and 2010. The sector has been one of the more pro-active in taking up the flexibility provided by the trading mechanism; with the Chemical Industry Association Broking and Trading Agency (CIABATA) being set up in 2001, to manage and administer the sector's Agreement, and to support its members' trading activities.

5.2 Model

In this stylised model of the chemicals production system, there are four mutually exclusive production sectors, three mutually exclusive categories of endogenous commodities, and no non-market environmental inputs (i.e. $M = \varnothing$).[8] The sectors and categories are defined in Table 5.4.

[7]Salmons (2001) provides a detailed evaluation of the agreement.

[8]In reality the chemicals sector is a major source of emissions of a range of pollutants. While these may be one of the underlying motivations for the regulatory intervention, they are not included in the regulatory target. Consequently, non-market environmental inputs are omitted from the model for the sake of simplicity.

Table 5.4 Definition of sectors and categories

Production sectors		Commodity categories	
$I^1 \subset I$	Energy producers	$K^1 \subset K$	Energy products
$I^2 \subset I$	Feedstock producers	$K^2 \subset K$	Feedstock products
$I^3 \subset I$	Chemicals producers	$K^3 \subset K$	Chemicals products
$I^4 \subset I$	Chemicals users		

The following assumptions are made regarding the relationships between the various production sectors and commodity categories. That is, which sectors produce / use each category. These are reflected in Figure 5.1, which provides a schematic representation of the production system.

$I^{k+} \subset I^1$ for all $k \in K^1$ $\qquad$ $I^{k-} \subset I^1$ for all $k \in K^1 \cup K^2$

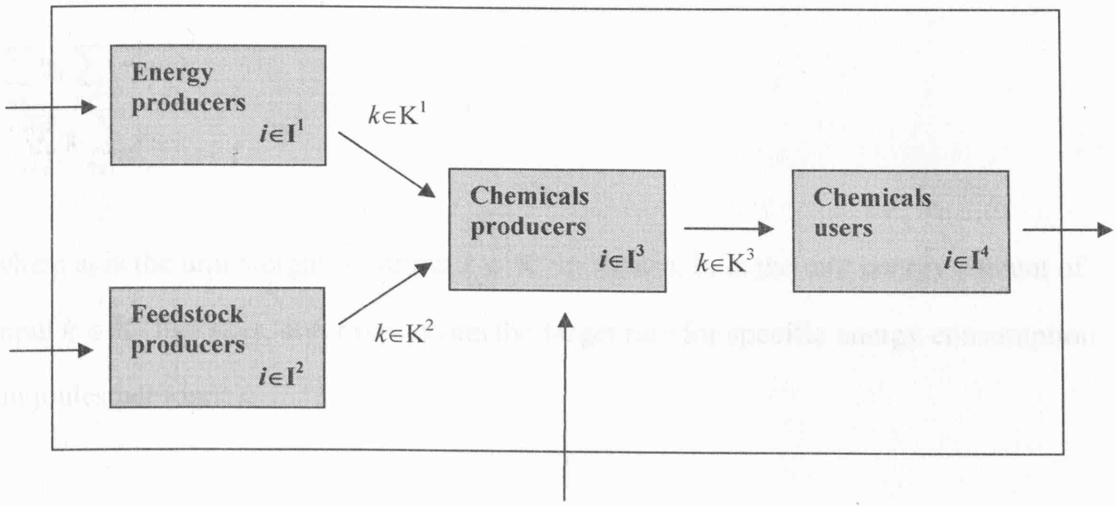
$I^{k+} \subset I^2$ for all $k \in K^2$ $\qquad$ $I^{k-} \subset I^4$ for all $k \in K^3$

$I^{k+} \subset I^3$ for all $k \in K^3$

Thus, firms in the chemicals sector use feedstock products, energy products (and a range of exogenous inputs such as plant and equipment, and labour) to produce a variety of chemicals products, which are then used as inputs by firms in other sectors of the economy.

Of course, this representation is a considerable simplification of what is, in reality a much more complex production system. In particular, it does not allow for energy to be produced as a by-product of chemicals production. Nor does it allow for chemicals products to be used as inputs by firms in the chemicals sector. Notwithstanding these simplifications however, the model is sufficiently rich to provide a number of interesting insights regarding the impacts of the trading scheme.

Figure 5.1 Production system structure



The production technologies of all firms in sectors I^1, I^2 and I^4 are identical (within sectors); all exhibiting constant returns to scale. Consequently, the inverse supply curves for energy products and for feedstock products are all flat; as are the inverse demand curves for chemicals products. This implies that the market prices of all endogenous commodities are constant. In order to clarify the notation, a distinction is made between the prices of chemicals sector's input commodities $k \in K^1 \cup K^2$ – which are denoted by c_k, and the prices of its output commodities $k \in K^3$ – which are denoted by p_k.

Since the prices of all endogenous commodities are constant, and feedstock products $k \in K^2$ are excluded from the aggregate performance rule (see below), the requirement that firms in sector I^3 use an exogenous input can be dropped.[9] The production

[9] The omission of the exogenous inputs is done only to simplify the notation. It is not intended to imply that the other inputs are unnecessary, or unimportant for production of chemicals products.

technologies of chemicals producers are heterogeneous, with $f^i(\boldsymbol{w}_i)$ homogeneous of degree $\psi_i < 1$ (and hence strictly concave).

The regulatory intervention takes the form of a relative standard for the maximum aggregate rate of specific energy consumption by the chemicals sector. That is:

$$\frac{\sum_{k \in K^1} b_k \sum_{i \in I^{k-}} (-w_{ki})}{\sum_{k \in K^3} a_k \sum_{i \in I^{k-}} y_{ki}} \quad \leq \quad r$$

where a_k is the unit weight of output $k \in K^3$ in tonnes, b_k is the unit energy content of input $k \in K^1$ in joules, and r represents the target rate for specific energy consumption (in joules per tonne).

Thus, the values of the parameters in the aggregate performance rule are:

α_k	$=$	0	β_k	$=$	b_k	for all $k \in K^1$
	$=$	0		$=$	0	for all $k \in K^2$
	$=$	$r\,a_k$		$=$	0	for all $k \in K^3$
δ	$=$	0				

The resultant rule is:

$$r \sum_{k \in K^3} a_k \sum_{i \in I^{k+}} y_{ki} \quad + \quad \sum_{k \in K^1} b_k \sum_{i \in I^{k-}} w_{ki} \quad \geq \quad 0 \qquad \ldots (5.1)$$

where the first term represents the total amount of energy consumption allowed under the performance rule, and the second term is the amount used.

It is assumed that $\theta_k = 0$ for all $k \in K$, and that $\varepsilon_i = 0$ for all $i \in I^1 \cup I^2 \cup I^4$. That is, all obligations and all initial property rights are assigned to the firms in the chemical sector, and that there is no reallocation of rights to firms in the other sectors. Consequently, the individual performance rules are:

$$r \sum_{k \in K^{3i}} a_k y_{ki} \quad + \quad \sum_{k \in K^{1i}} b_k w_{ki} \quad + \quad \varepsilon_i \quad \geq \quad v_i \qquad \text{for all } i \in I^3 \qquad \ldots (5.2.a)$$

$$0 \quad \geq \quad v_i \qquad \text{for all } i \in I^1 \cup I^2 \cup I^4 \qquad \ldots (5.2.b)$$

Thus, firms in the energy production, feedstock production, and chemicals user sectors are all inactive. It is also clear from (5.2.a) that in this application, the performance credits are denominated in units of energy (i.e. joules). Rearranging the individual performance rule for firm $i \in I^3$ gives:

$$\frac{\sum_{k \in K^{1i}} b_k (-w_{ki}) \; + \; v_i}{\sum_{k \in K^{3i}} a_k y_{ki}} \quad \leq \quad r \quad + \quad \frac{\varepsilon_i}{\sum_{k \in K^{3i}} a_k y_{ki}}$$

The left-hand side of the inequality is the firm's net specific energy consumption, after taking account of any transfers of performance credits, while the right-hand side represents its "effective" target rate. With the exception of section 5.5, it is assumed that the performance adjustment factors are set to zero for all chemical producers (i.e. $\varepsilon_i = 0$ for all $i \in I^3$), and hence that each producer must achieve the actual target rate (r). However, it is clear that by setting non-zero values for the adjustment factors, it is possible to vary the effective target rate between different firms (or groups of firms). For example, if $\varepsilon_i < 0$ for a particular producer then the second term is negative, and hence it faces a more stringent target than the sector average.

The potential impact of differentiating the effective target rate in this way is considered in section 5.5. An important point to note is that, while the adjustment factor (ε_i) is constant, the impact of the factor on the effective target rate is not. The inclusion of the firm's output in the denominator of this term means that the magnitude of the adjustment diminishes as its output increases.

Since the production technologies of firms in the energy production, feedstock production, and chemicals user sectors all exhibit constant returns to scale, their profits are all equal to zero.[10] Consequently, the aggregate profit for the production system as a whole is equal to the aggregate profit of the chemicals sector. That is:

$$\Pi(\tilde{w} \mid \tilde{y}) \quad = \quad \sum_{i \in I^3} \sum_{k \in K^{i+}} p_k y_{ki} \quad + \quad \sum_{i \in I^3} \sum_{k \in K^{i-}} c_k w_{ki} \qquad \ldots (5.3)$$

where the concatenated vector ($\tilde{w} \mid \tilde{y}$) denotes the production plan for the chemicals sector.

An interesting point to note is that the inputs and outputs of firms in sectors I^1, I^2 and I^4 do not appear in (5.1), (5.2) or (5.3). Thus the preceding assumptions have had the effect of defining the aggregate performance rule in terms of exogenous commodities for a "reduced" production system comprising only the chemicals producers.

5.3 Analysis

5.3.1 Regulated aggregate cost minimum

Under the assumptions set out in section 5.2, the regulated aggregate cost minimum problem (RC) can be decomposed into the following two-stage problem:[11,]

$$\Pi(r) \quad = \quad \text{Max} \quad \sum_{i \in I^3} \sum_{k \in K^{i+}} p_k y_{ki} \quad + \quad C(\tilde{y}; r)$$

[10] It also means that the values of inputs and outputs for individual firms in sectors I^1, I^2 and I^4 are indeterminate. Only the aggregate values can be determined.

$$C(\tilde{y}; r) \quad = \quad \text{Max} \quad \sum_{i \in I^3} \sum_{k \in K^{i-}} c_k w_{ki}$$

$$\text{s.t.} \quad y_{ki} - f^i(w_i) \leq 0 \qquad \text{for all } i \in I^3, k \in K^{i+}$$

$$r \sum_{k \in K^3} a_k \sum_{i \in I^{k+}} y_{ki} + \sum_{k \in K^1} b_k \sum_{i \in I^{k-}} w_{ki} \geq 0$$

where $C(\tilde{y}; r)$ is the aggregate cost for the chemicals sector of producing the output plan $\tilde{y}$, conditional on achieving the target rate for aggregate specific energy consumption.

The necessary first order conditions for the two stages are respectively:

$$p_k = -C_{ki}(\tilde{y}^*; r) \qquad \text{for all } i \in I^3, k \in K^{i+}$$

$$c_{k'} + \lambda^* b_{k'} = -\mu_{ki}^* f^i_{k'}(w_i^*) \qquad \text{for all } I \in I^3, k' \in K^{i-}, k \in K^{i+}$$

where μ_{ki}^* is the shadow value of the production constraint, λ^* is the shadow value of the aggregate performance rule. Noting that $C_{ki} = \lambda^* r\, a_k - \mu_{ki}^*$ by the envelope theorem, these conditions can be combined to give:

$$-(p_k + \lambda^* r\, a_k) f^i_{k'}(w_i^*) = c_{k'} + \lambda^* b_{k'} \qquad \text{for all } i \in I^3, k \in K^{i+}, k' \in K^{1i-}$$

$$= c_{k'} \qquad \text{for all } i \in I^3, k \in K^{i+}, k' \in K^{2i-}$$

Thus, cost minimization requires both that the "effective price" of each energy input be raised above the pre-regulation market price – by an amount proportional to its energy content, and that the "effective price" of each output be raised above the pre-regulation market price – by an amount proportional to its unit weight.[12] For all feedstock inputs it

---

[11] The decomposition of the problem into two stages is not necessary. However, it facilitates the interpretation of the shadow value of the performance rule constraint (λ).

[12] Note that because the pre-regulation prices of all commodities are now exogenous (unlike the general case considered in Chapter 3), it is possible to make definitive statements about the impact of the regulatory intervention on prices.

is clear that this implies a decline in their respective marginal products. For energy inputs, the impact on the marginal product depends on the values of the performance rule parameters relative to the exogenous market prices, with:

$$-f^i_{k'}(\mathbf{w}^0_i) \;=\; \frac{c_{k'}}{p_k} \;\geq\; \frac{c_{k'} \;+\; \lambda^* b_{k'}}{p_k \;+\; \lambda^* r a_k} \;=\; -f^i_{k'}(\mathbf{w}^*_i) \qquad \text{iff} \qquad \frac{c_{k'}}{p_k} \;\geq\; \frac{b_{k'}}{r a_k}$$

However, this cannot be the case for all $k' \in K^{1i-}$, or else it would imply that the performance rule is satisfied at the pre-regulation solution. Consequently, the marginal product must increase for some of the energy inputs.

Without placing further restrictions on the properties of the production functions, it is not possible to draw any general conclusions regarding the impact on the absolute values of each firm's inputs and outputs. However, it is possible to say something about the impact on the relative mix of inputs. Taking the ratio of the various marginal product conditions yields the following sets of conditions for the marginal rates of technical substitution (MRTS) between the firm's inputs:

$$\frac{f^i_k(\mathbf{w}^*_i)}{f^i_{k'}(\mathbf{w}^*_i)} \;=\; \frac{c_k \;+\; \lambda^* b_k}{c_{k'} \;+\; \lambda^* b_{k'}} \qquad \text{for all } k, k' \in K^1$$

$$\frac{f^i_k(\mathbf{w}^*_i)}{f^i_{k'}(\mathbf{w}^*_i)} \;=\; \frac{c_k \;+\; \lambda^* b_k}{c_{k'}} \qquad \text{for all } k \in K^1, k' \in K^2$$

$$\frac{f^i_k(\mathbf{w}^*_i)}{f^i_{k'}(\mathbf{w}^*_i)} \;=\; \frac{c_k}{c_{k'}} \qquad \text{for all } k, k' \in K^2$$

Thus, for any two inputs the change in the marginal rate of technical substitution will depend on the relative values of the energy content of each input. If $b_k > b_{k'}$ (i.e. the numerator input has the higher energy content), then the MRTS increases, and hence there is a reduction in the relative use of that input compared to the denominator input –

assuming that all other inputs are kept fixed.[13] In particular, there is a reduction the relative use of any energy input $k \in K^1$ compared to any feedstock input $k' \in K^2$.

The decomposition of the aggregate cost minimization problem into two stages allows an alternative interpretation to be derived for λ^*, the shadow price of the aggregate performance rule constraint.[14] Differentiating the aggregate profit function $\Pi(r)$ at the solution $\tilde{y}^* = \tilde{y}(r)$, and applying the first order and envelope conditions from the stage 2 problem yields:

$$\lambda^* \quad = \quad \frac{\Pi'(r)}{Y^*} \quad = \quad \frac{C_r(\tilde{y}^*;r)}{Y^*} \qquad \text{where} \quad Y^* \quad = \quad \sum_{j \in J} a_j \sum_{i \in I^j} y^*_{ji}$$

Thus λ^*, and hence the price of performance credits in the market equilibrium, can be interpreted as the marginal cost of energy efficiency per unit output (evaluated at the optimal value of the sector output plan $\tilde{y}^*$), or the *specific marginal cost of energy efficiency*. It is important to note that this is not the same as the change in the value of average profit per tonne. However, since Y^* and $\tilde{y}$ are both constant in the conditional cost minimization problem, it follows that:

$$\lambda^* \quad = \quad \frac{\partial}{\partial r}\left(\frac{C(\tilde{y}^*;r)}{Y^*}\right)$$

Consequently, λ^* does represent the <u>short run</u> change in the average aggregate cost of production per tonne (i.e. holding the outputs of individual agents constant). This contrasts with the case of an absolute limit for energy consumption, where the shadow

[13] For any concave production function, the elasticity of substitution between inputs is non-negative. Hence, an increase in the MRTS between two inputs leads to a reduction in the relative quantities used. Furthermore, for any CES production function (including Cobb-Douglas functions), the caveat that all other inputs be kept fixed can be dropped, as for these functions the MRTS for any two inputs is independent of all other inputs.

[14] In Chapter 3, an interpretation was derived which related λ^* to the marginal cost of abatement

value of the constraint is equal to the change in the total aggregate cost of production resulting from a marginal increase in the limit (i.e. the marginal cost of abatement).

5.3.2 Impact on individual producer

The analysis of the impact of a rise in the price of performance credits on the behaviour of an individual chemical producer $i \in I^3$ is facilitated by the decomposition of its regulated profit maximization problem (RB_i) into the following three stages:[15]

$$\pi(r, \mathbf{c}, p, q) \quad = \quad \underset{v}{\text{Max}} \quad \pi(v; r) \;+\; q\, v \qquad (\textit{net profit maximization})$$

$$\pi(v; \mathbf{c}, r, p) \quad = \quad \underset{y}{\text{Max}} \quad p\, y \;+\; c(y, v; r) \qquad (\textit{operating profit maximization})$$

$$c(y, v; \mathbf{c}, r) \quad = \quad \underset{\boldsymbol{w}}{\text{Max}} \quad \mathbf{c}'\boldsymbol{w} \qquad (\textit{production cost minimization})$$

$$\text{s.t.} \quad y - f(\boldsymbol{w}) \leq 0$$

$$v - r\, y - \mathbf{b}'\boldsymbol{w} \leq 0$$

In the first stage, the firm chooses the quantities of its feedstock and energy inputs in order to minimize the cost of producing a given level of output while satisfying its individual performance rule, conditional on the number of credits transferred.[16] In the second stage, it chooses the output level that maximizes its operating profit, again conditional on the number of performance credits that it has transferred. Finally, in the third stage, it chooses the quantity of performance credits to transfer.

[15] In order to simplify the notation, the firm and commodity indexes (i and k) have been omitted, as have the # superscripts for the solution values of the variables and the Lagrange multipliers. It is also assumed that the output is measured in tonnes (i.e. a = 1), and that the number of inputs used by the firm is equal to K (i.e. $|K^{i-}| = K$).

a) Production cost minimization

The decomposition of the problem into the three stages has some important consequences for the production cost minimization problem. First, for any positive value of v, there is a threshold value $y^{min}(v) = v / r$, such that the constraint set is empty for all values of $y < y^{min}(v)$. Second, there is a threshold value $v^{min} < 0$, such that the performance rule constraint is just binding under the firm's unregulated production plan $(y^u, \mathbf{w}^u)$, i.e. $v^{min} = r\, y^u + \mathbf{bw}^u$. Third, depending on the value of v, there may be a range of output values – denoted by the closed interval $[y^-(v), y^+(v)]$ – within which the cost-minimizing input vector is feasible (i.e. the performance rule constraint is non-binding). It follows directly that $y^+(v^{min}) = y^u$, and that $y^-(v) = 0$ for all $v < 0$.

Taking note of these points, the Lagrangian for the production cost minimization problem is:

$$\mathcal{L} \quad = \quad \mathbf{c'}w \; + \; \mu\,[\, f(w) \; - \; y\,] \; + \; \lambda\,[\, r\,y \; + \; \mathbf{b'}w \; - \; v \;]$$

By the envelope theorem:

$$c_y(y, v) \quad = \quad \mathcal{L}_y \quad = \quad - \;\; \mu(y, v) \; + \; r\,\lambda(y, v) \qquad \ldots (5.4.a)$$

$$c_v(y, v) \quad = \quad \mathcal{L}_v \quad = \quad - \;\; \lambda(y, v) \qquad \ldots (5.4.b)$$

and hence:

$$c_{yy} \quad = \quad - \;\; \mu_y \; + \; r\,\lambda_y \qquad \ldots (5.4.c)$$

$$c_{yv} \quad = \quad c_{vy} \quad = \quad - \;\; \mu_v \; + \; r\,\lambda_v \qquad \ldots (5.5.a)$$

---

[16] The decomposition is for expositional purposes only. It is not intended to imply that firms actually take decisions in this sequence.

$$c_{vv} \quad = \quad -\lambda_v \qquad \ldots (5.5.b)$$

The second partial derivatives of the Lagrangian are:

$$\mathcal{L}_{kk'} = \mu f_{kk'} \qquad \mathcal{L}_{k\mu} = f_k \qquad \mathcal{L}_{k\lambda} = b_k$$

$$\mathcal{L}_{kv} = 0 \qquad \mathcal{L}_{\mu v} = 0 \qquad \mathcal{L}_{\lambda v} = -1$$

$$\mathcal{L}_{ky} = 0 \qquad \mathcal{L}_{\mu y} = -1 \qquad \mathcal{L}_{\lambda y} = r$$

from which it follows that:

$$\lambda_v = \frac{G_{11}}{G} \qquad \mu_v = \frac{-G_{12}}{G}$$

$$\lambda_y = \frac{-r\,G_{11} - G_{21}}{G} \qquad \mu_y = \frac{r\,G_{12} + G_{22}}{G}$$

$$\text{where} \quad G = |\mathbf{G}| = \left|\begin{array}{cc|ccc} 0 & 0 & b_1 & \cdots & b_K \\ 0 & 0 & f_1 & \cdots & f_K \\ \hline b_1 & f_1 & \mu f_{11} & & \mu f_{1K} \\ \vdots & \vdots & & \ddots & \\ b_K & f_K & \mu f_{K1} & & \mu f_{KK} \end{array}\right|$$

and G_{11}, G_{12}, G_{21} and G_{22} are the (unsigned) cofactors of the elements in the top left-hand corner of **G**. Using the standard results for partitioned matrices, and the properties of the Hessian (**H**) and the bordered Hessian (**B**) for a homogeneous function that were derived in Chapter 4, the following expressions can be derived for the determinant and the cofactors.

$$G = -\mu^{K-2} H \, (\nabla f + \boldsymbol{\beta})' \, \mathbf{H}^{-1} (\nabla f + \boldsymbol{\beta})$$

$$G_{11} = -\mu^{K-1} H \, \nabla f' \, \mathbf{H}^{-1} \nabla f = \mu^{K-1} \frac{H}{1-\psi} \nabla f' \boldsymbol{w}$$

$$G_{12} \quad = \quad G_{21} \quad = \quad -\mu^{K-1}\,H\ \nabla f'\,\mathbf{H}^{-1}\,\mathbf{b} \quad = \quad \mu^{K-1}\,\frac{H}{1-\psi}\,\mathbf{b}'\boldsymbol{w}$$

$$G_{22} \quad = \quad -\mu^{K-1}\,H\ \mathbf{b}'\,\mathbf{H}^{-1}\,\mathbf{b}$$

Therefore, it follows that:

$$c_{yy} \quad = \quad -\left(\frac{r^2 G_{11} + r(G_{12}+G_{21}) + G_{22}}{G}\right)$$

$$= \quad \mu^{K-1}\left(\frac{H}{G}\right)(r\nabla f + \mathbf{b})'\,\mathbf{H}^{-1}\,(r\nabla f + \mathbf{b}) \qquad \ldots (5.6.a)$$

$$c_{vv} \quad = \quad -\left(\frac{G_{11}}{G}\right)$$

$$= \quad \mu^{K-1}\left(\frac{H}{G}\right)\nabla f'\,\mathbf{H}^{-1}\,\nabla f \qquad \ldots (5.6.b)$$

$$c_{yv} \quad = \quad c_{vy} \quad = \quad \frac{r\,G_{11} + G_{12}}{G}$$

$$= \quad \mu^{K-1}\left(\frac{H}{G}\right)\frac{(r\nabla f + \mathbf{b})'\boldsymbol{w}}{1-\psi} \qquad \ldots (5.6.c)$$

Since f($\boldsymbol{w}$) is strictly concave, **H** is negative semi-definite, and so too is its inverse $\mathbf{H}^{-1}$. Hence the quadratic forms in equations (5.6.a) and (5.6.b) are both negative. Noting that G and H always have the same sign, it follows that $c_{yy} < 0$ and $c_{vv} < 0$. That is, the marginal cost of production with respect to output increases in magnitude as output increases, and the marginal cost of production with respect to sales of performance credits increases in magnitude as the sales of credits increases.[17]

[17] Note that $c(y, v) \leq 0$. Hence a decrease in the value of $c(y, v)$ represents an increase in the magnitude of the cost.

Unfortunately, it is not possible to make such a definitive statement about the sign of c_{yv}, as this depends on the sign of the expression $(r\nabla f + \mathbf{b})'w$, which may be negative or positive, depending on the values of y and v. If $v \leq 0$, then the expression is unambiguously negative since $(r\nabla f + \mathbf{b})'\mathbf{w} < r\, f(\mathbf{w}) + \mathbf{b}'\mathbf{w} = v$. By continuity, this will also be the case for small positive values of v. However, for larger values of v it is possible for the expression to be positive. For example, if $f(w) = (-w_1)^{0.25}(-w_2)^{0.25}$, $r = 1$, $\mathbf{b} = (1,0)$ and $y = 1$, then it is straightforward to show that $(r\nabla f + \mathbf{b})' w < 0$ for all $v < 0.5$, and that $(r\nabla f + \mathbf{b})' w > 0$ for all $v > 0.5$. Thus, if the firm is a net buyer of credits, then a small increase in the quantity of credits transferred (i.e. a reduction in the number purchased) will lead to an increase in the magnitude of its marginal cost of production with respect to output. However, if the firm is a net seller of credits, then the impact may be in either direction.

b) Operating profit maximization

Turning to the operating profit maximization problem, the necessary first order condition is that:

$$p + c_y(y, v) = 0 \qquad \ldots (5.7)$$

This of course is the usual condition that price equals the marginal cost of production. The only difference being that the marginal cost is conditional on the value of v.

Substituting the solution $y(v)$ into the first order condition (5.7) and differentiating the resultant identity yields the following relationship between the change in the firm's output level, and the change in its transfers of performance credits:

$$\frac{dy}{dv} = -\frac{c_{yv}}{c_{yy}}$$

Consequently, the impact of increased transfers of performance credits depends on the sign of c_{yv}. If the firm is a net buyer of performance credits, then an increase in the value of v (i.e. a reduction in the quantity purchased) will unambiguously lead to a decrease in its output. However, if the firm is a net seller of credits then the impact on output depends on the values of y and v. In the preceding analysis these values were completely arbitrary and independent. Of course, this is no longer the case, and it may be that for any given value of v, the firm's optimal output level $y(v)$ is always such that the sign of c_{yv} is clearly positive, or clearly negative? Investigation of this possibility is facilitated by combining the firm's production cost and operating profit problems, noting that the production constraint is always binding. That is:

$$\pi(v; r) \quad = \quad \underset{w}{\text{Max}} \quad p\, f(\boldsymbol{w}) + \mathbf{c}'\boldsymbol{w} \qquad \text{s.t.} \qquad v - r\, f(\boldsymbol{w}) - \mathbf{b}'\boldsymbol{w} \leq 0$$

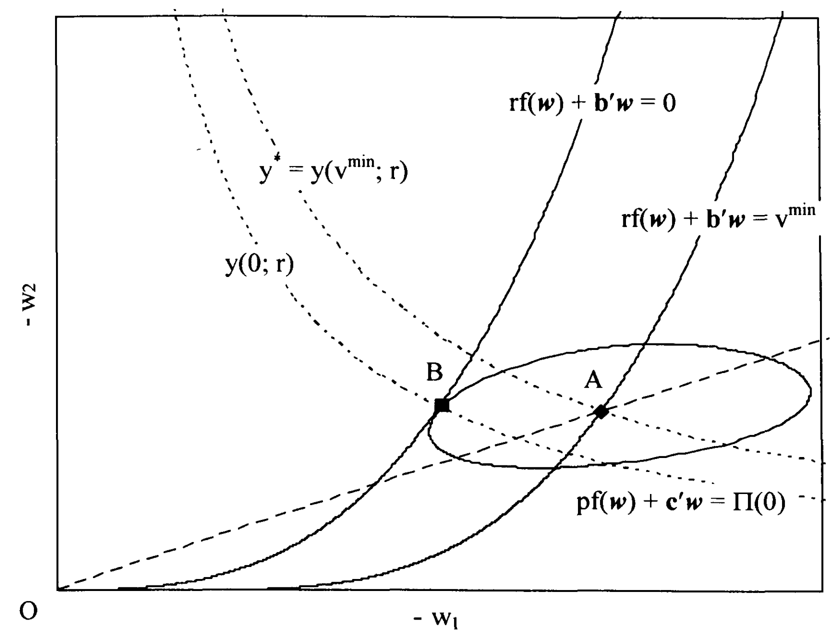
The necessary and sufficient first order conditions for this combined problem are:

$$(\, p\, f_k(\mathbf{w}) + c_k\,) + \lambda\,(\, r\, f_k(\mathbf{w}) + b_k\,) = 0 \qquad \text{for all } k \in \mathrm{K} \qquad \ldots (5.8)$$

Figure 5.2 provides a graphical illustration of the problem for the case of a firm with two inputs – one energy product ($k = 1$), and one feedstock product ($k = 2$), and a Cobb-Douglas production function of the form $f(w_1, w_2) = (-w_1)^{1/4} (-w_2)^{1/4}$. The point A denotes the firm's input choice in the absence of any regulatory intervention, and the extended dashed line OA represents the cost-minimizing combinations of inputs for different output levels.

The individual performance rule constraint is shown for two values of performance credit transfers. By definition, when $v = v^{\min}$ (< 0), the solution coincides with unregulated solution, at point A. When $v = 0$, the solution is at point B, where the constraint is tangential to the iso-profit contour (the ellipse centred on A). The two dotted lines are the iso-output contours that pass through the respective solutions. As one would expect (since $v < 0$ for all intervening values of transfers), the firm's output declines continually as the solution moves from A to B. But what happens when $v > 0$, and the solution trajectory moves on beyond B?

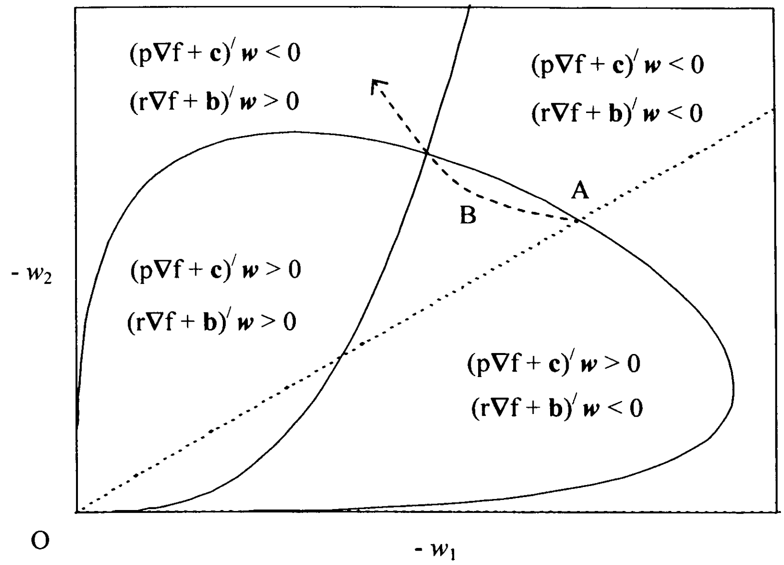
Figure 5.2 Combined conditional profit maximization problem with two inputs



The only general conclusion that can be drawn from the first order condition (5.8) is that in any solution, the expressions $(r\nabla f + \mathbf{b})'\mathbf{w}$ and $(p\nabla f + \mathbf{c})'\mathbf{w}$ must have opposite signs. However, as can be seen in Figure 5.3, the input combinations that satisfy this condition include combinations for which $(r\nabla f + \mathbf{b})'\boldsymbol{w} > 0$.

It is straightforward to show that with the assumed Cobb-Douglas production function, the minimum value of $(-w_2)$ for which the constraint is satisfied is directly proportional to the value of v. Since there is a maximum value of $(-w_2)$ for which $(p\nabla f + \mathbf{c})'\boldsymbol{w} > 0$, it follows that there is a threshold value of $v = v^{+}$, above which the solution must lie in the top left region – i.e. with $(r\nabla f + \mathbf{b})'\boldsymbol{w} > 0$. It has already been established that for all values of $v \leq 0$, the solution lies in the bottom right region – i.e. with $(r\nabla f + \mathbf{b})'\boldsymbol{w} < 0$. Consequently, as the value of v increases from v^{min}, the solution follows the trajectory shown in Figure 5.3, where A and B relate to the corresponding points in Figure 5.2. Thus, one can conclude that if the firm's prior sales of performance credits are relatively low (i.e. $v < v^{+}$), then any additional sales of credits will lead to a decrease in its output. However, if it has already sold a large number of performance credits, then additional sales will lead to an increase in output.

Figure 5.3 Solution trajectory for combined conditional profit maximization problem



For a given quantity of performance credits transferred, the firm's optimal levels of output and energy consumption must satisfy $\mathbf{b'w}(v) \equiv v - \mathrm{r}y(v)$. Since this is an identity, it can be differentiated to yield:

$$\frac{\mathrm{d}}{\mathrm{d}v}\mathbf{b'w}(v) = 1 + \mathrm{r}\left(-\frac{\mathrm{d}y}{\mathrm{d}v}\right)$$

Thus the impact of a change in the number of performance credits transferred on the firm's energy use is made up of two parts: a *substitution effect*, and an *output effect*. While the substitution effect is always positive (i.e. an increase in sales of credits results in a one-for-one reduction in the amount of energy used), the sign of the output effect is ambiguous. If $(r\nabla f + \mathbf{b})'\mathbf{w} < 0$, and hence $c_{yv} < 0$, then the output effect is also positive. In this case the firm will reduce the amount of energy that it uses by more than the quantity of performance credits that it sells. Conversely, it will increase its energy consumption by more than the quantity of credits that it purchases. On the other hand, if $(r\nabla f + \mathbf{b})'\mathbf{w} > 0$, then the magnitude of the change in energy use will be less than the magnitude of the change in performance credits transferred.

Of course, at this stage the value of v is completely arbitrary, and it may be that the number of credits that the firm chooses to transfer in the market equilibrium will always be sufficiently small to ensure that $(r\nabla f + \mathbf{b})'\mathbf{w} < 0$. In order to see whether this is the case, it is necessary to consider the firm's unconditional profit maximization problem.

c) Net profit maximization

The necessary first order condition for the firm's net profit maximization problem is:

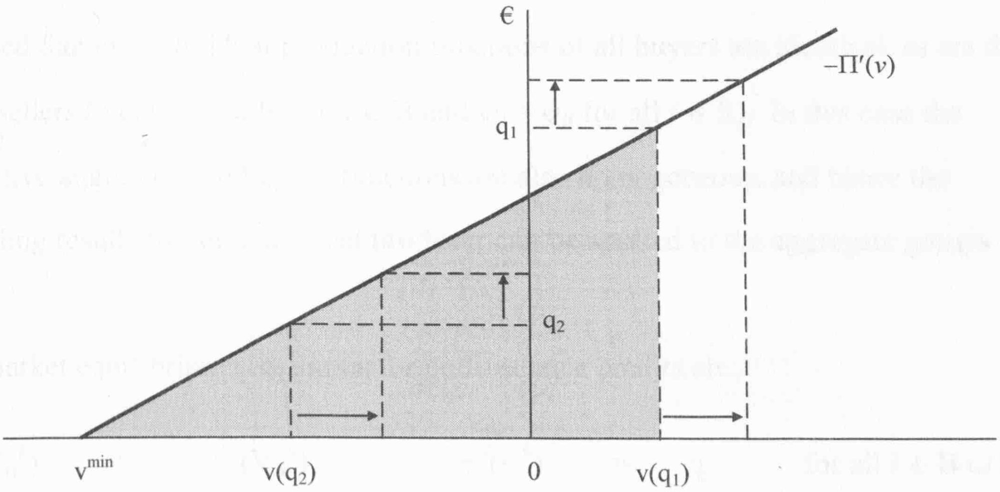
$$q = -\pi'(v) \qquad \dots (5.9)$$

where $-\pi'(v)$ represents the firm's marginal willingness to pay (WTP) for additional purchases of performance credits, or it marginal willingness to accept (WTA) for additional sales. Noting that $-\pi'(v) \equiv \lambda(v)$ and that $\lambda'(v) > 0$ with $\lambda(v^{min}) = 0$, it follows that a solution v(q) will always exist.[18] Figure 5.4 provides a graphical illustration of the solution; from which it is clear that the sign of v(q) depends on the relative values of q and $-\pi'(0)$, i.e. firm's WTP / WTA prior to engaging in any transfers. If $q > -\pi'(0)$, then the firm is a seller of credits. On the other hand, if $q < -\pi'(0)$, then it is a buyer. However, in either case, an increase in the price of credits leads to an increase in the

[18] The Lagrange multiplier $\lambda(v)$ relates to the combined conditional profit maximization problem. The derivation of the sign for $\lambda'(v)$ follows the same argument as that used in section 4.4 of Chapter 4, noting that an increase in v is equivalent to a decrease in δ.

quantity transferred. Of course, for a buyer this corresponds to a decrease the number of credits purchased.

Figure 5.4 Cost of individual performance rule



For a given quantity of credits transferred, the reduction in the firm's operating profit (i.e. excluding the financial value of the transfers) versus the pre-regulation level is:

$$\pi(v^{min}) \quad - \quad \pi(v) \quad = \quad \int_{v}^{v^{min}} \pi'(\varsigma)\, d\varsigma \quad = \quad \int_{v^{min}}^{v} -\pi'(\varsigma)\, d\varsigma$$

Thus, for a given market price of performance credits, the cost of the regulatory intervention to the firm is equal to the area under the graph of its WTA/WTP function between v^{min} and its optimal transfer quantity v(q) – shown as the grey shaded area in Figure 5.4.

5.3.3 Market equilibrium for performance credits

For the analysis of the market equilibrium for performance credits it is easier to work with the aggregate (inverse) demand and supply functions. To this end, the chemical producers are divided into two mutually exclusive groups – those that are net sellers of performance credits in the market equilibrium, and those that are net buyers (denoted respectively by the subsets $S \subset I^3$ and $B \subset I^3$). The aggregate production functions and the aggregate conditional profit functions for the two groups are denoted by $F^S(\boldsymbol{W}_S)$, $F^B(\boldsymbol{W}_B)$, $\Pi^S(V_S)$ and $\Pi^B(V_B)$, where $\boldsymbol{W}_S$, $\boldsymbol{W}_B$, V_S and V_B are the respective aggregate input vectors and aggregate transfers of performance credits. For simplicity it is assumed that the individual production functions of all buyers are identical, as are those of all sellers (i.e. $\psi_i = \psi_B$ for all $i \in B$ and $\psi_i = \psi_S$ for all $i \in S$). In this case the respective aggregate production functions are also homogeneous, and hence the preceding results for an individual producer can be applied to the aggregate groups.

The market equilibrium conditions for performance credits are:

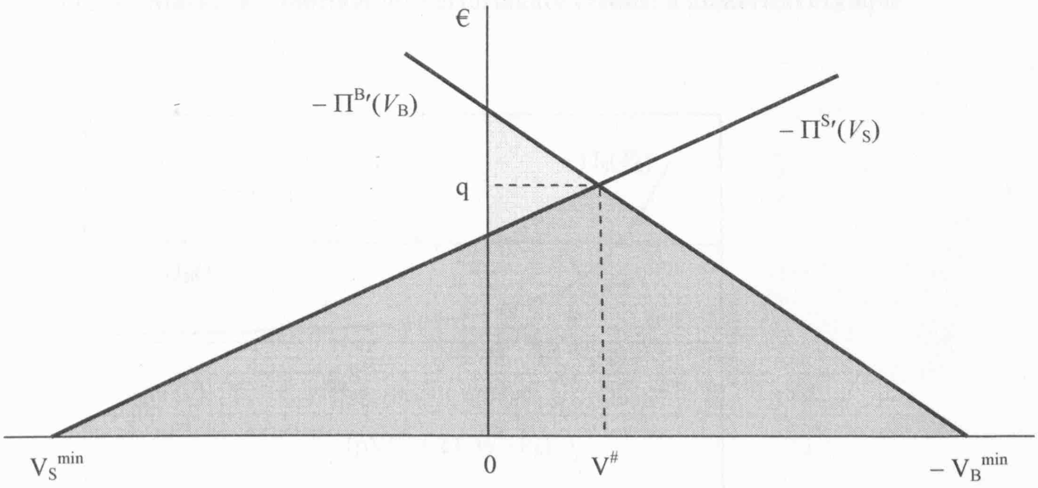
$$\Pi^{B\prime}(V_B^{\#}) \quad = \quad \Pi^{S\prime}(V_S^{\#}) \quad = \quad \pi^{i\prime}(\mathbf{v}_i^{\#}) \quad = \quad q \qquad \text{for all } i \in B \cup S$$

$$V_S^{\#} \quad = \quad V^{\#} \quad = \quad -V_B^{\#}$$

The market equilibrium is shown graphically in Figure 5.5, where the light grey shaded area represents the gains from trading. Since $v_i = 0$ for all $i \in I^3$ if $V_S = V_B = 0$, this is equal to the reduction in the aggregate cost of the regulatory intervention compared with an implementation mechanism that imposes the target as a common, individual performance standard.

The equilibrium market price lies in the range $-\Pi^{S\prime}(0) < q < -\Pi^{B\prime}(0)$. The exact price depends on the relative slopes of the WTA and WTP functions, as does the division of the total gain between the buyers and sellers of the credits. The dark grey shaded area represents the aggregate cost of the regulatory intervention when it is implemented efficiently. Since the aggregate WTA and WTP functions are both convex, it follows that the magnitude of the aggregate cost is less than or equal to $\frac{1}{2} \times | V_S^{min} + V_B^{min} | \times q$. Thus, an upper bound for the aggregate cost can be obtained by determining the number of credits that each firm would require to in order to continue with its pre-regulation production plan; aggregating these across all affected firms; and multiplying by half the market price of performance credits.[19]

Figure 5.5 Market equilibrium



Based on the preceding analysis for an individual firm, it is clear that the aggregate output of the buyers of performance credits is higher under the trading mechanism than under a system of common individual standards. However, the impact on the output of the sellers is ambiguous, as is the overall impact on the aggregate output of the sector.

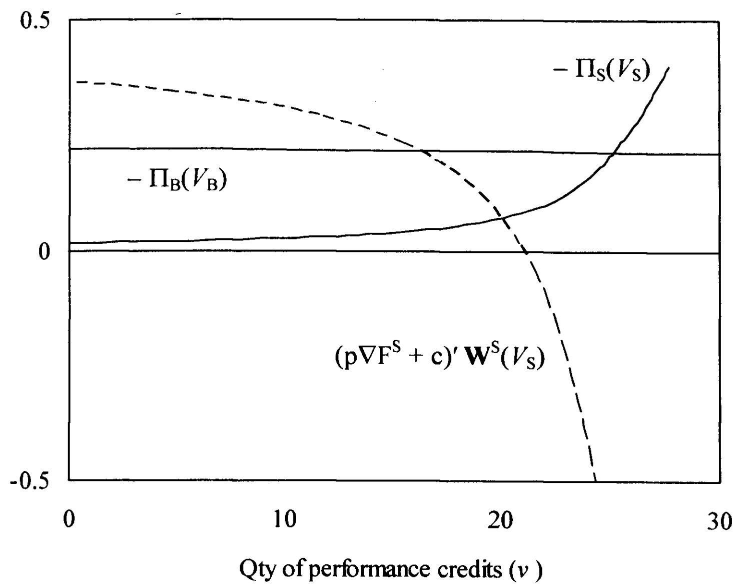
As can be seen in Figure 5.6, which shows the market equilibrium for a particular numerical example, it is possible to have $(p\nabla F^S + \mathbf{b})'\mathbf{W}^S(V_S^{\#}) < 0$.[20] This has two implications. First, if there is an exogenous shift in the demand for performance credits – causing the market price to rise, then aggregate output of the sellers will rise as they sell more credits. Second, it is not clear whether the aggregate output of the sellers is higher under the trading scheme than under a system of common, individual standards,

[19] For the linear aggregate WTA and WTP functions shown in Figure 5.5 this is equal to the aggregate cost. However, in general these functions will be strictly convex, in which case it overstates the cost.

[20] The aggregate production functions are $F^S(\boldsymbol{W}_S) = (-W_{1S})^{0.2}(-W_{2S})^{0.05}$ and $F^B(\boldsymbol{W}_B) = (-W_{1B})^{0.4}(-W_{2B})^{0.4}$; the exogenous prices are p =10 and $\mathbf{c}$ = (3, 0.01); and the performance rule parameters are r = 30, $\mathbf{b}$ = (0, 10).

or whether it is lower. While output decreases for the initial tranche of sales (0-22 in Figure 5.6), it increases for the remaining sales (22-25 in Figure 5.6), and it is not clear *a priori* which response dominates.

Figure 5.6 Market equilibrium for performance credits: a numerical example

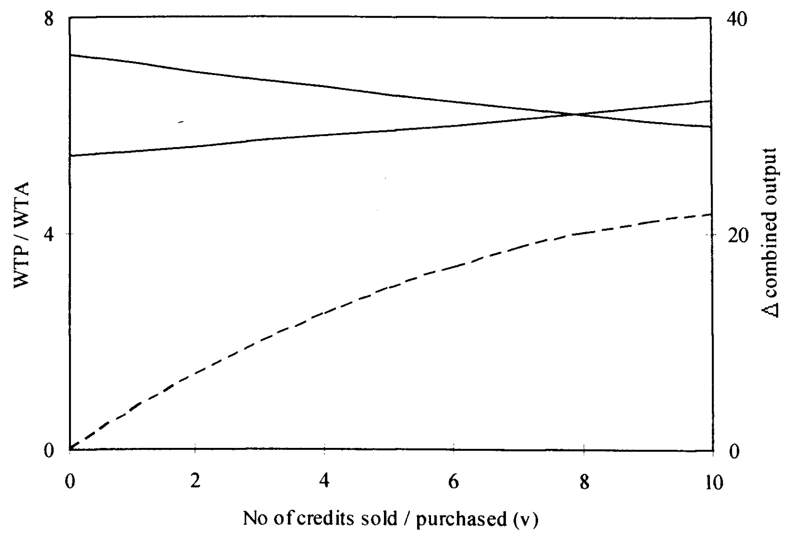


Even if one assumes that the net impact on the aggregate output of the sellers is negative (i.e. their output is lower under trading), it is not possible to draw any conclusions about the impact of trading on the aggregate output of the sector as a whole. As can be seen in Figure 5.7, which shows the aggregate WTP and WTA functions for two particular numerical examples, together with the aggregate output for different transfer quantities, the impact depends on the shapes of aggregate production functions and the values of the various exogenous parameters.[21]

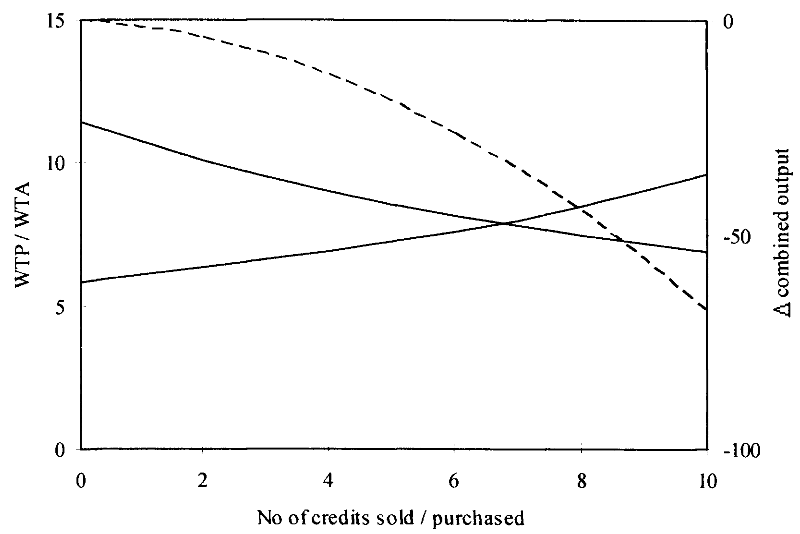
[21] In case (a), the exogenous prices are p = 5 and **c** = (1,1); the rule parameters are r = 0.22 and **b** = (5,0); and the production functions are $F^S(\boldsymbol{W}_S) = (-W_{1S})^{0.2}(-W_{2S})^{0.2}$ and $F^B(\boldsymbol{W}_B) = (-W_{1B})^{0.25}(-W_{2B})^{0.2}$. In case (b), the prices are p = 10 and **c** = (1,1); the rule parameters are r = 0.3 and **b** = (0.5, 0); and the production functions are $F^S(\boldsymbol{W}_S) = 1.5(-W_{1S})^{0.3}(-W_{2S})^{0.1}$ and $F^B(\boldsymbol{W}_B) = (-W_{1B})^{0.3}(-W_{2B})^{0.1}$.

Figure 5.7 Impact of trading on aggregate output

Case a) Combined output increases



Case b) Combined output decreases



In the first case, aggregate output is higher under the trading scheme than under the system of common, individual standards; in the second case it is lower. It follows directly that total energy consumption is higher under trading in the first case, and lower in the second.

5.4 Illustrative simulation

For the purposes of this simulation, the chemicals sector is divided into two distinct sub-sectors – specialty chemicals production, and basic chemicals production; each having three constituent firms. Each sub-sector produces a single product, and uses a different type of energy product (e.g. electricity versus fuel oil). For simplicity it is assumed that there is only one other input – a feedstock product – and that the production technologies of all six firms can be described by a Cobb-Douglas production function of the general form $y = Kw_1^{\mu}w_2^{\nu}$ (where w_1 and w_2 denote the feedstock and energy inputs, and K, μ and ν are firm production specific parameters, with $\mu + \nu < 1$).[22]

Table 5.5 Parameter values for simulation

	Market prices (€/unit)			Production function parameters			Conversion factors	
Firm	Output	Feed stock	Energy	μ	ν	K	a	b
S1	15	8	5	0.525	0.175	1.5	0.1	5.0
S2	15	8	5	0.475	0.200	1.5	0.1	5.0
S3	15	8	5	0.550	0.150	1.5	0.1	5.0
B1	50	8	10	0.200	0.525	0.5	1.0	2.5
B2	50	8	10	0.175	0.625	0.5	1.0	2.5
B3	50	8	10	0.225	0.575	0.5	1.0	2.5

[22] It should be noted that the functional forms and parameter values that have been used in the simulation have been chosen so as facilitate an analytical solution and to accentuate the effects of trading. While they broadly reflect the actual characteristics of the UK chemicals sector (see section 5.1) – at least in relative terms – they are not intended to provide a faithful representation.

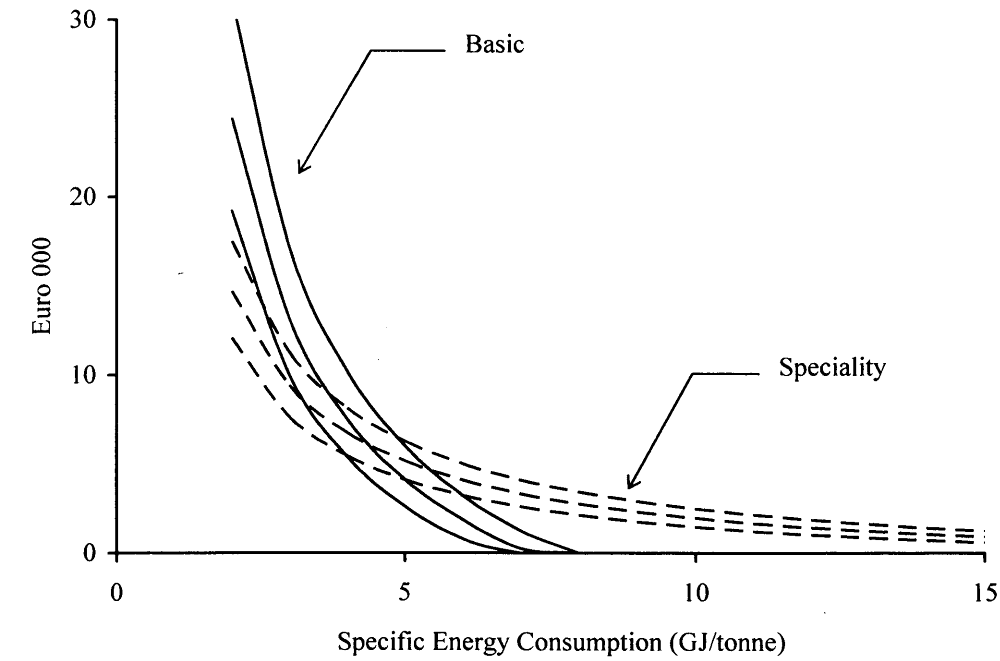
Table 5.5 provides details of the various market prices and parameter values that have been adopted for the simulation, while Table 5.6 shows the resultant "pre-regulation" sector profile. Thus, prior to the introduction of the performance rule, the aggregate profit for the sector is € 64.1 million, and the specific energy consumption rate is 11.3 gigajoules per tonne (GJ/t). In all of the following simulations it is assumed that the target aggregate rate (r) has been set at 7.0 GJ/t, which represents a 38% improvement.

Table 5.6 Pre-regulation sector profile

Firm / Sector	Output (Kt)	Share	Energy Cons. (PJ)	Share	SEC (GJ/t)	Profit (€m)
S1	258	7.4%	6.77	17.2%	26.3	11.6
S2	215	6.2%	6.44	16.3%	30.0	10.5
S3	274	7.8%	6.17	15.6%	22.5	12.3
Speciality	747	21.4%	19.38	49.1%	26.0	34.4
B1	597	17.1%	3.92	9.9%	6.6	8.2
B2	1189	34.1%	9.29	23.5%	7.8	11.9
B3	955	27.4%	6.86	17.4%	7.2	9.6
Basic	2741	78.6%	20.07	50.9%	7.3	29.7
Total	3488	100.0%	39.45	100.0%	11.3	64.1

Figure 5.8 shows the specific marginal costs of energy efficiency for the six firms when transfers of performance credits are prohibited, and each firm has to meet the common standard through its own actions. As can be seen, the selected values for the prices and parameters give rise to a considerable variation between firms, with the shapes of the marginal cost curves differing significantly between the two sub-sectors.

Figure 5.8   Specific marginal cost of energy efficiency



In order to provide a benchmark for assessing the potential benefit of the credit trading mechanism, Table 5.7 shows the outcome when the target rate of 7.0 GJ/t is applied as a common standard to all firms. Overall, there is a 16% reduction in output tonnage and a 49% reduction in energy consumption versus the pre-regulation levels. The achieved aggregate SEC rate is actually slightly lower than the target, reflecting the fact that the pre-regulation rate for firm B1 was already below 7.0 GJ/t.

The total cost for the chemicals sector as a whole is € 10.8 million. However, as one would expect from Figure 5.8, the mechanism has a disproportional impact on the specialty chemicals sub-sector, which is forced to reduce its energy consumption by 84%, and suffers a 30% decline in the value of profits. In contrast, profits for the basic chemicals sub-sector are reduced by less than 2%, reflecting the fact that its pre-regulation SEC rate (at 7.3 GJ/t) is only slightly above the target. Correspondingly, there is a large variation in marginal costs between the firms – ranging from € 4,100 for firm S2 to zero for firm B1 (which does not have to undertake any action).

Table 5.7 Common energy efficiency target

Firm	Change in output	Change in energy cons.	SEC (GJ/t)	Cost (€ million)	Credits Purchased / (Sold) (PJ)	Marginal cost (€ 000)
S1	- 42.3%	- 84.0%	7.0	3.6	n/a	3.33
S2	- 47.6%	- 87.7%	7.0	3.9	n/a	4.11
S3	- 29.6%	- 78.6%	7.0	2.8	n/a	2.60
Specialty	- 38.7%	- 83.5%	7.0	10.3	n/a	
B1	0.0%	0.0%	6.6	0.0	n/a	0.00
B2	- 18.5%	-26.9%	7.0	0.5	n/a	1.24
B3	- 4.2%	- 6.1%	7.0	0.0	n/a	0.25
Basic	- 9.1%	- 14.5%	6.9	0.5	n/a	
Total	- 15.8%	- 48.5%	6.9	10.8	n/a	

Table 5.8 shows the outcome when the target is implemented by the credit trading mechanism, without the inclusion of any performance adjustment factors (i.e. with $\varepsilon_i = 0$ for all firms). This time, there is a 21% reduction in total output, and a 51% reduction in total energy consumption. Interestingly, both of these reductions are greater than the corresponding reductions under the common standard (i.e. they are consistent with case b) in Figure 5.7). As one would expect, firms in the basic chemicals sub-sector generate performance credits, which they sell to firms in specialty sub-sector. Consequently, while the achieved aggregate SEC rate exactly matches the target, the achieved rates for individual firms (i.e. before taking into account the credits bought or sold) vary between 5.7 GJ/t and 11.7 GJ/t.[23]

[23] For example, the introduction of trading allows firm A1 to increase its energy consumption to 1.76 petajoules (versus 1.08 petajoules under the common target), and its output to 171 kilotonnes (versus 154 kilotonnes), giving a physical efficiency rate of 10.3 gigajoules per tonne. However, when the 0.56

The total cost of achieving the target rate, at € 8.3 million, is 24% lower than that incurred under the common target, with each sub-sector enjoying lower costs. However, the benefits of trading are disproportionately weighted towards the basic chemicals sub-sector. Indeed, the revenues that the firms in this sector receive from selling performance credits actually exceed the costs that they incur in improving their energy efficiency by € 1 million.

Table 5.8 Credit trading without performance adjustments

Firm	Change in output	Change in energy cons.	SEC (GJ/t)	Cost (€ million)	Credits Purchased / (Sold) (PJ)	Marginal cost (€ 000)
S1	- 34.6%	- 74.0%	10.3	3.3	0.56	1.77
S2	- 33.3%	- 75.3%	11.7	3.3	0.64	1.77
S3	- 25.9%	- 71.8%	8.8	2.7	0.36	1.77
Specialty	- 32.2%	- 73.7%	10.1	9.3	1.56	
B1	- 11.7%	- 23.0%	5.7	- 0.9	(0.70)	1.77
B2	- 22.7%	- 33.3%	6.8	0.4	(0.22)	1.77
B3	- 15.6%	- 26.8%	6.2	- 0.5	(0.63)	1.77
Basic	- 17.5%	- 29.0%	6.3	- 1.0	(1.56)	
Total	- 20.9%	- 51.0%	7.0	8.3	0.00	

While this simulation is relatively simplistic, it provides a number of interesting insights regarding the impacts of a move from a common individual performance standard, to an aggregate standard implemented by a credit trading mechanism. First, as expected, it reduces the costs of achieving the target rate. While the magnitude of the cost saving will depend on the relative characteristics of the firms covered by the regulatory intervention, it may be relatively significant. Second, the move is likely to change the environmental outcome. This is a major difference compared to the introduction of

petajoules of credits that it purchased are subtracted from its energy consumption, the firm exactly satisfies the post-trading performance target of 7.0 gigajoules per tonne.

trading for absolute limits, where the environmental outcome remains unchanged. However, it need not necessarily mean that the environmental outcome deteriorates. As this example illustrates, it may actually be better under the trading mechanism.

Finally, while all firms benefit from the introduction of the trading mechanism, the distribution of benefits may create problems for the political acceptability of the outcome. Indeed, as this example illustrates, it can lead to a situation where some firms actually benefit from the regulatory intervention. Of course, this problem is not unique to performance-based credit trading. It can also occur under a "cap and trade" permit mechanism for an absolute target, depending on the initial allocation of permits. Fortunately, as has been demonstrated in Chapter 3, it is possible to address the issue by using the performance adjustment factors.

5.5 Impact of performance adjustment factors

It is clear that if there is significant variation in the pre-regulation energy efficiency rates of firms, then the imposition of a common target rate is likely to result in widely differing cost burdens on individual firms. This is true under both the system of fixed individual standards and the trading mechanism. However, as the preceding example illustrates, the introduction of trading can exacerbate the inequity of the cost distribution.[24] Such an outcome may be hard to justify if the sector comprises a number of distinct sub-sectors producing different types of output with substantially different inherent energy intensities – as is the case in the chemicals sector. In this situation, the divergence in the rates of specific energy consumption between firms in different sub-sectors will reflect fundamental differences in the characteristics of the products produced, and the production processes used.

[24] The spread of costs rises from € 3.9 million to € 4.2 million when the trading mechanism is introduced.

Fortunately, the facility for including performance adjustment factors provides a flexible mechanism that can be used to address this issue. In particular, it allows a number of different approaches to be adopted for the determination of the values of the individual factors. If information is available, then the values could be calculated on the basis of the expected costs of improved energy efficiency to individual firms (or sub-sectors). In theory this would allow the calculation of adjustment factors that reflected some notion of equity such as "ability to pay" (Rose *et al* (1998) identify a number of potential equity criteria that could be applied). However, in practice the use of a more mechanistic approach – where adjustment factors are calculated on the basis of some pre-determined rule, may prove to be more acceptable to the firms in the sector.

One such rule might be that each sub-sector should be required to make the same percentage improvement in its aggregate energy efficiency rate, with the resultant sub-sector adjustment factors allocated to constituent firms such that each firm then faces the same *ex ante* "effective" target rate. This approach has the advantage of rewarding firms that have already taken steps to improve their energy efficiency, while recognising the intrinsic differences between sub-sectors.

Suppose that the chemicals sector comprises a fixed number (N) of mutually exclusive sub-sectors, and let $I^{3n} \subset I^3$ denote the set of chemical producers in sub-sector $n \in N$. Denoting the pre-regulation values of the firms' inputs and outputs by $\hat{w}_{ki}$ and $\hat{y}_{ki}$ respectively, then the pre-regulation SEC rate for sub-sector $n \in N$ is:

$$\hat{r}_n \quad = \quad \frac{\sum_{i \in I^{3n}} \sum_{k \in K^{1i-}} b_k(-\hat{w}_{ki})}{\sum_{i \in I^{3n}} \sum_{k \in K^{3i+}} a_k \hat{y}_{ki}}$$

while the pre-regulation SEC rate for the chemicals sector as a whole is:

$$\hat{r} \quad = \quad \frac{\sum_{n \in N} \sum_{i \in I^{3n}} \sum_{k \in K^{1i-}} b_k(-\hat{w}_{ki})}{\sum_{n \in N} \sum_{i \in I^{3n}} \sum_{k \in K^{3i+}} a_k \hat{y}_{ki}}$$

If the aggregate target rate for the sector is r then, under the proposed burden sharing rule, the target rate for firms in sub-sector $n \in \mathrm{N}$ would be $(\mathrm{r}/\hat{\mathrm{r}})\hat{\mathrm{r}}_n$, which implies that the "aggregate" performance adjustment factor for the sub-sector (denoted by ε_n) must be:

$$\varepsilon_n \quad = \quad \left(\frac{\mathrm{r}}{\hat{\mathrm{r}}}\right)(\hat{\mathrm{r}}_n - \hat{\mathrm{r}})\left(\sum_{i \in \mathrm{I}^{3n}} \sum_{k \in \mathrm{K}^{3i+}} \mathrm{a}_k \hat{\mathrm{y}}_{ki}\right)$$

Thus the sign of the adjustment factor for a particular sub-sector depends on whether its pre-regulation SEC rate is above or below the sector average; while the relative magnitude of the factor depends on the extent to which its rate differs from the average, and on its pre-regulation output level (in tonnes). It is straightforward to show that the performance adjustment factors calculated under this rule satisfy the requirement that they sum to zero. That is:

$$\begin{aligned}
\sum_{n \in \mathrm{N}} \varepsilon_n \quad &= \quad \left(\frac{\mathrm{r}}{\hat{\mathrm{r}}}\right) \sum_{n \in \mathrm{N}} (\hat{\mathrm{r}}_n - \hat{\mathrm{r}})\left(\sum_{i \in \mathrm{I}^{3n}} \sum_{k \in \mathrm{K}^{3i+}} \mathrm{a}_k \hat{\mathrm{y}}_{ki}\right) \\
&= \quad \left(\frac{\mathrm{r}}{\hat{\mathrm{r}}}\right) \sum_{n \in \mathrm{N}} \left(\frac{\sum_{i \in \mathrm{I}^{3n}} \sum_{k \in \mathrm{K}^{1i-}} \mathrm{b}_k(-\hat{\mathrm{w}}_{ki})}{\sum_{i \in \mathrm{I}^{3n}} \sum_{k \in \mathrm{K}^{3i+}} \mathrm{a}_k \hat{\mathrm{y}}_{ki}} - \frac{\sum_{n \in \mathrm{N}} \sum_{i \in \mathrm{I}^{3n}} \sum_{k \in \mathrm{K}^{1i-}} \mathrm{b}_k(-\hat{\mathrm{w}}_{ki})}{\sum_{n \in \mathrm{N}} \sum_{i \in \mathrm{I}^{3n}} \sum_{k \in \mathrm{K}^{3i+}} \mathrm{a}_k \hat{\mathrm{y}}_{ki}} \right) \left(\sum_{i \in \mathrm{I}^{3n}} \sum_{k \in \mathrm{K}^{3i+}} \mathrm{a}_k \hat{\mathrm{y}}_{ki}\right) \\
&= \quad \left(\frac{\mathrm{r}}{\hat{\mathrm{r}}}\right) \left(\sum_{n \in \mathrm{N}} \sum_{i \in \mathrm{I}^{3n}} \sum_{k \in \mathrm{K}^{1i-}} \mathrm{b}_k(-\hat{\mathrm{w}}_{ki}) \; - \; \sum_{n \in \mathrm{N}} \sum_{i \in \mathrm{I}^{3n}} \sum_{k \in \mathrm{K}^{1i-}} \mathrm{b}_k(-\hat{\mathrm{w}}_{ki}) \right) \\
&= \quad 0
\end{aligned}$$

If the sub-sector adjustment factor (ε_n) is allocated to constituent firms on the basis of their relative pre-regulation output levels, then the post-trading SEC target for firm $i \in \mathrm{I}^{3n}$ becomes:

$$\frac{\sum_{k\in K^{1i-}} b_k(-w_{ki}) \; + \; v_i}{\sum_{k\in K^{3i+}} a_k y_{ki}} \quad \leq \quad r\left[1 + \left(\frac{\hat{r}_n}{\hat{r}} - 1 \right) \frac{\sum_{k\in K^{3i+}} a_k \hat{y}_{ki}}{\sum_{k\in K^{3i+}} a_k y_{ki}} \right]$$

from which it can be seen that the *ex post* target rate will only be equal to the *ex ante* target rate $(r/\hat{r})\hat{r}_m$ if the output (in tonnes) of the firm remains at the pre-regulation level.

Returning to the numerical example, the application of this burden sharing rule would imply that the adjusted *ex ante* targets for the specialty and basic chemicals sub-sectors should be 16.1 GJ/t and 4.5 GJ/t respectively (i.e. a 38% improvement on the respective pre-regulation levels). Table 5.9 gives the outcome when the resultant sub-sector adjustment factors are allocated to firms on the basis of their pre-regulation output.[25] As expected, the adjustment factors do not result in any change to the overall cost of meeting the target, nor in any changes to the values of the physical variables. In particular, the specific energy consumption rates of the individual firms are completely unaffected. However, there is a major impact on the market for performance credits (although not the equilibrium price), and on the distribution of net costs between the sub-sectors.

Compared with the previous outcome (Table 5.8) the volume of transfers is almost four times greater, and the roles of the two sub-sectors are reversed (i.e. the specialty chemicals producers are now net sellers of performance credits). However, it is the impact on the distribution of net costs that is the most striking. Now, the specialty chemicals producers gain from the intervention, while the basic chemicals producers suffer substantial net costs. Furthermore, the spread of costs is actually greater than

[25] With only two sub-sectors, the magnitude of the adjustment factor is the same in each case, with only the sign differing – positive for the specialty chemicals sub-sector, and negative for the basic chemicals sub-sector. Using the values in Table 5.[x], the adjustment factor for the specialty sub-sector is calculated as (7.0 / 11.3) × (26.0 – 11.3) × 0.747 = 6.8 PJ.

before. Consequently, if the objective is a more equitable distribution of the cost, then the performance adjustment factors have clearly failed!

Table 5.9 Credit trading with performance adjustments

Firm	Change in output	Change in energy cons.	SEC (GJ/t)	Cost (€ million)	Credits Purchased / (Sold) (PJ)	Marginal cost (€ 000)
S1	- 34.6%	- 74.0%	10.3	(0.9)	(1.79)	1.77
S2	- 33.3%	- 75.3%	11.7	(0.1)	(1.32)	1.77
S3	- 25.9%	- 71.8%	8.8	(1.7)	(2.14)	1.77
Speciality	- 32.2%	- 73.7%	10.1	(2.7)	(5.24)	
B1	- 11.7%	- 23.0%	5.7	1.7	0.78	1.77
B2	- 22.7%	- 33.3%	6.8	5.7	2.72	1.77
B3	- 15.6%	- 26.8%	6.2	3.7	1.74	1.77
Basic	- 17.5%	- 29.0%	6.3	11.0	5.24	
Total	- 20.9%	- 51.0%	7.0	8.3	0.00	

5.6 Summary

The most significant observation that can be made from the analysis and the simulation is that the cost of achieving a given target arte for aggregate specific energy consumption is lower under the trading scheme than under a system of common, individual performance standards. While the magnitude of the saving will depend (*inter alia*) on the heterogeneity of the affected firms production functions, it may be significant. This, of course, is to be expected given the demonstrated cost efficiency of the trading mechanism. Of greater interest therefore are some of the other points highlighted by the application.

First, cost efficient achievement of the standard requires that the "effective prices" of all energy inputs, and of all outputs, be raised above their respective market prices. In the first case, the uplift is proportional to the energy content of the input; in the second, it is proportional to the unit weight of the output. Consequently, there is a reduction in the relative use of inputs with high energy content values. In particular, there is a reduction in the use of any energy input compared to any feedstock input.

Second, the shadow price of the aggregate performance rule – and hence the market price of performance credits – can be interpreted as the short run change in the average cost of production per tonne (i.e. holding output constant) resulting from a small change in the target rate for specific energy consumption.

Third, while a rise in the market price of performance credits will always lead to an increase in the number of credits transferred by a firm (i.e. an increase in the number sold, or a reduction in the number purchased), the resultant impact on its output and energy consumption will depend (ceteris paribus) on the number of credits that it has already transferred. If the firm is a buyer of credits, or a small seller, then its output will decline, and its energy consumption will decline by more than the increase in the number of credits transferred. However, if the firm is a large seller, then its output will increase, and the reduction in its energy consumption will be less than the change in its credit transfers.

Fourth, aggregate output – and hence aggregate energy consumption – may be higher, or lower, under the trading scheme than under a system of common, individual performance standards. Thus the introduction of trading may improve, or weaken, the environmental effectiveness of the regulatory intervention.

Fifth, it is possible to determine an upper bound for the aggregate cost of the intervention under the trading scheme from the market price of the performance credits, and the firms' pre-regulation levels of output and energy consumption. This will be close to the actual cost if the aggregate WTP / WTA curves for performance credits are approximately linear (i.e. they are not too convex).

Finally, while the introduction of the trading scheme will reduce the aggregate cost of the regulatory intervention, it may well exacerbate the inequity of the cost distribution. This problem can be addressed by setting non-zero values for the performance adjustment factors. However, as the simulation illustrates, care needs to be taken when using mechanistic burden sharing rules to calculate these factors, as apparently egalitarian rules may actually make the problem worse.

Chapter 6 Packaging recovery targets under extended producer responsibility

In this chapter the policy context is extended producer responsibility for packaging, and performance-based credit trading is used to implement a target rate for the recovery of waste packaging, expressed as a percentage of the amount used. However, as was the case in the previous chapter, the various insights that are derived are not specific to packaging, and they are directly transferable to other non-durable products. In particular, they would apply equally to the case of a recovery target for newspapers and magazines.[1]

The chapter starts with a brief introduction to the concept of extended producer responsibility, and a description of how the principle has been applied in the United Kingdom in the case of packaging. The second section of the chapter provides a detailed description of the model that has been adopted for the analysis. This encapsulates the salient features of the packaging system in the United Kingdom – in particular the institutional framework for waste management. However, as in the previous chapter, the objective is to use the model to explore the properties of the trading mechanism, rather than to provide a detailed and faithful representation of the actual system. The model differs from the generic model used in Chapter 3 in that it recognises that the production system may be affected by other regulatory interventions.

[1] As was noted in Chapter 1, such a target has been proposed in the United Kingdom – although it has not been implemented.

In particular, it allows for the possibility that subsidies may be used to reduce – or eliminate – the direct cost of waste collection services to householders.

Following a brief review of the conditions for a regulated aggregate cost minimum, the majority of the analysis presented in the third section focuses on the joint market equilibrium under the trading scheme. In particular, it looks at the impact of the regulatory intervention on the markets for packaging and diverted waste packaging, and considers the relationship between these markets and the market for performance credits. The understanding of this relationship allows a simple expression to be derived that can be used to estimate the aggregate cost of the regulatory intervention. The analysis is supplemented by an illustrative simulation in the fourth section of the chapter, in which a simple numerical example is used to show how the market equilibrium responds to increases in the target recovery rate under different assumptions regarding the assignment of initial property rights, and to explore the distributional impacts of the trading scheme.

6.1 Background

6.1.1 Extended Producer Responsibility

Under extended producer responsibility (EPR), producers are required to assume responsibility for the life-cycle environmental impacts of their products, with particular emphasis on the post-consumer stage. This responsibility may be defined in financial, or in physical terms. That is, producers may be held responsible for (a proportion of) the waste management costs associated with their products, or they may be required to take the products back and arrange for their disposal or recovery. Early examples of EPR include the German Packaging Ordinance and the Dutch Packaging Covenant, both implemented in 1991, and the 1993 Producer Responsibility initiative in the United

Kingdom for six priority waste streams.[2] More recently, the EU Directive on End of Life Vehicles[3] has set recycling targets, and stipulated that producers must pay "all or a significant part" of the costs of take-back and treatment from January 2007.

EPR can be viewed as an umbrella principle, or concept, which can support a range of different policy goals and objectives, and which can be implemented by a number of different policy instruments (Macauley & Walls, 2000). However, while the exact manifestation may vary between applications, any EPR initiative would be expected to exhibit two key characteristics:

- a shift in financial and / or physical responsibility for post-consumer waste upstream to the producer, and away from local authorities;
- the provision of incentives for producers to incorporate environmental considerations into the design of their products.

There is a continuum of institutional approaches that has been adopted for the implementation of EPR; ranging from voluntary unilateral commitments by individual firms or industry sectors, through negotiated agreements between government and industry, to statutory legislation. For example, in the case of packaging waste, a negotiated agreement was used in the Netherlands, while the initiatives in Germany and the United Kingdom were both underpinned by legislation.

While take-back requirements are the policy instrument most closely associated with EPR, they are not the only one that can be used. Depending on the underlying objective, a variety of economic instruments can be used to provide incentives for changes in behaviour. These include deposit / refund schemes, advanced disposal fees,

---

[2] The six priority waste streams were packaging, newspapers, tyres, batteries, vehicles and electronic equipment.

[3] 2000/53/EC

and material taxes. Of particular relevance to recycling targets is an upstream combined tax / subsidy (UCTS). As its name suggests, this hybrid instrument combines an advanced disposal fee (including appropriate external costs) with a subsidy to recycling (Dinan, 1993; Palmer & Walls 1999).

Take-back requirements can take several forms, but there is commonly a requirement to recover (i.e. divert from the waste stream going to landfill) a given proportion of "end-of-life" products, or component materials. However, whatever form they take, a common characteristic these initiatives is a recognition that while individual producers should have an obligation to meet the cost of recovery, it is not economically efficient to require them to undertake the necessary recovery themselves. One common approach is to allow producers to discharge their individual obligations by joining a collective compliance scheme (or producer responsibility organisation). The scheme then assumes responsibility for ensuring that sufficient waste is collected and recovered to satisfy the aggregate obligation; with the cost of these activities funded by the membership fees. The best known example of this approach is the Duales System Deutschland (DSD) for packaging in Germany.

An alternative approach, which maintains the concept of individual responsibility, has evolved over recent years in the United Kingdom in relation to the recovery targets for waste packaging that were set out in the 1994 EU Directive on Packaging and Packaging Waste – the so-called PRN scheme.

6.1.2 United Kingdom PRN scheme

The UK Packaging Regulations[4] were introduced in March 1997 in response to the EU Directive[5], which had set targets for recovery and recycling of packaging waste for the

[4] Producer Responsibility Obligations (Packaging Waste) Regulations 1997

[5] Directive on Packaging and Packaging Waste (94/62/EC)

year 2001. These statutory Regulations set a series of annual targets for the proportion of packaging that was to be recovered and recycled from 1998 to 2001. The targets – which are shown in Table 6.1 – were subsequently revised for the later years, and were extended through to 2003.

Table 6.1 Packaging recycling and recovery targets

	UK Packaging Regulations		EU Directive		
Year	Total recovery	Material recycling	Total recovery	Total recycling	Material recycling
1998	38%	7%			
1999 [(a)]	43%	10%			
2000 [(b)]	45%	11%			
2001 [(c)]	56%	18%	50%	25%	15%
2002	59%	19%			
2003	59%	19%			

(a) Original targets: recovery 38%, recycling 7%

(b) Original targets: recovery 43%, recycling 11%

(c) Original targets: recovery 52%, recycling 16%

Thus the Regulations impose targets at two levels: an aggregate recovery target across all packaging materials; and a common "threshold" recycling target (i.e. excluding energy recovery) that applies to individual materials. However, with the exception of plastics, the actual recycling rates of the individual packaging materials exceed threshold value. Therefore, in effect it is the aggregate recovery target that drives the system.[6]

[6] Proposed revisions to the EU Directive would set significantly higher targets for 2008, with the threshold recycling rates differentiated by material.

In addition to setting the recovery and recycling targets, the Packaging Regulations also impose a number of obligations on individual producers and users of packaging materials that satisfy certain threshold conditions.[7] These "obligated producers" are (*inter alia*) required to:

- register with the relevant Environment Agency each year, and provide data on the amount of packaging materials handled in the previous year;
- take reasonable steps to meet their individual recycling and recovery obligations;
- provide the Environment Agency with annual of certificates of compliance, supported by acceptable evidence that recycling and recovery has actually taken place;

By 2001 over 6000 obligated producers had registered. As an alternative to registering directly with the Environment Agency, an obligated producer can choose to register with a collective compliance scheme, which – in return for the payment of an annual fee – will assume responsibility for discharging its obligations. The recycling and recovery obligations for each compliance scheme are then calculated as the sum of its members' individual obligations. The large majority of producers have chosen to follow this route, with compliance schemes accounting for around 85% of total registrations.

The individual recovery obligations are determined according to the principle of shared responsibility, under which the total obligation is broken down between the different stages of the packaging chain, with 6% assigned to raw material manufacturers; 9% to material converters; 37% to packer-fillers; and 48% to sellers. So for example, if a

[7] The Packaging Regulations apply only to businesses that handle more than 50 tonnes of packaging, and that have a turnover of more than £2 million. It was estimated that companies below these threshold levels accounted for around 9% of total packaging in 2001. Hence, the recovery rate for obligated companies was set at 56% in order to ensure that the overall recovery rate exceeded 50%, as required by the Directive.

retailer handled 100 tonnes of packaging in 2001, its recovery obligation would be 26.9 tonnes (i.e. $100 \times 0.56 \times 0.48$).

The Packaging Regulations required obligated producers to provide proper evidence that waste packaging had actually been recovered. However, the precise definition of what would be considered acceptable was left to subsequent statutory Guidance from the Environment Agency. When this was published in July 1997, it introduced a standardised form of evidence – called *Packaging Waste Recovery Notes* (PRNs) – that would be issued by accredited reprocessor.[8] Under this voluntary scheme, reprocessors must submit quarterly reports on amount of waste packaging that they have received, and are allowed to issue PRNs up to this amount. By 2001, over 300 reprocessors and exporters had gained accreditation under the scheme. While they are not the only form of evidence that will be accepted by the Environment Agency, the administrative burden associated with the use of other forms has resulted in the almost universal use of PRNs as evidence of compliance.

An important point to note about the PRN scheme is that it was not originally intended as a trading mechanism. The consultations undertaken during the development of the Regulations had focussed on the level of the recovery and recycling targets, and on the assignment of the responsibility for meeting the targets. Much less attention was paid to how the targets would be achieved. In particular, no consideration was given to the potential for using a system of tradable credits to implement the targets. While there was a general acceptance of the advantages of a market-led approach, it was envisaged that this would be achieved largely through a system of collective compliance schemes,

---

[8] The Guidance also made provision for PRNs to be issued by overseas reprocessors that were covered by recognised accreditation schemes operated by local or national governments. However, this was replaced in 1999 by a separate accreditation scheme for exporters of waste packaging materials for reprocessing. Companies satisfying certain requirements are given Agency Accredited Exporter Status, and are allowed to issue Packaging Waste Export Recovery Notes (PERNs). For simplicity, the acronym PRN is used for both types of evidence.

which would use the fees that they received from their members to fund the necessary expansion of collection and reprocessing activities.[9]

The main motivations for the introduction of the PRN scheme were the deterrence of fraud by reprocessors, and the reduction of the administrative burden on the obligated businesses. PRNs were seen as providing evidence that a transaction involving diverted waste packaging had taken place; not as a separate commodity with its own value.[10] However, since there was nothing to prevent the PRNs from being transferred, and a producer only had to provide evidence that sufficient waste packaging had been diverted to meet its obligation – not that it had undertaken (or contracted out) the diversion itself, the creation of a transparent and guaranteed "common currency" soon led to the development of a separate market in evidence.

By 2001 the revenues from the sales of PRNs had risen to around £70 million. The large majority of these sales are covered by long term contracts between the reprocessors and compliance schemes. However, the number of spot trades and secondary transactions is increasing rapidly. For example, in the first five years of the scheme, annual sales of PRNs through the web-based Environment Exchange[11] grew at a compound rate of 96%. By the end of 2002 it had more than 600 members (including 16 compliance schemes), and accounted for approximately 7% of all PRN sales. Interestingly, around a quarter of the purchases were made by reprocessors. It is possible that these purchases may have been speculative. However, a more likely

---

[9] The Packaging Regulations did allow companies to meet their own recovery and recycling obligations; either through their own activities; or by contracting with third parties to collect waste packaging on their behalf and arrange delivery to reprocessors.

[10] This is reflected in the fact that neither the Packaging Regulations, nor the Environment Agency Guidance, considered the issue of the assignment of the initial property rights to the PRNs. However, the *de facto* assignment has been to the reprocessors – reflecting their administrative responsibility for issuing PRNs and for providing an audit trail.

[11] Details can be found at the Environment Exchange website http://www.t2e.co.uk

reason is that the reprocessors needed additional PRNs in order to satisfy contractual commitments to compliance schemes. Thus while the PRN scheme has some of the characteristics of a collective, contract-based mechanism, it is evolving into a market-based mechanism.

This evolution has not been without its critics, and a number of concerns have been expressed about the operation of the scheme (Bailey, 1999; O'Doherty & Bailey, 2001; DETR, 1998; ENDS, 2001). In particular, it has been claimed that the revenues from PRN sales have been taken by the reprocessors as "windfall profits", rather than being used to develop the reprocessing and collection infrastructures as was intended; and that the small number of firms in some reprocessing sectors, together with barriers to new entrants, has allowed the reprocessors to exercise market power – raising the price of PRNs and increasing the costs of the obligated producers.[12]

In response to these concerns, a number of amendments were made to the PRN system in 1999, with the intention of clarifying initial property rights and preventing the possibility of speculation by non-obligated parties. These prohibited reprocessors from selling PRNs to anyone other than obligated producers and compliance schemes (or their agents), and required that obligated producers delivering waste packaging be given first refusal on the resultant PRNs. Reprocessors were also required to provide an annual return to the Environment Agency showing the total revenue generated from sales of PRNs in the previous calendar year, and the proportions of this revenue used to fund the expansion of the collection infrastructure; additional investment in reprocessing capacity; and the development of end-use markets.

[12] The implications of strategic behaviour by reprocessors are investigated in Chapter 8.

6.2 Model

As with the previous application, a partial equilibrium model is adopted for the analysis; with only those commodities that are salient to the packaging system being treated as endogenous. The model comprises nine mutually exclusive sectors, and eight endogenous commodities. These are defined in Table 6.2.

Table 6.2 Definition of sectors and commodities

Sectors		**Commodities**	
$I^1 \subset I$	Packaging producers	$k = 1$	Packaging
$I^2 \subset I$	Packer-fillers	$k = 2$	Packaged good
$I^3 \subset I$	Content producers	$k = 3$	Content good
$I^4 \subset I$	Householders	$k = 4$	Consumption good
$I^5 \subset I$	Consumers	$k = 5$	Waste collection service
$I^6 \subset I$	Local authorities	$k = 6$	Waste disposal service
$I^7 \subset I$	Waste disposers	$k = 7$	Waste diversion service
$I^8 \subset I$	Waste diverters	$k = 8$	Diverted waste packaging
$I^9 \subset I$	Waste reprocessors		

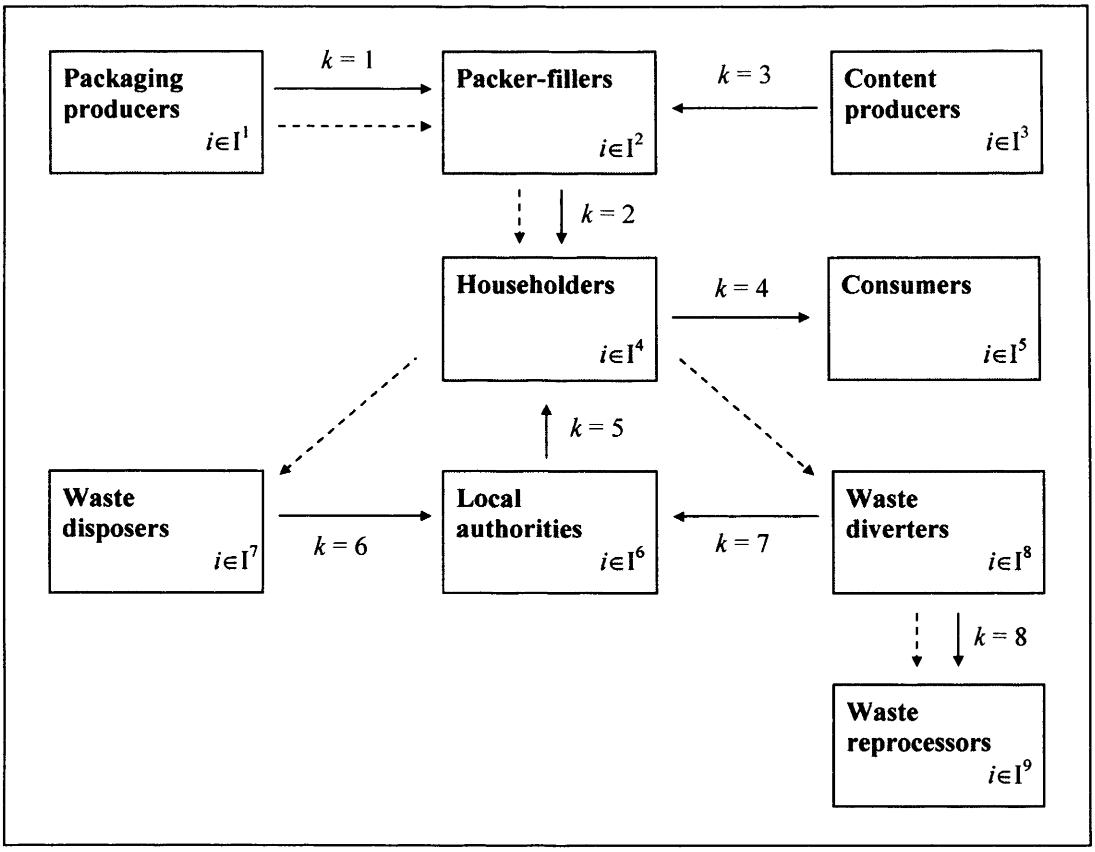
Each commodity $k \in K$ comes in a fixed number of variants – indexed by $m \in M$, where each variant is associated with a particular packaging material.[13, 14] For example, if the set of packaging materials is {glass, plastic, cardboard}, then the respective variants of packaging might be {bottle, punnet, box}; the respective variants of the content good might be {wine, strawberries, cornflakes}; and so on. Thus the double

[13] This assumption is not as restrictive as it may appear at first sight. For example it does not preclude the definition of composite packaging materials (e.g. plastic and paper), or the distinction between different variants of the same packaging material (e.g. green versus brown glass, etc.). However, it does mean that the packaging material associated with a particular content good is predetermined and cannot be changed.

[14] It should be noted that the definition of the set M in this chapter is different to that used in Chapters 2 – 4, where it denoted the set of non-market environmental commodities.

index $km \in$ KM is used to denote the variant of commodity $k \in$ K associated with packaging material $m \in$ M. This leads to some slight (and obvious) changes in notation. For example, the set I^{km+} denotes those agents that produce the variant of commodity k that is associated with material m, while M^i denotes the set of variants that are relevant to agent $i \in$ I.

Figure 6.1 Packaging system



The relationships between the various sectors and commodities (i.e. which sectors produce and use each commodity) are defined in Figure 6.1, which provides a schematic representation of the packaging system. The solid arrows denote the economic flows of goods and services, and the broken arrows denote the physical flows of packaging materials.[15]

[15] For the sake of simplicity, the flows of exogenous commodities have been omitted.

Thus, packer-fillers produce packaged goods, using *packaging* and *content goods* as inputs. These are then unpacked by householders, to release the *consumption goods* and generate a requirement for *waste collection services* for the discarded packaging. While responsibility for providing these services lies with the local authorities, it is assumed that they do not actually collect the waste. Rather, they buy-in *waste disposal services* (in the form of landfill) and *waste diversion services* from sub-contractors, who undertake the collection on their behalf. The *diverted waste packaging* is then sent for reprocessing.

While this stylised representation captures the salient features of the packaging system – at least for the United Kingdom, it is of course a simplification of reality. Some sectors (such as retailers) have been omitted for the sake of simplicity, and it has been implicitly assumed that the reprocessed material produced by the waste reprocessors is not used as an input to the production of packaging goods (i.e. that the system is "open loop"). The first simplification is not too troublesome, since – under perfect competition (as is assumed here) – it is always possible to vertically integrate firms. Thus packer-fillers can be interpreted as a composite of packer-fillers and retailers; packaging producers as a composite of raw material manufacturers and converters; and so on. While the second simplification is a reasonable reflection of reality for some packaging materials (e.g. plastic), it is not so for others, where waste packaging is collected explicitly to be reprocessed for use as a packaging input (e.g. aluminium beverage containers). However, it can be interpreted as an assumption that reprocessed and virgin materials are perfect substitutes in the production of packaging products and that the latter has a constant price (i.e. an infinitely elastic supply), which may not be too unreasonable.

It is assumed that all production functions are additively separable by associated packaging material $m \in M$. For example, while an individual packaging producer may make glass bottles and plastic punnets, the production of the two products is completely independent. The production functions of packer-fillers and households are all Leontief (i.e. inputs are used in fixed proportions), while those of the local authorities are additive.[16] For the sake of simplicity, all exogenous inputs (such as labour) are ignored. Consequently, the following production constraints apply:

$$\gamma_m y_{2mi} + w_{1mi} \leq 0 \text{ and } y_{2mi} + w_{3mi} \leq 0 \qquad \text{for all } i \in \mathrm{I}^2,\ m \in \mathrm{M}^i \quad \ldots (6.1.a)$$

$$y_{4mi} + w_{2mi} \leq 0 \text{ and } \gamma_m y_{4mi} + w_{5mi} \leq 0 \qquad \text{for all } i \in \mathrm{I}^4,\ m \in \mathrm{M}^i \quad \ldots (6.1.b)$$

$$y_{5mi} + w_{6mi} + w_{7mi} \leq 0 \qquad \text{for all } i \in \mathrm{I}^6,\ m \in \mathrm{M}^i \quad \ldots (6.1.c)$$

where the parameter γ_m represents the quantity of packaging material (in kilogrammes) required per unit of the packaged good.

The production technologies of the agents in the other sectors are not modelled explicitly. Rather, reduced-form functions are used for the agents' marginal costs of production (MC), in the case of packaging producers, content producers, waste diverters and waste disposers; or for their marginal willingness's-to-pay (WTP), in the case of consumers and waste reprocessors. These are denoted respectively by: [17]

$$\mathrm{c}^{kmi}(y_{kmi}) \leq 0 \quad \text{with} \quad \mathrm{c}^{kmi\,\prime} \leq 0 \qquad \text{for all } i \in \mathrm{I}^{1,3,7,8},\ km \in \mathrm{KM}^{i+}$$

$$\mathrm{p}^{kmi}(w_{kmi}) \leq 0 \quad \text{with} \quad \mathrm{p}^{kmi\,\prime} \leq 0 \qquad \text{for all } i \in \mathrm{I}^{5,9},\ km \in \mathrm{KM}^{i-}$$

It follows from the properties of the underlying production functions that all of these functions are concave, and that $\mathrm{c}^{kmi}(0) = 0$ and $\mathrm{p}^{kmi}(0) = -\infty$ for the respective

[16] The adoption of a Leontief production function for packer-fillers precludes the redesigning of packaging as a response to the policy intervention (i.e. reducing the parameter γ_m). Consequently, the only way for producers to reduce the amount of packaging that they use is to reduce production of the packaged good.

[17] The following notational definitions are adopted: $\mathrm{I}^{1,3,5} = \mathrm{I}^1 \cup \mathrm{I}^3 \cup \mathrm{I}^5$, and $\mathrm{I}^{1:3} = \mathrm{I}^1 \cup \mathrm{I}^2 \cup \mathrm{I}^3$, etc.

commodities / variants.[18] The definitions may seem rather counterintuitive – particularly in the case of the WTP functions. However, they merely reflect the convention of representing inputs by negative values. Hence an increase in the value of an input variable represents a reduction in the amount used.

In order to simplify the analysis, it is assumed that the marginal costs of production of all content producers, and of all waste disposers, are constant and equal (within sector and material), i.e.

$$c^{3mi}(y_{2mi}) \equiv -c_{3m} \qquad \text{for all } i \in I^3,\ m \in M^i$$

$$c^{6mi}(y_{6mi}) \equiv -c_{6m} \qquad \text{for all } i \in I^7,\ m \in M^i$$

Furthermore, if one views waste diversion services as a by-product of the production of diverted waste packaging, then it follows that:

$$c^{7mi}(y_{7mi}) \equiv 0 \qquad \text{for all } i \in I^8,\ m \in M^i$$

$$y_{7mi} - y_{8mi} \leq 0 \qquad \text{for all } i \in I^8,\ m \in M^i \qquad \ldots (6.2)$$

The regulatory intervention takes the form of a target rate for the minimum aggregate quantity of diverted waste packaging that is sent for reprocessing, relative to the aggregate quantity of packaging that is used by the packer-fillers. That is:

$$\frac{\sum_{m\in M}\sum_{i\in I^{8m+}} y_{8mi}}{\sum_{m\in M}\sum_{i\in I^{2m-}} (-w_{1mi})} \geq r$$

[18] The properties of the production functions are defined in assumptions A8-A12 in Chapter 3. The concavity of the WTP and marginal cost functions follows from the fact that the objective functions in the underlying maximization problems are concave, and the constraint functions are all convex. (see Theorem 21.23 in Simon & Blume, 1994)

where $r \in (0,1)$ is the target recovery rate. Thus the values of the parameters in the aggregate performance rule are:

$$\begin{aligned}
\alpha_{km} &= 0 & \beta_{km} &= r & &\text{for } k = 1, m \in M\\
&= 0 & &= 0 & &\text{for } k = 2, \ldots 7, m \in M\\
&= 1 & &= 0 & &\text{for } k = 8, m \in M\\
\delta_{\cdot} &= 0
\end{aligned}$$

The resultant linear constraint is:

$$\sum_{m \in M} \sum_{i \in I^{8m+}} y_{8mi} + r \sum_{m \in M} \sum_{i \in I^{2m-}} w_{1mi} \geq 0 \qquad \ldots (6.3)$$

It is assumed that the target diversion rate r is always greater than (or equal to) the market rate achieved in the absence of any regulation (denoted by r^0). Consequently, the aggregate performance rule is always binding (i.e. the constraint (6.3) holds as an equality).

Without any loss of generality, it is assumed that the assignment parameter for each commodity is the same for all associated packaging materials (i.e. $\theta_{km} = \theta_k$ for all $m \in M$). Consequently, the individual performance rule parameters are:

$$\begin{aligned}
\rho_{km} &= -\theta_k r & \sigma_{km} &= (1-\theta_k) r & &\text{for } k = 1, m \in M\\
&= 0 & &= 0 & &\text{for } k = 2, \ldots 7, m \in M\\
&= 1-\theta_k & &= -\theta_k & &\text{for } k = 8, m \in M
\end{aligned}$$

Furthermore, the performance adjustment factors (ε_i) are set to zero for all those agents that do not produce or use the two commodities included in the aggregate performance rule. Consequently, the individual performance rules are:

$$v_i \leq \theta_1 \, r \sum_{m \in M'} (-y_{1mi}) + \varepsilon_i \qquad \text{for all } i \in I^1 \qquad \ldots (6.4.a)$$

$$v_i \leq (1 - \theta_1) \, r \sum_{m \in M'} w_{1mi} + \varepsilon_i \qquad \text{for all } i \in I^2 \qquad \ldots (6.4.b)$$

$$v_i \leq (1 - \theta_8) \sum_{m \in M'} y_{8mi} + \varepsilon_i \qquad \text{for all } i \in I^8 \qquad \ldots (6.4.c)$$

$$v_i \leq \theta_8 \sum_{m \in M'} (-w_{8mi}) + \varepsilon_i \qquad \text{for all } i \in I^9 \qquad \ldots (6.4.d)$$

$$v_i \leq 0 \qquad \text{for all } i \in I^{3:7} \qquad \ldots (6.5.e)$$

Thus, content producers, householders, consumers, local authorities, and waste disposers are all inactive in the market for performance credits. Firms in the other four sectors may be active or inactive, depending on the values that are set for the two assignment parameters, and for their performance adjustment factors.

The price paid by householders for waste collection services may be subject to an *ad valorem* subsidy equal to $(1 - \varpi)$, where $0 \leq \varpi \leq 1$. If $\varpi = 1$ then there is no subsidy, and householders are charged directly for the service; if $\varpi < 1$ then the cost of the subsidy is funded by the imposition of lump-sum taxes on householders $i \in I^4$.[19]

[19] Direct charging for waste collections services is becoming increasingly common. It is in widespread use in seven European countries (ACR, 1999), and has been adopted by over 4000 communities across the USA, accounting for around 12% of the population (Miranda *et al*, 1998). However, in many countries – including the United Kingdom – waste collection services are still funded entirely from general taxation (i.e. $\varpi = 0$).

Consequently, the individual benefits of the various agents in the packaging system, including the financial value of performance credit transfers (where relevant), are:

$$\pi_i \;=\; \sum_{k\in \mathrm{K}^{i+}} \sum_{m\in \mathrm{M}^i} \left[\; \mathrm{p}_{km}\, y_{kmi} \;+\; \int_0^{y_{kmi}} \mathrm{c}^{kmi}(\xi)\, d\xi \;\right] \;+\; \mathrm{q}\, v_i \qquad \text{for all } i \in \mathrm{I}^{1,3,7,8} \quad \dots (6.6.\mathrm{a})$$

$$\pi_i \;=\; \sum_{m\in \mathrm{M}^i} (\,\mathrm{p}_{2m}\, y_{2mi} \;+\; \mathrm{p}_{1m}\, w_{1mi} \;+\; \mathrm{p}_{3m}\, w_{3mi}\,) \;+\; \mathrm{q}\, v_i \qquad \text{for all } i \in \mathrm{I}^{2} \quad \dots (6.6.\mathrm{b})$$

$$\pi_i \;=\; \sum_{m\in \mathrm{M}^i} (\,\mathrm{p}_{4m}\, y_{4mi} \;+\; \mathrm{p}_{2m}\, w_{2mi} \;+\; \varpi\, \mathrm{p}_{5m}\, w_{5mi}\,) \;+\; \mathrm{T}_i \;+\; \mathrm{q}\, v_i$$

$$\text{for all } i \in \mathrm{I}^{4} \quad \dots (6.6.\mathrm{c})$$

$$\pi_i \;=\; \sum_{m\in \mathrm{M}^i} (\,\mathrm{p}_{5m}\, y_{5mi} \;+\; \mathrm{p}_{6m}\, w_{6mi} \;+\; \mathrm{p}_{7m}\, w_{7mi}\,) \;+\; \mathrm{q}\, v_i \qquad \text{for all } i \in \mathrm{I}^{6} \quad \dots (6.6.\mathrm{d})$$

$$\pi_i \;=\; \sum_{k\in \mathrm{K}^{i-}} \sum_{m\in \mathrm{M}^i} \left[\; \int_0^{w_{kmi}} \mathrm{p}^{kmi}(\xi)\, d\xi \;+\; \mathrm{p}_{km}\, w_{kmi} \;\right] \;+\; \mathrm{q}\, v_i \qquad \text{for all } i \in \mathrm{I}^{5,9} \quad \dots (6.6.\mathrm{e})$$

where $\mathrm{p}_{km} \geq 0$ is the market price of commodity / variant $km \in \mathrm{KM}$, $\mathrm{q} \geq 0$ is the market price of performance credits, and $\mathrm{T}_i \leq 0$ is the lump-sum tax imposed on agent $i \in \mathrm{I}^4$. In the case of firms, these expressions represent profits; in the case of local authorities and householders, they represent surpluses; and in the case of consumers, they represent the financial value of the utility derived from the consumption goods.

Since the full cost of any subsidy is borne by the householders, the aggregate gross economic benefit for the packaging system is given by:

$$\Pi \;=\; \sum_{i\in \mathrm{I}^{5,9}} \sum_{k\in \mathrm{K}^{i-}} \sum_{m\in \mathrm{M}^i} \left[\; \int_0^{w_{kmi}} \mathrm{p}^{kmi}(\xi)\, d\xi \;\right] \;+\; \sum_{i\in \mathrm{I}^{1,3,7,8}} \sum_{k\in \mathrm{K}^{i+}} \sum_{m\in \mathrm{M}^i} \left[\; \int_0^{y_{kmi}} \mathrm{c}^{kmi}(\xi)\, d\xi \;\right] \qquad \dots (6.6)$$

That is, it is equal to the sum of the gross benefits accruing to the consumers of the consumption good and to the reprocessors of the diverted waste packaging, less the sum of the gross production costs incurred by the packaging producers, content good producers, waste disposers and waste diverters.

6.3 Analysis

The majority of the analysis presented in this section focuses on the joint market equilibrium under the trading scheme. However, in order to put this into context, it is instructive to consider briefly the conditions for the regulated aggregate cost minimum.

6.3.1 Regulated aggregate cost minimum

By Proposition 3.2 of Chapter 3, the solution to the regulated cost minimum problem is interior and all of the constraints are binding (i.e. all of the shadow prices are strictly positive). Consequently, under the assumptions set out in section 6.2, the following sets of necessary conditions can be derived for each packaging material $m \in$ M:[20]

$$- c^{1mi\prime}(y_{1mi}^{*}) \quad = \quad \eta_{1m}^{*} \quad = \quad \frac{1}{\gamma_m}\left[\, \eta_{2m}^{*} - \eta_{3m}^{*} \,\right] - r\,\lambda^{*}$$

$$\text{for all } i \in I^{1m+} \qquad \ldots (6.7.a)$$

$$\eta_{2m}^{*} + \gamma_m\, \eta_{5m}^{*} \quad = \quad \eta_{4m}^{*} \quad = \quad - p^{4mi\prime}(w_{4mi'}^{*})$$

$$\text{for all } i \in I^{4m-} \qquad \ldots (6.7.b)$$

[20] Together with the production constraints (6.1) and (6.2), the aggregate performance constraint (6.3), and the resource constraints for each commodity / variants $km \in$ K, these conditions are also sufficient for a regulated aggregate cost minimum.

$$- c^{8mi\prime}(y_{8mi}^{*}) \; - \; \eta_{7m}^{*} \quad = \quad \eta_{8m}^{*} \quad = \quad - p^{8mi'\prime}(w_{8mi'}^{*}) \; + \; \lambda^{*}$$

$$\text{for all } i \in I^{8m+},\ i' \in I^{8m-} \qquad \ldots (6.7.c)$$

$$c_{3m} \quad = \quad \eta_{3m}^{*} \qquad \ldots (6.7.d)$$

$$c_{6m} \quad = \quad \eta_{6m}^{*} \quad = \quad \eta_{5m}^{*} \quad = \quad \eta_{7m}^{*} \qquad \ldots (6.7.e)$$

where λ^* is the shadow value of the aggregate performance rule, and η_{km}^* is the shadow value of the resource constraint for commodity / variant $km \in$ KM. Assuming that the respective marginal cost and marginal benefit functions all have strictly negative derivatives, then it follows that the solution is unique for the argument variables.[21]

Condition (6.7.a) requires that the net marginal private benefit of each packaging product must exceed its marginal cost of production by an amount equal to the shadow price of the aggregate performance rule multiplied by the target diversion rate. In contrast, condition (6.7.c) requires that, for each packaging material, the net marginal cost of waste diversion (i.e. the excess cost over landfill) must exceed the marginal private benefit by an amount equal to the shadow price of the diversion constraint. The important point to note is that, for both sets of conditions, the magnitude of the difference between the (net) marginal cost and the (net) marginal benefit is the same for all packaging materials. Finally, condition (6.7.b) requires that the marginal cost of production of each consumption good must incorporate the shadow price of waste collection for the associated packaging, which is equal to the marginal cost of landfill.

[21] Because the production functions of the agents in the other sectors all exhibit constant returns to scale (either explicitly or implicitly), the individual values of their respective inputs and outputs are indeterminate. Only the aggregate values for these sectors can be determined.

6.3.2 Market equilibrium

For the joint market equilibrium with performance credits, the equivalent necessary first-order conditions for each packaging material $m \in \mathrm{M}$ are:[22]

$$-\,c^{1mi\prime}(y_{1mi}^{\#}) \quad = \quad p_{1m} - \theta_1\, r\, q \quad = \quad \frac{1}{\gamma_m}\,[\, p_{2m} \;-\; p_{3m}\,] \;-\; r\,q$$

$$\text{for all } i \in \mathrm{I}^{1m+} \qquad \ldots\ (6.8.\text{a})$$

$$p_{2m} \;+\; \varpi\, \gamma_m\, p_{5m} \quad = \quad p_{4m} \quad = \quad -\,p^{4mi\prime}(w_{4mi'}^{\#})$$

$$\text{for all } i \in \mathrm{I}^{4m-} \qquad \ldots\ (6.8.\text{b})$$

$$-\,c^{8mi\prime}(y_{8mi}^{\#}) \;-\; p_{7m} \quad = \quad p_{1m} + \theta_8\, q \quad = \quad -\,p^{8mi\prime}(w_{8mi'}^{\#}) \;+\; q$$

$$\text{for all } i \in \mathrm{I}^{8m+}\,,\ i' \in \mathrm{I}^{8m-} \qquad \ldots\ (6.8.\text{c})$$

$$c_{3m} \quad = \quad p_{3m} \qquad \ldots\ (6.8.\text{d})$$

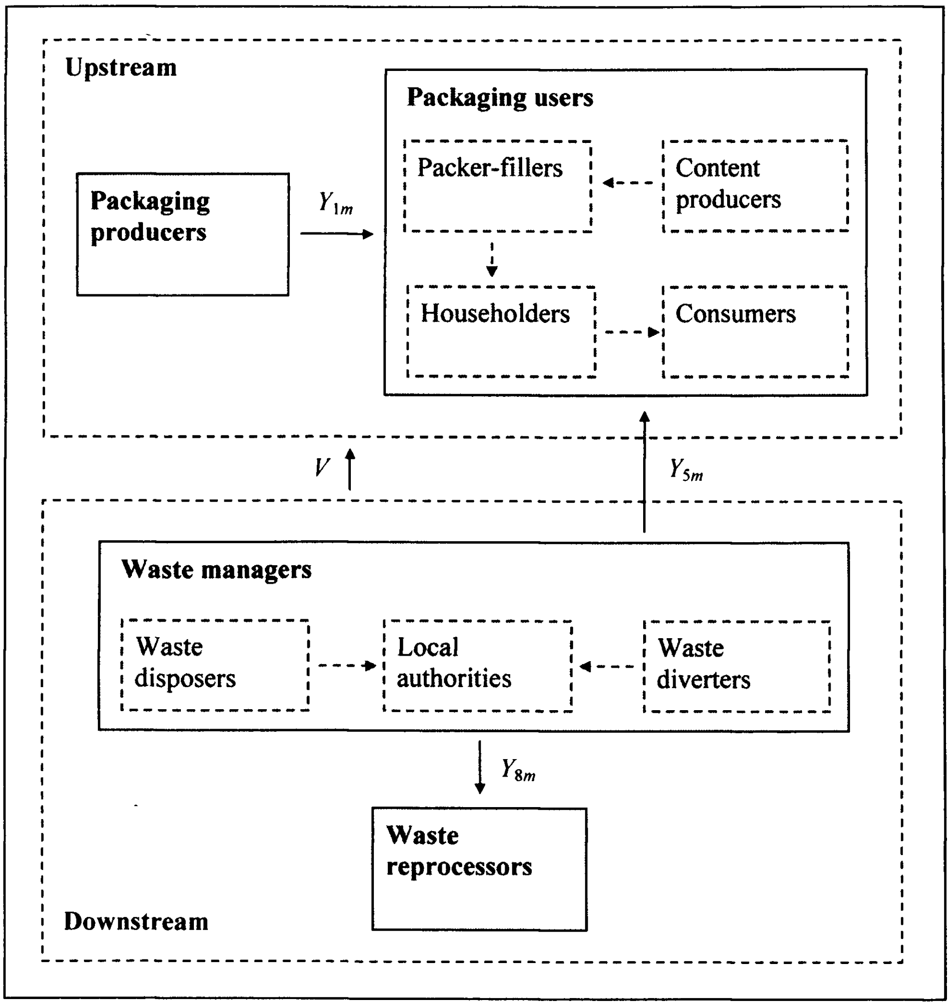
$$c_{6m} \quad = \quad p_{6m} \quad = \quad p_{5m} \quad = \quad p_{7m} \qquad \ldots\ (6.8.\text{e})$$

There are two main points to note from these conditions. First, as expected, the values of the assignment parameters θ_1 and θ_8 affect only the market prices of packaging and diverted waste packaging respectively. They do not affect the equilibrium values of the real variables, or any of the other market prices. Second, if $\varpi < 1$ then the resultant market equilibrium cannot coincide with the regulated aggregate cost minimum, except in the unlikely scenario that the marginal cost of landfill (and hence the waste collection cost) is equal to zero. In particular, the amount of packaging used will be greater than the cost efficient level. Put another way, a necessary condition for the cost efficiency of

[22] Together with the production constraints (6.1) and (6.2), the individual performance rules (6.4), and the market clearing conditions for each commodity / variant $km \in \mathrm{KM}$, these conditions are also sufficient.

the trading mechanism is that householders face the full marginal cost of waste collection.[23]

Figure 6.2 Simplified packaging system



[23] If waste collection is subsidised (i.e. $\varpi < 1$), then the cost efficiency of the trading mechanism will be restored if a tax is imposed on each packaging product, equal to the uncollected cost of waste collection (i.e. $t_{1m} = (1 - \varpi)c_{6m}$). A similar conclusion is drawn by Fullerton & Wu (1998) in relation to the overall welfare efficiency of the packaging system. Using a general equilibrium model, they show that achievement of the social optimum requires either that households pay the full social cost of disposal (i.e. including the costs of any externalities); or that a tax is imposed on producers' use of packaging, accompanied by a subsidy to encourage recyclable designs.

Further analysis of the market equilibrium is facilitated by amalgamating some of the sectors, and aggregating agents within sectors. The resultant, simplified packaging system is shown in Figure 6.2. There are now only four sectors; with packaging users representing an amalgamation of packer-fillers, content producers, householders and consumers; and waste managers representing an amalgamation of local authorities, waste disposers and waste diverters. Correspondingly, the number of markets is reduced to four: packaging, waste collection services, diverted waste packaging, and performance credits.

For each material $m \in M$, denote the aggregate WTP functions for the associated consumption good and for the diverted waste packaging by $P^{4m}(\cdot)$ and $P^{8m}(\cdot)$ respectively, and denote the aggregate MC functions for packaging and diverted waste packaging by $C^{1m}(\cdot)$ and $C^{8m}(\cdot)$ respectively. Then for a given price of performance credits, the market equilibrium conditions – defined in terms of the respective (positive) aggregate output quantities – for each variant of the three commodities are:

$$\tilde{P}^{1m}(Y_{1m}^{\#};\, p_{5m}) \quad = \quad \tilde{p}_{1m} \quad = \quad \tilde{C}^{1m}(Y_{1m}^{\#}) \;+\; r\,q \qquad \text{... (6.10.a)}$$

$$\tilde{P}^{5m}(Y_{5m}^{\#};\, \tilde{p}_{1m}) \quad = \quad \varpi\, p_{5m} \quad = \quad \varpi\, \tilde{C}^{5m}(Y_{5m}^{\#}) \qquad \text{... (6.10.b)}$$

$$\tilde{P}^{8m}(Y_{8m}^{\#}) \quad = \quad \tilde{p}_{8m} \quad = \quad \tilde{C}^{8m}(Y_{8m}^{\#}) \;-\; q \qquad \text{... (6.10.c)}$$

where

$$\tilde{p}_{1m} \quad \equiv \quad p_{1m} + (1-\theta_1)\, r\, q$$

$$\tilde{p}_{8m} \quad \equiv \quad p_{8m} - (1-\theta_8)\, q$$

and:

$$\tilde{p}^{1m}(Y_{1m};\, p_{5m}) \quad \equiv \quad \text{Max}\ [\,0\,,\ \frac{1}{\gamma_m}\left(-p^{4m}\left(\frac{-Y_{1m}}{\gamma_m}\right) - c_{3m}\right) \ - \ \varpi\, p_{5m}\ ]$$

$$\tilde{p}^{5m}(Y_{5m};\, \tilde{p}_{1m}) \quad \equiv \quad \text{Max}\ [\,0\,,\ \frac{1}{\gamma_m}\left(-p^{4m}\left(\frac{-Y_{1m}}{\gamma_m}\right) - c_{3m}\right) \ - \ \tilde{p}_{1m}\ ]$$

$$\tilde{p}^{8m}(Y_{8m}) \quad \equiv \quad -\,P^{8m}(-Y_{8m})$$

$$\tilde{C}^{1m}(Y_{1m}) \quad \equiv \quad -\,C^{1m}(Y_{1m})$$

$$\tilde{C}^{5m}(Y_{5m}) \quad \equiv \quad c_{6m}$$

$$\tilde{C}^{8m}(Y_{8m}) \quad \equiv \quad -\,C^{8m}(Y_{8m}) \ - \ c_{6m}$$

The "effective prices" $\tilde{p}_{1m}$ and $\tilde{p}_{8m}$ represent respectively the net price of packaging to the packaging users (i.e. including the cost of acquiring performance credits to meet their share of the resultant obligation), and the net price of diverted waste packaging to the waste reprocessors (i.e. including the revenue received from selling their share of the resultant property right). It is clear from (6.10.a) and (6.10.c) that while the actual market prices are affected by the values of the assignment factors θ_1 and θ_8 (as was noted above), the effective prices are not. It is also clear that if $q > 0$, then the actual and effective prices will coincide if and only if the value of the respective assignment factor is set equal to one.

By construction, all of the aggregate WTP and MC functions are non-negative – as are their arguments. This allows for a more intuitive interpretation of the market equilibrium in terms of (inverse) supply and demand curves. Noting that for each material, the price of the waste collection service is constant (albeit with a possible divergence between consumer and producer prices), condition (6.10.a) can be restated in terms of an independent (i.e. fixed) "excess WTP" function for packaging, and condition (6.10.c) can be restated in terms of an independent "excess MC" function for diverted waste packaging. That is, for each material $m \in$ M:

$$\hat{\mathrm{P}}^{1m}(Y_{1m}) \equiv \tilde{\mathrm{P}}^{1m}(Y_{1m}; \varpi\, c_{6m}) - \tilde{C}^{1m}(Y_{1m}) = r\,q \qquad \text{... (6.11.a)}$$

$$\hat{C}^{8m}(Y_{8m}) \equiv \tilde{C}^{8m}(Y_{8m}) - \tilde{\mathrm{P}}^{8m}(Y_{8m}) = q \qquad \text{... (6.11.b)}$$

It follows from the properties of the underlying MC and WTP functions that the excess WTP functions are all decreasing in output, while the excess MC functions are all increasing in output. Since – by definition – excess WTP and excess MC are equal to zero in the absence of any regulation, conditions (6.11.a) and (6.11.b) imply that for each material, the aggregate quantity of packaging used is lower than its pre-regulation level, while the aggregate quantity of diverted waste packaging is greater.

It is straightforward to show that in any market equilibrium:

$$\sum_{m\in M}\left[\int_0^{Y^{\#}_{1m}} \hat{\mathrm{P}}^{1m}(\xi)\,d\xi + \int_0^{Y^{\#}_{8m}}\left(-\hat{C}^{8m}(\xi)\right)d\xi\right]$$

$$= \sum_{m\in M}\left[\int_0^{Y^{\#}_{1m}}\left(\tilde{\mathrm{P}}^{1m}(\xi; c_{6m}) - \tilde{C}^{1m}(\xi)\right)d\xi + \int_0^{Y^{\#}_{8m}}\left(\tilde{\mathrm{P}}^{8m}(\xi) - \tilde{C}^{8m}(\xi)\right)d\xi\right]$$

$$= \sum_{m\in M}\left[\int_0^{Y^{\#}_{1m}} C^{1m}(\xi)\,d\xi + \int_0^{Y^{\#}_{3m}} C^{3m}(\xi)\,d\xi + \int_0^{Y^{\#}_{6m}} C^{6m}(\xi)\,d\xi + \int_0^{Y^{\#}_{8m}} C^{8m}(\xi)\,d\xi\right]$$

$$+ \sum_{m\in M}\left[\int_{W^{\#}_{4m}}^{0} P^{4m}(\xi)\,d\xi + \int_{W^{\#}_{8m}}^{0} P^{8m}(\xi)\,d\xi\right] + (1-\varpi)\sum_{m\in M}\left[c_{6m}Y^{\#}_{5m}\right]$$

$$= \Pi^{\#} + (1-\varpi)\sum_{m\in M}\left[c_{6m}Y^{\#}_{5m}\right]$$

Thus, if households face the full marginal cost of waste collection (i.e. $\varpi = 1$), then the sum of the surpluses across all of the markets for packaging and all of the markets for diverted waste packaging is equal to the aggregate gross economic benefit for the system. However, if households do not face the full marginal cost, then the sum overstates the aggregate benefit by an amount equal to the uncharged collection cost. It follows directly that the aggregate cost of the regulatory intervention is:

$$\Pi^{u} - \Pi^{r} \quad = \quad \sum_{m \in M} \left[\int_{Y_{1m}^{r}}^{Y_{1m}^{u}} \hat{P}^{1m}(\xi)\, d\xi \;+\; \int_{Y_{8m}^{u}}^{Y_{8m}^{r}} \hat{C}^{8m}(\xi)\, d\xi \right]$$

$$- \;(1 - \varpi) \sum_{m \in M} c_{6m} \left[Y_{5m}^{u} - Y_{5m}^{r} \right] \qquad \ldots (6.12)$$

where the superscripts "u" and "r" denote the unregulated and regulated equilibrium values.

The second term on the right-hand side – which represents the reduction in the uncharged cost of waste collection – is non-negative. Consequently, the sum of the deadweight losses in the various markets for packaging and diverted waste packaging provides an upper bound on the aggregate cost of the regulatory intervention. If households are charged the full cost of waste collection (i.e. $\varpi = 1$), then the sum of the deadweight losses is equal to aggregate cost.

Turning to the market for performance credits, and denoting the (positive) aggregate net quantity of performance credits transferred from the downstream sector to the upstream sector by V, the market equilibrium condition is:

$$\widetilde{P}^{c}(V^{\#}; E_{U}, r) \quad = \quad q \quad = \quad \widetilde{C}^{c}(V^{\#}; E_{D}) \qquad \ldots (6.13)$$

where

$$\tilde{P}^c(V; E_U, r) \equiv \frac{1}{r}\,\hat{P}^1\!\left(\frac{V+E_U}{r}\right) \qquad \text{if} \quad V + E_U < V^{max}(r)$$

$$\equiv 0 \qquad\qquad \geq V^{max}(r)$$

$$\tilde{C}^c(V; E_D) \equiv 0 \qquad \text{if} \quad V - E_D \leq V^{min}$$

$$\equiv \hat{C}^8(V - E_D) \qquad\qquad > V^{min}$$

$\hat{P}^1(\cdot)$ and $\hat{C}^8(\cdot)$ are the horizontal sums of the excess WTP and excess MC functions across all packaging materials; E_U and E_D are the sums of the performance adjustment factors of the agents in the upstream and downstream sectors respectively; $V^{max}(r)$ is gross obligation (i.e. before taking account of performance adjustment factors) of the upstream sector at the pre-regulation level of packaging use; and V^{min} is the quantity performance credits generated by the downstream sector at the pre-regulation level of diversion. By definition, $V^{max}(r)$ is strictly greater than V^{min} for all values of $r > r^0$.

For all $[V^{min} + E_D] \leq V \leq [V^{max}(r) - E_U]$, the quantity of packaging used by the upstream sector is given by the function $Y^1(V) \equiv (V + E_U)/r$, and the quantity of waste packaging diverted by the downstream sector by the function $Y^8(V) \equiv V - E_D$. Noting that $Y_1{}^r = Y^1(V^{\#})$ and $Y_1{}^u = Y^1(V^{max}(r))$, and that $Y_8{}^r = Y^1(V^{\#})$ and $Y_8{}^u = Y^1(V^{min})$, it follows by the Fundamental Theorem of Calculus that for a given value of r:

$$\int_0^{\infty} \text{Min}\left[\tilde{C}^c(\xi; E_D),\ \tilde{P}^c(\xi; E_D, r)\right] d\xi$$

$$= \int_{V^{min}}^{V^{\#}} \tilde{C}^c(\xi; E_D)\, d\xi \;+\; \int_{V^{\#}}^{V^{max}(r)} \tilde{P}^c(\xi; E_U, r)\, d\xi$$

$$= \int_{Y^8(V^{min})}^{Y^8(V^{\#})} \hat{C}^8(Y^8(\xi))\, d\xi \;+\; \frac{1}{r}\int_{Y^1(V^{\#})}^{Y^1(V^{max}(r))} \hat{P}^1(Y^1(\xi))\, d\xi$$

$$= \int_{Y_8^u}^{Y_8^r} \hat{C}^8(\xi)\,d\xi \quad + \quad \int_{Y_1^r}^{Y_1^u} \hat{P}^1(\xi)\,d\xi$$

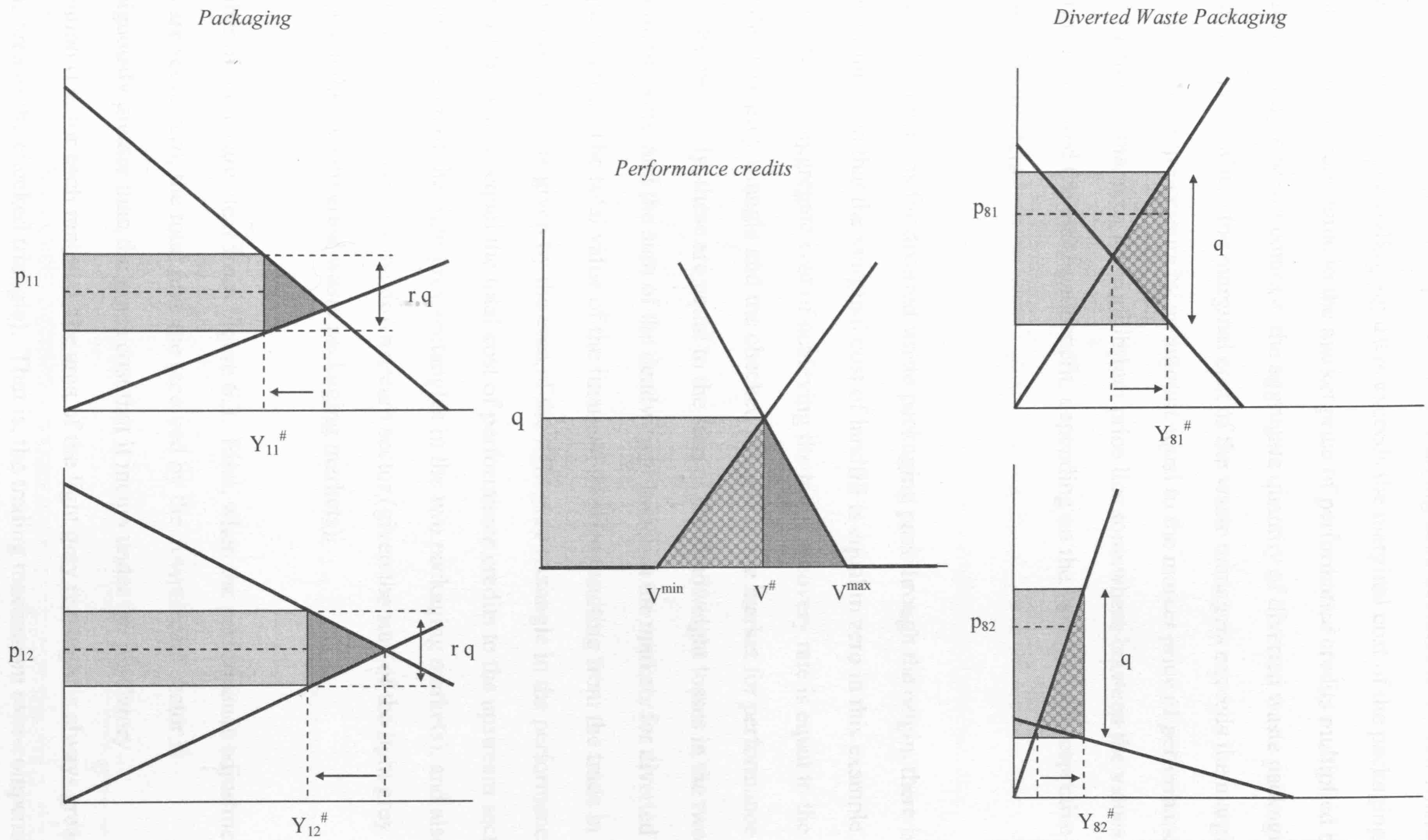
$$= \quad [\,\Pi^u - \Pi^r\,] \quad + \quad (1-\varpi) \sum_{m \in M} c_{6m} \left[\, Y_{5m}^u - Y_{5m}^r \,\right]$$

That is, the area under the lower envelope of the of the WTP and MC curves for performance credits is equal to sum of the deadweight losses in the markets for packaging and diverted waste packaging, and hence provides an alternative measure of the upper bound for the aggregate cost of the regulatory intervention – or the actual cost if householders are charged directly for waste collection. This is a significant observation, as it means that it is possible to calculate the value of the upper bound by considering the single market for performance credits, rather than having to determine the aggregate supply and demand curves in each of the markets for packaging and diverted waste packaging – of which there may be many.

Of course, the calculation of the area under the envelope curve still requires the determination of functional forms of the aggregate WTP and MC curves for performance credits, which may not be possible. However, a simple approximation of the area is provided by the expression $\frac{1}{2} \times [\, V^{max}(r) - V^{min}\,] \times q$. Since the WTP and MC curves are both convex, this estimate will be greater than or equal to the actual area under the envelope. However, this potential reduction in the precision of the estimate is more than offset by the simplicity of the expression; with the values of all three components readily being available.

These various observations are illustrated in Figure 6.3, for the case of two packaging materials, assuming that the performance adjustment factors of all agents are set to zero (i.e. $E_D = E_U = 0$).

Figure 6.3 Market equilibrium with two packaging materials



For each material, the aggregate quantity of packaging used declines to a point where the marginal WTP of the packaging users exceeds the marginal cost of the packaging producers by an amount equal to the market price of performance credits multiplied by the target recovery rate. In contrast, the aggregate quantity of diverted waste packaging increases to a point where the marginal cost of the waste managers exceeds the marginal WTP of the waste reprocessors by an amount equal to the market price of performance credits. In all four markets, the equilibrium price lies somewhere between the values of the marginal cost and the marginal benefit, depending on the values of the respective assignment parameters.

Since the MC functions for diverted waste packaging pass through the origin, there is an implicit assumption that the marginal cost of landfill is equal to zero in this example. Consequently, the aggregate cost of achieving the target recovery rate is equal to the sum of the dark grey triangle and the checked triangle in the market for performance credits. Respectively, these are equal to the sum of the deadweight losses in the two packaging markets, and the sum of the deadweight losses in the markets for diverted waste packaging. The total value of the financial transfer resulting from the trade in performance credits is given by the area of the light grey rectangle in the performance credit market. This is equal the total cost of performance credits to the upstream sector (given by the sum of the light grey rectangles in the two packaging markets), and also to the total revenue received by the downstream sector (given the sum of the light grey rectangles in the two diverted waste packaging markets).

A number of points are clear from Figure 6.3. First, when the performance adjustment factors are set to zero, the total revenue received by the downstream sector is unambiguously greater than the gross cost that it incurs under the regulatory intervention (i.e. for each material, the area of the light grey rectangle is always greater than the area of the checked triangle). That is, the trading mechanism over-compensates the downstream sector. However, it is not possible to draw any unambiguous conclusion regarding the relative magnitudes of the financial transfer and the total cost incurred by the system as a whole. If the reduction in packaging use in the upstream sector is relatively low, then the value of the financial transfer will be greater than the aggregate cost. However, if there is a significant reduction in packaging use, it is possible that the value of the transfer will understate the aggregate cost.

Second, unless all materials have the same diversion rate in the market equilibrium (which will be equal to the target rate), some materials will generate more performance credits than is required to meet the obligations arsing from their use, while others will generate fewer credits than is required. Consequently, the trading of performance credits will result in a financial transfer from materials with low equilibrium diversion rates (for which the demand for credits will exceed the supply) to materials with high diversion rates (which have excess supply).

Finally, it is clear that the same equilibrium outcome could have been achieved by imposing a uniform tax on all packaging goods – equal to the market price of performance credits multiplied by the target recovery rate (r × q), and providing a uniform subsidy payment for all diverted waste packaging – equal to the market price of credits (q). Such a scheme would be revenue-neutral, since $(r\ q)\ Y_1^{\#} = q\ (r\ Y_1^{\#}) = q\ Y_8^{\#}$.

Condition (6.13) can be re-stated to give the following implicit relationship between the number of credits that are traded and the target recovery rate:

$$\tilde{P}^c(V;\mathrm{E_U}, r) \;-\; \tilde{C}^c(V;\mathrm{E_D}) \;\equiv\; \frac{1}{r}\,\hat{P}^1\!\left(\frac{V+\mathrm{E_U}}{r}\right) \;-\; \hat{C}^8(V-\mathrm{E_D}) \;=\; 0$$

Two things are clear from this relationship. First, changes in the target diversion rate affect the number of credits traded through changes to the WTP function only. Second, the impact will depend on the functional forms of the aggregate excess WTP function for packaging and the aggregate excess marginal cost function for diverted waste packaging. An interesting case arises when these are both linear, taking the forms:

$$\hat{p}^1(\xi) \equiv a - b\,\xi$$

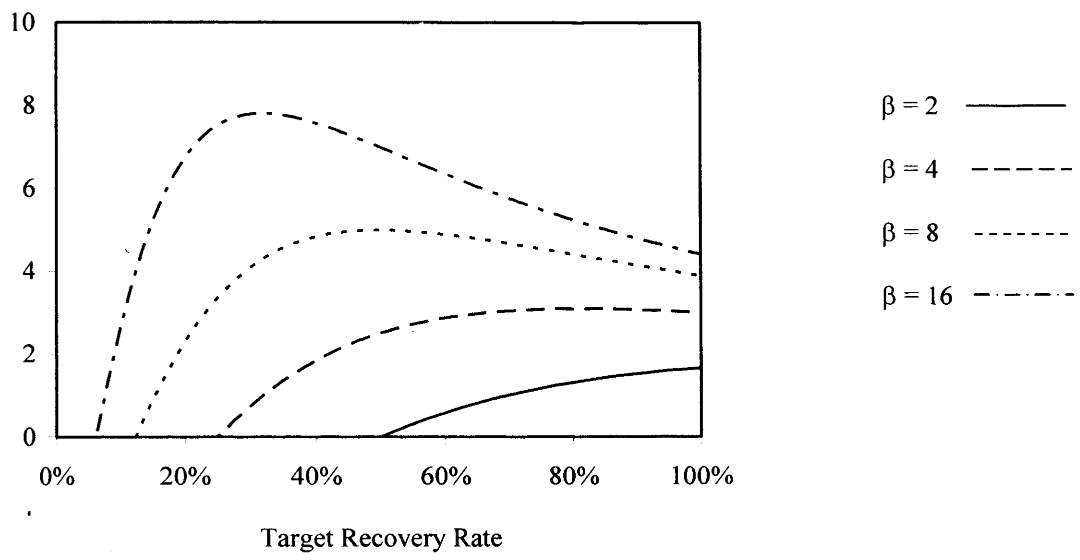
$$\hat{C}^8(\xi) \equiv c\,\xi - d$$

If $E_U = E_D = 0$, then the market equilibrium quantity of performance credits traded is:

$$V^{\#}(r) = \frac{r(a + d\,r)}{b + c\,r^2}$$

It is straightforward to show that if $\beta > 1 + 2\alpha$ (where $\beta = c / b$ is the relative value of the slopes of the two functions, and $\alpha = d / a$ is the relative value of the two intercepts) then there will be a range of values for the target recovery rate, over which increases in the target rate lead to reductions in the number of performance credits traded. It follows directly from the supply side of (6.13) that this will cause the market price of credits to fall, leading to a reduction in the total value of the financial transfer.

Figure 6.4 Market price of performance credits (q)



This is illustrated in Figure 6.4, which shows the relationship between the target diversion rate and the price of performance credits for a range of values of β (with $\alpha = 1$). When the excess MC function is twice as steep as the excess WTP function, the price rises continuously as the target rate increases. However, if the excess MC function is sixteen times steeper, then the price of performance credits falls as the target rate increases above 30%.

6.4 Simulation

With the exception of the final part, the previous section provided a static analysis of the market equilibrium for a given value of the target recovery rate. In this section, a simple numerical example is used to investigate how the market equilibrium changes as the target recovery rate is increased, and to explore the distributional impacts of the trading scheme. In this example there are only two packaging materials – glass and plastic; with the respective packaged goods being a bottle of champagne and a punnet of strawberries. The following aggregate functional forms and parameter values are assumed:

Glass (m = 1)	*Plastic (m = 2)*
$C^{11}(Y_{11}) = -0.8\,Y_{11}$	$C^{12}(Y_{12}) = -0.4\,Y_{12}$
$C^{31}(Y_{31}) = -2$	$C^{32}(Y_{22}) = -2$
$C^{61}(Y_{61}) = -1$	$C^{62}(Y_{62}) = -1$
$C^{81}(Y_{81}) = -2\,Y_{81}$	$C^{82}(Y_{82}) = -4\,Y_{82}$
$P^{41}(W_{41}) = -12 - W_{41}$	$P^{42}(W_{42}) = -10 - W_{42}$

$P^{81}(W_{81}) = -6.5 - W_{81}$ $\qquad$ $P^{82}(W_{82}) = -6.5 - W_{82}$

$\gamma_1 = 1$ $\qquad$ $\gamma_1 = 1$

Thus, the MC functions for packaging and diverted waste packaging, and the WTP functions for consumption goods and diverted waste packaging are all linear. The only differences between the two materials relate to the marginal cost functions of the two packaging goods (with glass bottles having the steeper slope); the marginal cost functions of waste diversion services (with plastic having the steeper slope); and the WTP functions for the respective consumption goods (with champagne having the higher "choke price"[24]).

It is assumed that households are charged the full marginal cost of waste collection (which is equal to the constant marginal cost of landfill). Consequently, for these functional forms and parameter values, the excess aggregate WTP functions for packaging, and the excess aggregate MC functions for diverted waste packaging are:

$\hat{P}^{11}(Y_{11}) = 9 - 1.8\, Y_{11}$ $\qquad$ $\hat{P}^{12}(Y_{12}) = 7 - 1.4\, Y_{12}$

and

$\hat{C}^{81}(Y_{81}) = 3\, Y_{81} - 7.5$ $\qquad$ $\hat{C}^{82}(Y_{82}) = 5\, Y_{82} - 7.5$

Consequently, the "market" rates of recovery in the absence of any diversion target are 50% for glass and 30% for plastic, with an overall recovery rate of 40%.[25] Taking the

---

[24] That is, the price at which demand falls to zero.

[25] The market rate of diversion is derived by finding the values of v and z that set the excess demand and excess supply equal to zero (e.g. $Y_{11}{}^{u} = 9 / 1.8 = 5$ and $Y_{81}{}^{u} = 7.5 / 3 = 2.5$), and dividing the first value by the second.

horizontal sums of respective material-specific functions yields the following aggregate excess WTP and excess MC functions:

$$\hat{p}^1(Y_1) = 9 - 1.8\,Y_1 \qquad \text{for} \quad Y_1 < 1.1111$$

$$= 7.875 - 0.7875\,Y_1 \qquad \text{for} \quad Y_1 \geq 1.1111$$

$$\hat{c}^8(Y_8) = 1.875\,Y_8 - 7.5$$

Thus, if the performance adjustment factors are all set to zero – as is assumed initially, then the inverse demand and supply curves for performance credits are:

$$\tilde{p}^c(V) = \frac{9}{r} - \frac{1.8}{r^2}V \qquad \text{for} \quad V < 1.1111\ r$$

$$= \frac{7.875}{r} - \frac{0.7875}{r^2}V \qquad \text{for} \quad V \geq 1.1111\ r$$

$$\tilde{c}^c(V) = 1.875\,V - 7.5$$

When r = 0.4 (i.e. the target is set equal to the pre-regulation recovery rate), $V^{\#} = 4$ and $\tilde{p}^c(4) = \tilde{c}^c(4) = 0$. That is, the market price of performance credits is zero.

As the target recovery rate is increased, the price rises continuously – at a declining rate of increase (see Figure 6.5). In the light of the discussion at the end of the previous section this is to be expected, noting that the relative value of the intercepts of the excess MC and WTP functions is $\alpha = 7.5 / 7.875 = 0.95$, and that the relative value of the respective slopes is $\beta = 1.875 / 0.7875 = 2.4$, and hence that $\beta < 1 + 2\alpha$.

Figure 6.5 Price of performance credits (€)

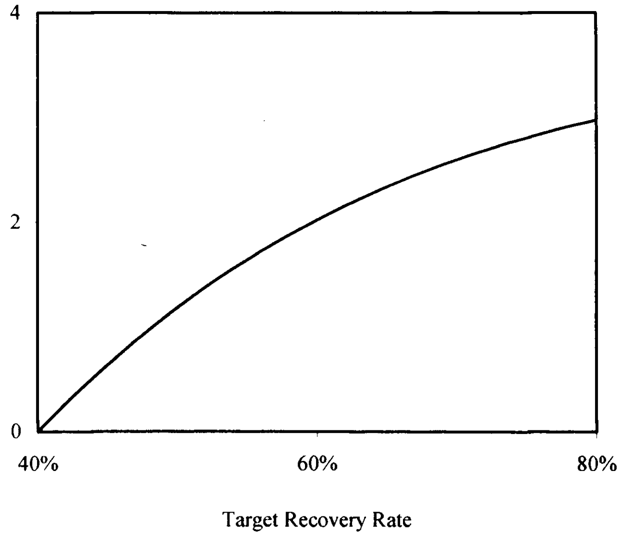


Figure 6.6 shows how the achieved diversion rate of each material changes as the aggregate target rate is increased. As one would expect, the diversion rate increases for both materials – rising from 50% to 95% for glass, and from 30% to 63% for plastic. Somewhat surprisingly the absolute increase is greatest for glass, which had the higher pre-regulation recovery rate (although the increase is lower in percentage terms). Consequently, the "recovery gap" between the two materials at an 80% aggregate target recovery rate has grown to over 30% points.

Figure 6.6 Diversion rates by material

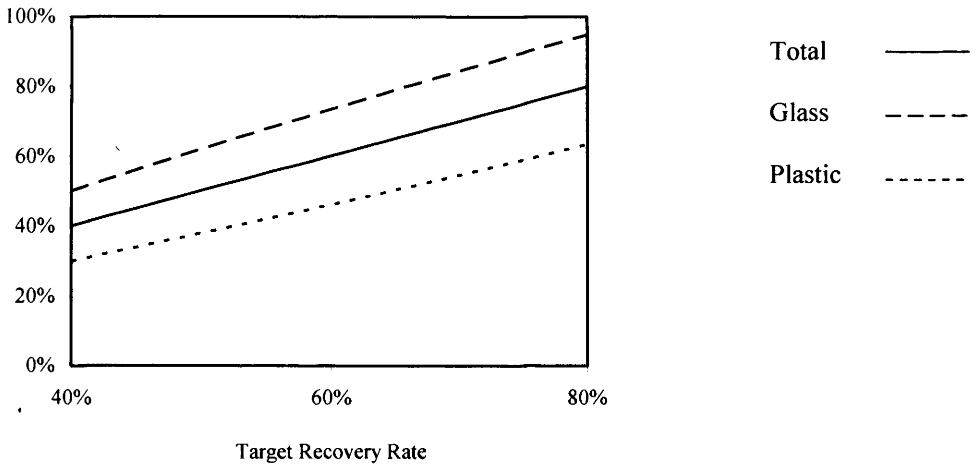
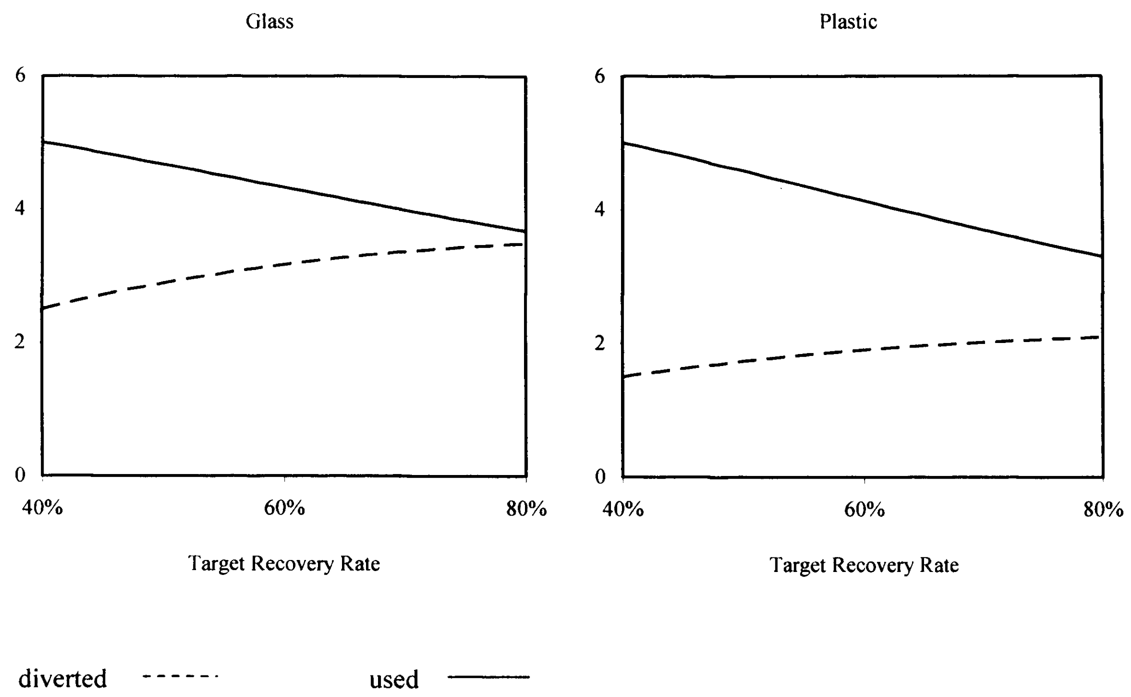


Figure 6.7 shows what this means in terms of the quantities of each packaging material used and diverted, with the difference representing the amount going to landfill. The general picture is the same for both materials – a decline in the quantity produced and an increase in the quantity diverted, and hence a reduction in the quantity ending up in landfill. However, the relative contribution of the two responses differs between the two materials. For example, at an 80% aggregate recovery rate, packaging reduction accounts for 55% of the increase in the diversion rate for plastic, but only 48% for glass.[26]

Figure 6.7 Quantity of packaging used and diverted (M. tonnes)

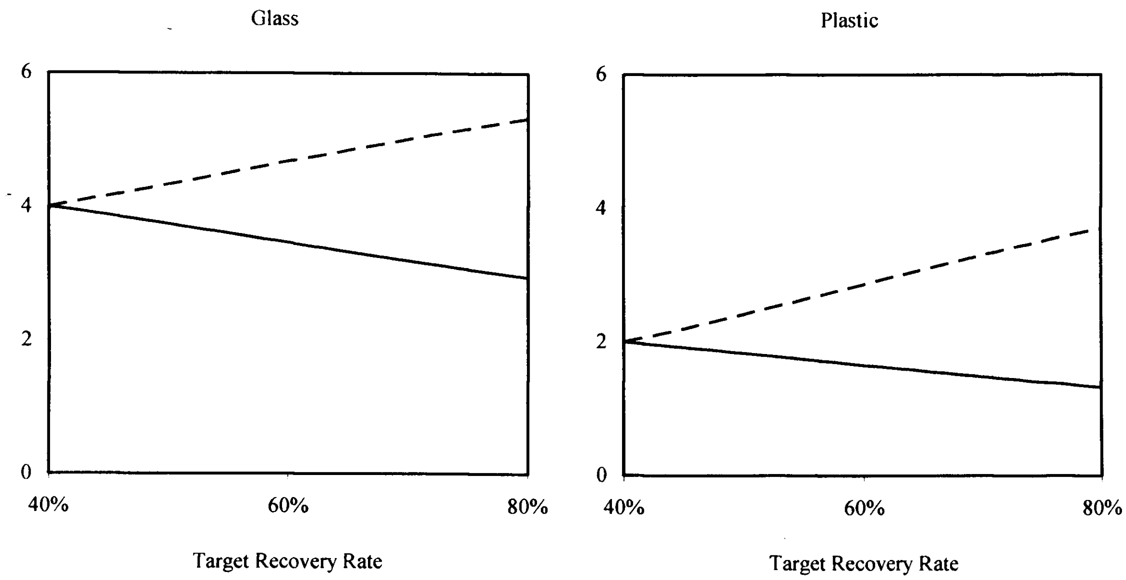


The impact on the prices of the two packaging goods – bottles and punnets – depends on the assignment of the obligation to purchase performance credits (i.e. the value of θ_1). As can be seen in Figure 6.8, if the obligation is placed entirely on the packaging producers (i.e. $\theta_1 = 1$), then the market price of packaging increases as the target

[26] The contribution of packaging reduction is calculated as $[\ln(Y_1^u) - \ln(Y_1^\#)] / [\ln(r) - \ln(r^u)]$, and the contribution of increased diversion as $[\ln(Y_8^\#) - \ln(Y_8^u)] / [\ln(r) - \ln(r^u)]$.

recovery rate rises. In contrast, if the obligation is placed entirely on the packer-fillers (i.e. $\theta_1 = 0$), the price falls.

Figure 6.8 Price of packaging (€)

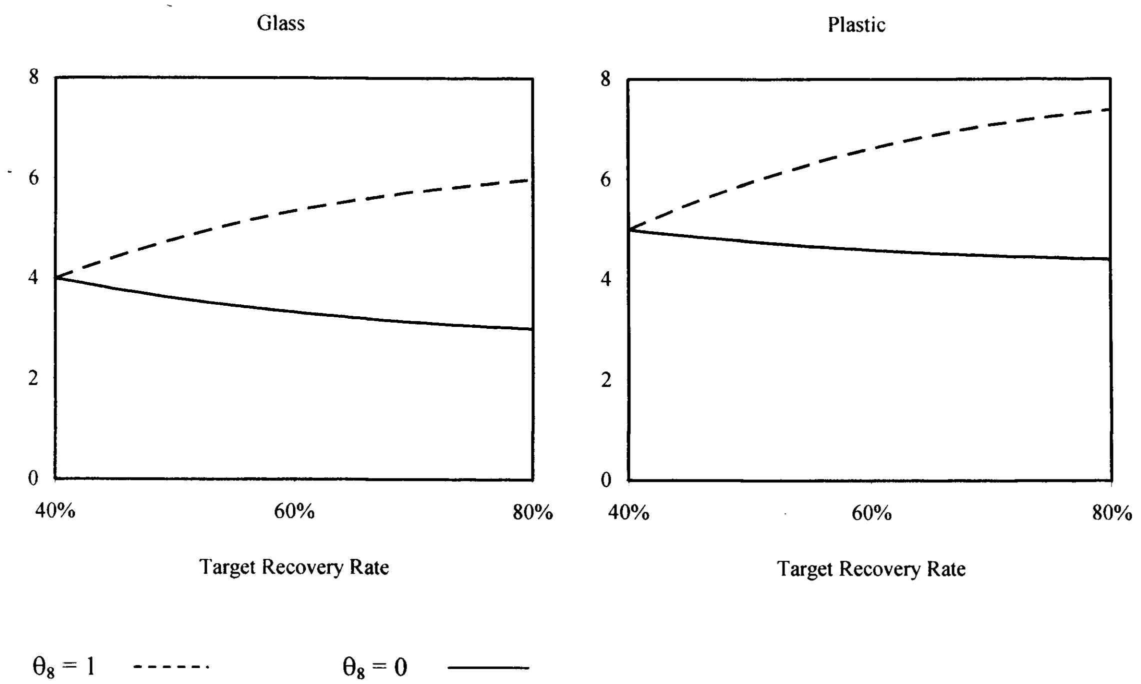


$\theta_1 = 1$ ------ $\theta_1 = 0$ ———

For both materials, the difference between the two price trajectories is equal to the price of performance credits multiplied by the target recovery rate; with the higher trajectory representing the effective price of packaging to the users (i.e. the net cost of packaging, taking account of the additional performance credit that must be purchased). Thus, the net cost increase is greater for plastic punnets (+85% versus +33%), which is consistent with the greater contribution of packaging reduction to the increase in the diversion rate for plastic.

A similar picture can be seen in Figure 6.9, which shows the impact on the price of diverted waste packaging. When the rights to the performance credits are assigned entirely to the material reprocessors (i.e. $\theta_8 = 1$), the price of diverted waste packaging increases as the target diversion rate rises. In contrast, when the rights are assigned entirely to the waste managers (i.e. $\theta_8 = 0$), the price declines.

Figure 6.9 Price of diverted waste packaging (€)



Again, since the difference between the two price trajectories is equal to the price of performance credits at the prevailing target recovery rate, it is the lower price trajectory which represents the effective price of diverted waste packaging to the reprocessors (i.e. taking into account the value of the additional performance credit that can be sold). Consequently the net cost decrease is greatest for glass (-25% versus -12%), which again is consistent with the greater contribution of waste diversion to the increase in the diversion rate for glass.

Table 6.3 shows the aggregate costs of meeting a range of different recovery targets (i.e. reduction in the aggregate surplus versus the pre-regulation level). As one would expect, the cost rises as the target recovery rate increases. Also shown are the estimated costs that have been calculated using the simple formula identified in section 6.3. Because all of the underlying WTP and MC functions are linear in this example, the estimated values are exactly equal to the actual costs.

Table 6.3 Aggregate cost of achieving diversion target

Target diversion rate	Actual cost (€ M.)	V^{MIN} (M. Tonnes)	$V^{MAX}(r)$ (M. Tonnes)	q (€ / Tonne)	Estimated cost (€ M.)
40%	0.00	4	4	0.00	0.00
50%	0.59	4	5	1.18	0.59
60%	2.02	4	6	2.02	2.02
70%	3.89	4	7	2.60	3.89
80%	5.94	4	8	2.97	5.94

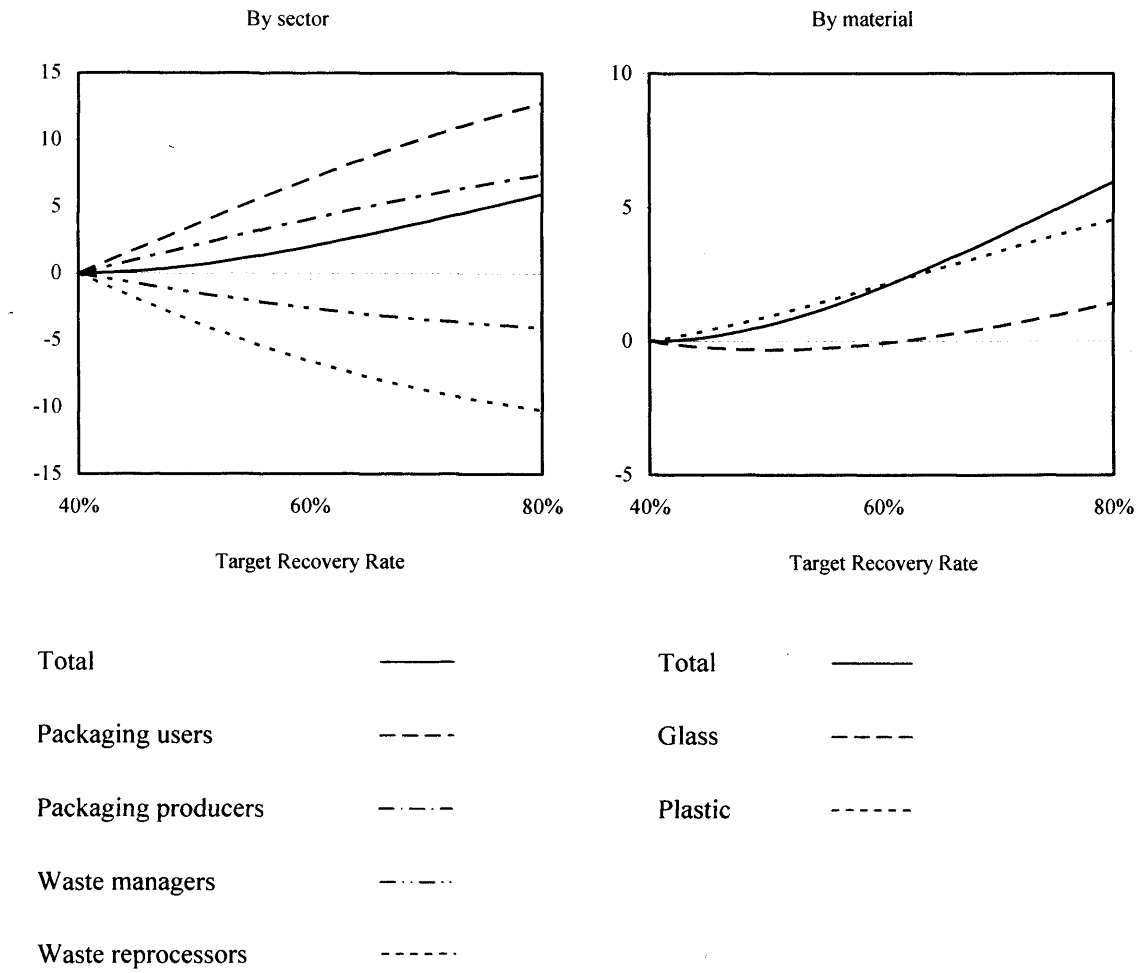
NOTE: $V^{MIN} = Y_8^u = 4$

$V^{MAX}(r) = Y_1^u \times r = 10 \times r$

Figure 6.10 shows the distribution of the aggregate cost between the four sectors, and between the two materials, when the performance adjustment factors are all set equal to zero. In each case, the solid line represents the aggregate cost, while the broken lines represent the costs incurred by the different sectors / materials, or the benefits accruing to them (i.e. negative costs).

As can be seen, the distribution of the aggregate cost is highly inequable. For example, at a 60% recovery rate the spread of the costs between sectors is seven times greater than the aggregate cost. This is driven by two factors. First, as has already been noted (in the previous section), when the performance adjustment factors are all set equal to zero, the financial transfer received by the downstream sector necessarily exceeds the additional costs that it incurs as a result of the regulatory intervention. In this example, the scale of this overcompensation is significant, and it increases as the target recovery rate becomes more stringent.

Figure 6.10 Distribution of aggregate cost (€ M.)



Second, the distribution of the net benefit within the downstream sector is unequal, as is the distribution of the net cost within the upstream sector; with the waste reprocessors receiving the majority of the benefit, and the packaging users shouldering the majority of the cost burden. However, this merely reflects the fact that under the chosen functional forms, the slopes of the WTP functions are steeper than those of the MC functions in all of the markets. If the situation was to be reversed, then the packaging producers would bear the greatest cost, and the waste managers would receive the greatest benefit. The important point to note however, is that the distribution of the total cost is completely unaffected by the choice of values for the assignment parameters θ_1 and θ_8.

A similar picture can be seen in relation to the distribution of the aggregate cost between materials, with the reduction in the net surplus associated with glass packaging being significantly lower than that associated with plastic packaging.[27] Indeed for target recovery rates below around 60%, there is an increase in the net surplus associated with glass. This cross-subsidization between materials arises because of the recovery target applies to the aggregate quantity of packaging used across all materials. If each material was required to meet the target individually, then it would not occur.

In order to get a better understanding of what lies behind these impacts, Table 6.4 provides a breakdown of the net cost into its two components – the gross cost (i.e. the reduction in operating profit / consumer surplus), and the financial transfer resulting from the trade of performance credits.

Two points are particularly striking in this breakdown. First, the scale of the financial transfers is significantly greater than the gross costs. For example, the total value of the transfer (at around € 5.4 M.) is nine times greater than the total gross cost, while the upstream sectors' cost of acquiring performance credits in relation to glass is almost thirty times greater than the reduction in their combined operating profits! Second, the distribution of the gross cost is much more even (although by no means equal) – both between upstream and downstream sectors (37% versus 63%), and between glass and plastic (55% versus 45%).

[27] Under the assumptions that have been made regarding the separability of production and utility functions by associated packaging material, the total surplus for the system is equal to the sum of the total net surpluses associated with each material, where the net surplus is equal to the gross surplus plus / minus the monetary value of performance credits sold / bought.

Table 6.4 Operating cost and financial transfers for 50% diversion target (€ M.)

		Operating cost	Trading transfer	Net cost
Glass	Upstream	0.096	2.747	2.843
	Downstream	0.230	(3.399)	(3.169)
		0.326	(0.652)	(0.326)
Plastic	Upstream	0.123	2.692	2.815
	Downstream	0.138	(2.039)	(1.901)
		0.262	0.652	0.914
Total		0.588	0	0.588

It is clear therefore that the huge distributional impacts are caused by the scale of the financial transfers resulting from the trade of performance credits, which goes far beyond the value required to compensate the downstream sector for the additional gross costs that it incurs. While this has no bearing on the cost efficiency of the trading mechanism, such an inequitable outcome is likely to undermine its political acceptability. However, as has been noted in previous chapters, it is possible to alter the distributional impact of the scheme by setting non-zero values for the performance adjustment factors in the agents' individual performance rules.

In theory, any distributional outcome can be achieved by setting appropriate values for the performance adjustment factors. In particular, as can be seen in Table 6.5, it is possible to achieve an outcome in which there is no over-compensation of downstream sectors (i.e. the net cost of downstream sectors is zero), and there is no cross-subsidy between materials.

Table 6.5 Operating cost and financial transfers for 50% diversion target, with performance adjustment factors (€ M.)

		Operating cost	Trading transfer	Net cost	PAF (M. tonnes)
Glass	Upstream	0.096	0.230	0.326	(2.141)
	Downstream	0.230	(0.230)	0	2.696
		0.326	0	0.326	0.555
Plastic	Upstream	0.123	0.138	0.261	(2.173)
	Downstream	0.138	(0.138)	0	1.618
		0.261	0	0.261	(0.555)
Total		0.588	0	0.588	0

NOTE: q = € 1.175

Of course, in order to achieve such an outcome in practice, the regulatory authority would need to have some idea of the scale and distribution of the gross cost between sectors and materials. However, this information can be derived – at least approximately – from the market for performance credits, noting that the expression for the upper bound of the aggregate cost can be decomposed into:

$$\left[V^{max} - V^{min} \right] \times \frac{q}{2} = \sum_{m \in M} \left[V^{d}_{m} - V^{min}_{m} \right] \times \frac{q}{2} + \sum_{m \in M} \left[V^{max}_{m} - V^{u}_{m} \right] \times \frac{q}{2}$$

where $$\sum_{m \in M} V^{d}_{m} = V^{\#} = \sum_{m \in M} V^{u}_{m}$$

The components of the two summations on the right-hand side represent the operating costs of the downstream and upstream sectors respectively, for each packaging material. Thus, the only information needed to estimate the operating costs are the pre-regulation quantities of packaging used and diverted for each material (which are used to calculate the values of V^{max} and V^{min} respectively); the numbers of performance credits that are bought and sold for each material; and the market price of credits. This is illustrated in Figure 6.6, which shows the calculations for a 50% diversion target. The estimated breakdown of operating costs is exactly the same as the actual breakdown in Figure 6.4. However, as with the total cost estimate, this reflects the use of linear WTP and MC functions in this particular example.

Table 6.6 Estimated distribution of total operating cost for 50% diversion target

		V^{max}	V^{min} (M. tonnes)	V	Estimated op. cost (€ M.)
Upstream	Glass	2.500	-	2.337	0.096
	Plastic	2.500	-	2.290	0.123
		5.000	-	4.627	0.219
Downstream	Glass	-	2.500	2.892	0.230
	Plastic	-	1.500	1.735	0.138
		-	4.000	4.627	0.368
Total					0.588

NOTE: q = € 1.175

6.5 Summary

In this chapter a stylised model of the packaging system has been utilized to explore the use of performance-based credit trading to implement a diversion (or recovery) target for waste packaging – expressed as a percentage of the total amount used. While the model encapsulates the salient features of the packaging system in the United Kingdom, it is not intended to provide a detailed and faithful representation of the actual system. However, it is sufficiently rich to investigate the impacts – both real and financial – of the mechanism on different parts of the system, and to demonstrate how information provided by the market for performance credits can be used to assess the magnitude and distribution of the aggregate cost of the intervention.

The first point to note is that performance-based credit trading can provide a cost-efficient implementation mechanism for an aggregate waste diversion target under extended producer responsibility. However, in order for it to do so, it is necessary that households face the full marginal cost of waste collection – either through direct charging for the service, or through the imposition of an advance disposal charge on producers.[28] If this is not the case, then the amount of packaging used will be greater than the cost efficient level, unless the marginal cost of landfill (and hence the cost of waste collection services) is equal to zero.

Under the trading scheme, the total quantity of each packaging material that is used declines from its pre-regulation level, to a point where the marginal WTP of the packaging users exceeds the MC of production by an amount equal to the market price of performance credits multiplied by the target recovery rate. In contrast, the total quantity of each packaging material that is diverted increases, to a point where the marginal cost of diversion exceeds the marginal WTP of the waste reprocessors by an amount equal to the market price of performance credits. As the target diversion rate increases, the quantity of packaging used declines, while the amount that is diverted rises. Consequently, the quantity of waste packaging sent to landfill declines. While the contributions of source reduction and increased diversion to the reduction in landfill will depend on the relative costs of the two responses, the general trends are the same for all packaging materials. Thus, for all materials the percentage increase in the quantity of waste packaging diverted is less than the percentage increase in the

[28] The requirement that households be charged for waste collection is only necessary for the cost efficiency of the trading scheme. All of the other conclusions regarding the properties of the trading scheme remain valid in the absence of household charging.

diversion target. For those materials with particularly high costs of diversion, and hence significant source reduction, the difference may be substantial.[29]

For any given target diversion rate, the same outcome could have been achieved by imposing a uniform tax on all packaging goods, and providing a uniform subsidy payment for all diverted waste packaging. Such a scheme would be revenue-neutral. Indeed, in this application, performance-based credit trading can be viewed as a "private sector subsidy scheme", with the subsidy being paid by the upstream sector rather than by the government.

The impact of an increase in the target diversion rate on the actual market prices of packaging and diverted waste packaging depends on the values of the respective assignment parameters. Prices could rise, decline, or stay the same. However, the impacts on the effective prices of packaging to the users, and on the effective prices of diverted waste packaging to the reprocessors, are unaffected by the choice of parameter values; with the former increasing, and the latter declining. The impact on the market price of performance credits depends on the shapes of the underlying demand and supply functions, and it is possible that the price could decline as the target rate is increased.

If households face the full marginal cost of waste collection, then the aggregate cost of the regulatory intervention under the trading scheme is equal to the sum of the deadweight losses across all of the markets for packaging and diverted waste packaging. However, if households are not charged the full cost of waste collection, then the sum of the deadweight losses only provides an upper bound on the aggregate cost – overstating it by an amount equal to the reduction in the value of the implicit subsidy. The sum of the deadweight losses is also equal to the area under the lower envelope of the aggregate

[29] It is straightforward to show that the percentage increase in the quantity of diverted waste packaging is equal to the percentage increase in the diversion rate, less the percentage decrease in the quantity used.

inverse demand and supply curves for performance credits. Thus, it is only necessary to consider this single market in order to determine (the upper bound of) the aggregate cost, rather than all of the individual markets for packaging and diverted waste packaging – of which there may be many. Furthermore, this area can be easily approximated using information about the pre-regulation quantities of packaging used and diverted, and the market price of performance credits.

An important point to note is that the total value of the financial transfer resulting from the performance credit transactions is not – in general – equal to the aggregate cost of the intervention under the trading scheme. Unfortunately, it is not possible to draw any unambiguous conclusion regarding the relative magnitudes of the financial transfer and the aggregate cost. If the reduction in packaging use by the upstream sectors is relatively low, then the value of the financial transfer will be greater than the aggregate cost. However, if there is a significant reduction in packaging use, it is possible that the value of the transfer could understate the aggregate cost.

When the performance adjustment factors are set to zero, the total revenue received by the downstream sectors is unambiguously greater than the gross cost that they incur under the regulatory intervention (i.e. the sum of the deadweight losses in the markets for diverted waste packaging). That is, the trading mechanism over-compensates the downstream sectors. Furthermore, unless all materials have the same diversion rate in the market equilibrium, the trading of performance credits will result in a financial transfer from materials with low equilibrium diversion rates (for which the demand for credits will exceed the supply) to materials with high diversion rates (which have excess supply).

The scale of the over-compensation of the downstream sectors can be significant, as can the scale of the cross-subsidization between materials. Indeed, it is possible that the financial transfer arising from the trading of performance credits may dwarf the aggregate cost of the intervention. However, both the over-compensation and the cross-subsidization can be eliminated by setting appropriate values for the performance adjustment factors. In order to do this, the authorities need to know the scale of the aggregate cost, and its distribution between sectors and materials. Prior to the implementation of the trading scheme, the calculations must be based on forecasts – over which there may be little consensus. However, once the scheme is operational, all of the information that is required to determine the values of the adjustment factors is readily available from the market for performance credits.

Part Three

Performance-based credit trading and market power

Chapter 7 Strategic interaction and market-based mechanisms

In Part Two of the thesis, the properties of performance-based credit trading were explored under the assumption of perfect competition in all markets – i.e. that all of the agents in the production system act as price-takers. That assumption is relaxed in this third part of the thesis, and the implication of *strategic interaction* between some – or all – of the agents is investigated. Of particular interest are those forms of interaction that lead to an agent being able to exercise *market power*. That is, the ability to alter prices away from competitive levels, to its own benefit.[1]

There are many ways in which market power could be introduced into the model. Different assumptions can be made about the nature of the market power (e.g. monopoly / monopsony price-setting, dominant firm price leadership, Cournot quantity oligopoly, Stackelberg quantity duopoly, etc.), and about the market – or markets – that are affected. The potential combinations are numerous, and each case must be considered individually in order to determine how it might affect the outcome. One

[1] When there is strategic interaction between agents, each agent recognises that the benefit it receives depends not only on its own actions, but also on the actions of other agents. These may be actions that the other agents have already taken; actions that they are expected to take at the same time; or actions that they may take (or not take) as a result of the agent's current action. Strategic interaction between agents does not necessarily lead to any of those agents being able to exercise market power. For example, in the Bertrand oligopoly model, the outcome of the interaction is that all firms set their prices equal to the competitive level.

particular combination is analysed in Chapter 8, which revisits the application of performance-based credit trading to packaging recovery targets considered in the previous chapter. However, before moving on to this analysis, it is useful to review the previous work that has been undertaken on the topic of strategic interaction and market-based implementation mechanisms.

A number of authors have considered how market power can affect the outcome under a market-based implementation mechanism. In all but one of these analyses, the authors have assumed that the regulatory target takes the form of an absolute limit for aggregate emissions of some pollutant, and that the mechanism is a "cap and trade" permit trading system. These analyses cover a number of different forms of strategic interaction, with varying degrees of market power. Essentially however, they can be divided into two broad groups: those that focus on the implications of market power in the permit market for the outcome in that market; and those that focus on the implications of using tradable emission permits for the outcome in the output market, when the latter is imperfectly competitive. These groups are considered in turn in the next two sections.

Only Mauleg (1990) has considered the issue of market power in the context of a relative performance standard – assessing the implications of using a performance-based credit trading scheme to implement a target rate for emissions per unit output when the product market is a homogeneous Cournot oligopoly. Since, this is clearly more relevant to the topic of this thesis, a more detailed review of his analysis is provided in the third section of the chapter.

The final section pulls together the findings of the various analyses in an attempt to draw some general conclusions regarding the implications of market power for the impacts of market-based implementation mechanisms in general, and of performance-based credit trading in particular.

7.1 Implications for the permit market

Table 7.1 identifies the analyses that have focused on the implications of market power in the permit market for the outcome in that market. In all four cases, the regulatory target takes the form of an absolute limit for aggregate emissions. However, the analyses differ in the assumptions that they make about the nature of market power in the permit market, and about the competitive structure of the output market.

Table 7.1 Analyses focussing on the permit market

<table>
<tr><th>Reference</th><th>Regulatory target</th><th>Permit market</th><th>Output market</th></tr>
<tr><td>Hahn (1984)</td><td rowspan="4">Absolute</td><td rowspan="3">Monopoly / monopsony</td><td rowspan="2">Perfect competition</td></tr>
<tr><td>Hagem and Westskog (1998)</td></tr>
<tr><td>Misiolek and Elder (1989)</td><td>Dominant firm</td></tr>
<tr><td>von der Fehr (1993)</td><td>Duopoly</td><td>Duopoly</td></tr>
</table>

Hahn (1984) considers the simplest case of a static model, in which one firm has the power to set the price in the permit market – either as a seller (monopolist) or as a buyer (monopsonist), while the output market is perfectly competitive.[2,3] He demonstrates that unless the number of permits allocated to the price-setter is equal to the quantity that it holds in the market equilibrium (i.e. it does not buy or sell any permits), then the price that it sets will deviate from its marginal abatement cost. If it sells permits in the equilibrium, then it sets a permit price that is higher than its marginal cost of abatement.

[2] Hahn does not speculate how such a situation might arise. However, a sufficient condition would be that the initial allocation of permits equalises the marginal abatement costs of all but one of the firms. Then, by default, the remaining firm will either be the only buyer of permits, or the only seller, depending on the value of its marginal abatement cost.

[3] Misiolek & Elder (1989) repeat Hahn's analysis in diagrammatic form, using the same underlying assumptions and notation, as they explicitly acknowledge.

If it buys permits, then it sets a price that is lower. In either case, this means that marginal abatement costs are not equalised across all firms, and hence that the total cost of achieving the aggregate limit exceeds the minimum level. Put another way, a necessary condition for cost efficiency is that the initial permit allocation to the price-setter is efficient.

Figure 7.1 Permit market equilibrium with a single price-setter

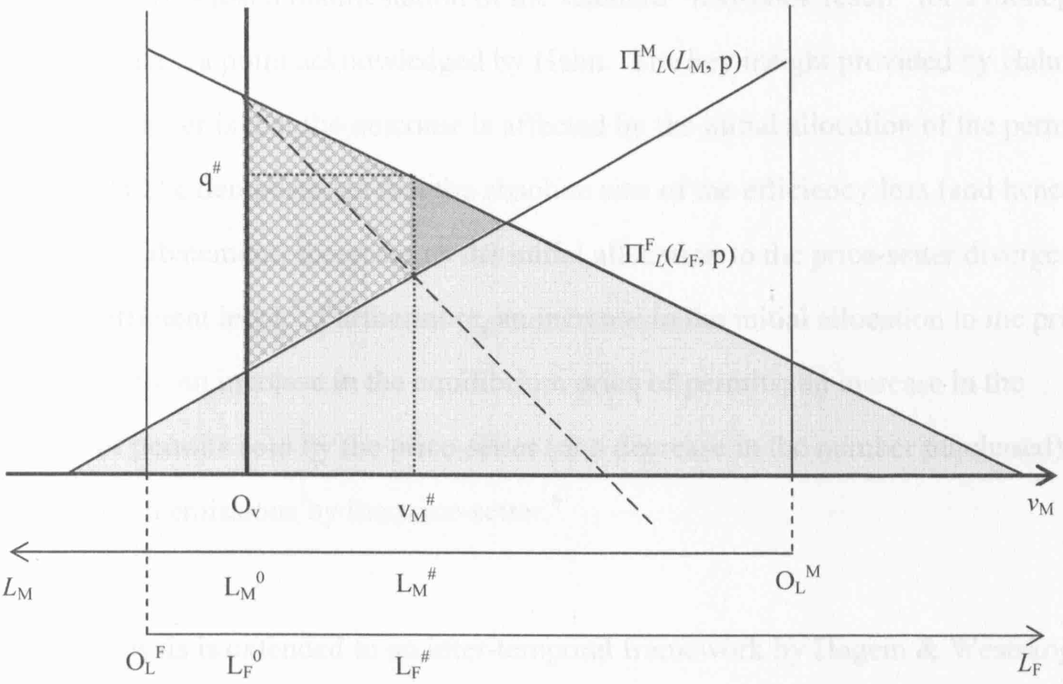


Figure 7.1 illustrates the impact on the market for emission permits for the case where the price-setter is a seller. There are three sets of axes. The first two (with origins O_v and $O_L{}^M$) relate respectively to the price-setter's sales of permits and to its permit holding, while the third (with origin $O_L{}^F$) relates to the aggregate permit holding of the price-taking firms – known collectively as the Fringe.[4] The functions $\Pi^M{}_L(\cdot)$ and $\Pi^F{}_L(\cdot)$

[4] The axes relating to the price-setters permit holding have been "flipped" so that positive values are to the left of the origin

are the marginal abatement costs of the price-setter and the Fringe respectively. For the given initial allocations (L_M^0, L_F^0), these define the inverse supply and demand functions for permit transfers (origin O_v), with the dashed line representing the price-setter's marginal revenue. The light grey shaded area denotes the minimum total abatement cost, which would be achieved if the permit market was perfectly competitive. The gains from trading are given by the checked area; while the dark grey area represents the efficiency loss (or deadweight loss) arising from the market power.

Of course, this is just a manifestation of the standard "text-book result" for a monopolist / monopsonist – a point acknowledged by Hahn. The key insight provided by Hahn's analysis however is that the outcome is affected by the initial allocation of the permits. In particular, he demonstrates that the absolute size of the efficiency loss (and hence the total cost of abatement) increases as the initial allocation to the price-setter diverges from the efficient level.[5] Furthermore, an increase in the initial allocation to the price-setter leads to: an increase in the equilibrium price of permits; an increase in the quantity of permits sold by the price-setter (or a decrease in the number purchased); and an increase in emissions by the price-setter.[6]

Hahn's analysis is extended to an inter-temporal framework by Hagem & Westskog (1998), under the assumption that an absolute limit is set for aggregate emissions over a time horizon comprising a fixed number of periods. In this case, the price-setter chooses a vector of permit prices – one for each period. However, assuming (as Hagem & Westskog do) that there are no restrictions on the banking and borrowing of permits,

---

[5] Implicit in Hahn's analysis is an assumption that the price in the output market (p) is unaffected by the strategic behaviour in the permit market, and hence can be treated as a fixed parameter in the derivation of the comparative statics results.

[6] All of these impacts can easily be confirmed by reference to Figure 7.1. An increase in the price-setter's initial allocation causes the permit transfer origin (Ov) to shift to the left. Consequently, so too does the intersection of the price-setter's marginal revenue and marginal abatement cost curves; with all of the identified impacts following directly.

each firm's discounted abatement cost is determined by its total permit allocation and net transfers over the time horizon, and is completely unaffected by their timing. Furthermore, since all firms are assumed to have perfect foresight about future permit prices, it must be the case that the price increases at the rate of interest (i.e. the discounted permit price is constant). Otherwise there would be opportunities for inter-temporal arbitrage. Consequently, the firms' permit-transfer choices can be decoupled from their permit-use and output choices, and one can assume (for analytic purposes) that all permit allocations and transfers take place at a single point in time.

Hagem & Westskog demonstrate that all firms – including the price-setter – equalize their discounted marginal abatement costs across time periods. However, as in the single-period case, unless net permit transfers are equal to zero, the marginal abatement cost of the price-setter diverges from that of the price-takers. Consequently, while the trading mechanism achieves inter-temporal efficiency within firms, it does not usually achieve intra-temporal efficiency between firms. Although it is not acknowledged by the authors, the de-coupling of the permit market means that all of Hahn's conclusions relating to the impacts of changing the initial permit allocations (between firms) carry over to the inter-temporal setting.

Returning to a single-period framework, Misiolek & Elder (1989) consider the situation in which the price-setting firm in the permit market is also dominant in the output market, and can set the price in that market as well. In this case, the firm can affect the outcome in the output market both directly, by setting the output price; and indirectly, by affecting the production cost of the Fringe through the price that it sets for emission permits. The use of the permit price in this way to raise the production cost of the Fringe is an example of so-called "exclusionary manipulation" (EM), first analysed by Salop & Scheffman (1983; 1987).

In Salop & Scheffman (1987), the dominant firm (or predator) in the output market can also set the price of a factor input; adjusting its purchases to clear the market.[7] They demonstrate that the predator will set a price for the input such that the marginal factor cost exceeds the marginal revenue product. In this sense, the predator "over-purchases" the factor relative to the amount that it would use as a "simple monopolist-monopsonist" (SMM).[8] While they recognize that it is not possible to draw any general conclusions about the relationship between the two equilibria, it is likely to be the case that the predator purchases more of the factor when it engages in exclusionary manipulation.[9] However, since its purchases are too low in the SMM case (relative to the competitive level), the exclusionary manipulation may have the effect of improving overall efficiency.

Misiolek & Elder apply the results of this analysis to the case of emission permits. While they set out the optimization problem for the predator, they do not attempt to derive the conditions for the solution. Rather, they take the "over-buying" result derived by Salop & Scheffman as given, and use a graphical framework to assess the impact of the exclusionary manipulation on the permit market relative to two benchmarks: a) the SMM outcome; b) the outcome when the permit market (but not the output) market is perfectly competitive. In doing so, they implicitly assume that the price chosen by the predator in the output market is unaffected by the outcome in the permit market (i.e. is the same in all three equilibria). Hence, they can compare the respective equilibria directly using static (inverse) supply and demand curves. While

[7] They assume that the factor is supplied by a competitive industry, and that it has an upward-sloping supply curve.

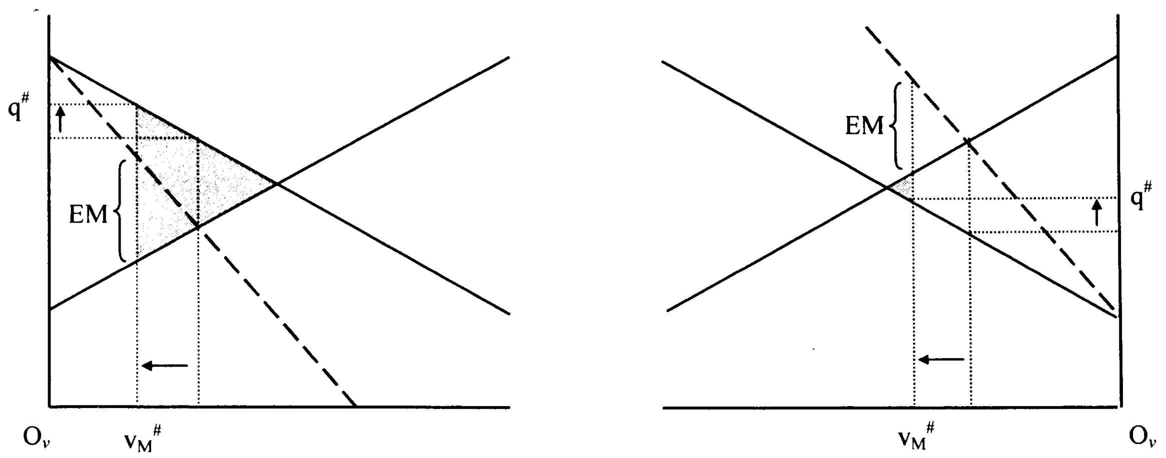
[8] In the "simple monopoly-monopsony" (SMM) case, the firm acts as a dominant seller in the output market, and as a dominant buyer in the factor market, but ignores the interaction between the two markets. That is, it ignores the effect of a change in the factor price on the residual demand for its output, and the effect of a change in the output price on the residual supply of the factor.

[9] Unless the output price chosen by the predator is unaffected by the exclusion manipulation, both the marginal factor cost curve and the marginal revenue product curve will shift. Hence the intersection of the EM curves does not generally represent the equilibrium in the SMM case.

this is not an impossibility, it is not generally the case.[10] Consequently, their conclusions should be taken as indicative of the likely impacts allowing for shifts in the curves as the output price changes.

Figure 7.2 Impact of exclusionary manipulation on permit market

a) Predator sells permits **b) Predator buys permits**



The impact of exclusionary manipulation on the equilibrium in the permit market is illustrated in Figure 7.2, where – in keeping with Misiolek & Elder – it is assumed that the product price chosen by the price-setter is unaffected by the outcome in this market. If the predator is a seller of permits in the SMM case, then exclusionary manipulation causes it to further reduce the number of permits that it sells, so that the marginal revenue from an additional permit sale exceeds its marginal abatement cost (i.e. it "under-sells"). Consequently, it holds more permits, has higher emissions and undertakes less abatement. Since, the price-setter's sales are too low (relative to the competitive level) in the SMM case, the exclusionary manipulation exacerbates the

[10] For example, if the predator's conditional cost function is $C^P(Y_P, L_P) = \alpha\, Y_P^2 + \beta\, Y_P\, L_P + \chi\, L_P^2$; product demand is $Y(p) = a - b\,p$; and the Fringe's supply is $Y^F(p, L_P) = d\,(L - L_P) + e\,p$, where L is the total number of permits; then the output price is independent of the value of L_P if $\beta = -\,d\,[\,\alpha + 1/(b + e)\,]$.

efficiency loss. The excess of the price-setter's marginal revenue over its marginal abatement cost (EM) represents the loss of profit that it incurs in the product market (at the equilibrium product price) if the Fringe acquires an additional permit. If this is sufficiently large, then it is possible for the predator to switch from being a seller of permits (under SMM) to being a buyer.

If the predator is a buyer of permits in the SMM case, then the impact is the same as for any other factor input. The predator buys more permits than it would if it did not engage in exclusionary manipulation, and hence undertakes less abatement and has higher emissions. However, as noted by Salop & Scheffman, since the dominant firm purchases too few permits under SMM (relative to the competitive level), this can lead to an improvement in abatement efficiency – even to the attainment of the optimal allocation of abatement effort. In this case, the excess of the marginal factor cost over the price-setter's marginal abatement cost (EM) represents the increase in its profit in the product market if the Fringe sells an additional permit.

Finally in this group of analyses, von der Fehr (1993) considers the case where two firms that compete as Cournot duopolists in the product market can also influence the price of emission permits through their respective purchase quantity choices.[11] He investigates a number of different structures for the interaction between the two firms; which in each case is modelled as a two-stage game, with the solution being given by the sub-game perfect equilibrium.

In the first case, the firms simultaneously choose their permit purchases in the first stage; before choosing their output levels in the second stage. Von der Fehr shows that if quantities are strategic substitutes in the output market, then each firm will "over-purchase" emission permits – again in the sense that the marginal factor cost exceeds

[11] He assumes that the firms are both buyers of emission permits. These are supplied by other regulated firms that are all price-takers with upward sloping marginal abatement cost curves.

the firm's marginal abatement cost.[12] The intuition for this behaviour is that by increasing its purchases of permits, a firm reduces its marginal cost of production, and hence increases its output. The resultant reduction in the market price causes its rival to reduce its output, to the benefit of the firm. Thus, each firm has an additional incentive to purchase permits – over and above the avoided abatement cost. This is an example of a so-called "Top Dog" strategy, in which firms competing in strategic substitutes over-invest in strategic variables that make them more aggressive.[13] While he does not consider the impact on the output market explicitly, von der Fehr notes that since this strategic effect tends to increase the firms outputs there is an amelioration of the efficiency loss in this market.

This result is driven by the impact of the firms' permit purchases on their own marginal production costs. In the first stage, each firm takes the other's permit purchase quantity as given, and hence does not take account of any impact that its decision may have on its rival's marginal cost. In order to capture this impact, von der Fehr considers an alternative game structure, in which Firm 1 acts as a Stackelberg leader in the permit market. In this case, Firm 1 (the "leader") chooses its permit purchase quantity in the first stage; with Firm 2 (the "follower") making its choice in the second stage, at the same time as both firms make their output choices.

Von der Fehr demonstrates that if output quantities are strategic substitutes, and the marginal factor cost of permits for the follower is increasing in purchases by leader, then the leader will over-purchase emission permits relative to the benchmark.[14] In this

[12] Von der Fehr compares the outcome against the "non-strategic" benchmark in which each firm's marginal permit expenditure equals its marginal abatement cost. This would occur if the firms were to make their output and permit decisions simultaneously (i.e. in a one-stage game).

[13] See Tirole (1988) – chapter 8.3.

[14] The benchmark differs from the previous case in that the leader's marginal permit expenditure takes account of the change in permit price resulting from its own purchases and the induced change in the follower's purchases.

case there are two reasons for the leader to "over-invest" in emission permits. In addition to the "Top Dog" effect, by increasing its purchases the leader forces up the price of permits. This leads the follower to reduce the quantity of permits that it purchases; causing its marginal cost curve to shift upwards, and reducing its output. The resultant reduction in the aggression of its rival in the product market benefits the leader, giving it an additional incentive to purchase permits. While not acknowledged as such by von der Fehr, this second effect is clearly equivalent to the exclusionary manipulation of the permit price analysed by Misiolek & Elder (1989).

Finally, von der Fehr considers how the strategic potential of emission permits can affect the choice of abatement technologies. In this case, the leader chooses how much to invest in abatement equipment in the first stage; with both firms simultaneously choosing their output levels and permit purchases in the second stage. He shows that if output quantities are strategic substitutes and if expenditure on abatement equipment reduces the marginal cost of abatement but increases the marginal cost of production, then the leader will under-invest in abatement equipment relative to the non-strategic benchmark (i.e. in which the marginal cost of investment equals the marginal benefit).

The intuition for this behaviour is similar to that underlying the first case – only this time the leader benefits from the response of its rival in the permit market as well as the output market. By reducing its expenditure on abatement equipment, the leader shifts its marginal cost of production downwards, and its (inverse) demand for emission permits upwards. The resultant increases in the leader's output and permit purchases cause the product price to fall and the permit price to rise. Consequently, the follower reduces both its output quantity and the number of permits that it acquires – with both responses benefiting the leader.

7.2 Implications for an imperfectly competitive output market

Six authors have considered the implications of using a market-based mechanism to implement a given regulatory target for the outcome in the product market, when this market is subject to strategic interaction (see Table 7.2). In particular, they consider the impacts on output levels, abatement costs, profits, and – in some cases – gross welfare; with most of the analyses using the outcome under an equivalent consent-based mechanism as the benchmark for the comparison. Apart from Mauleg (1990), whose analysis is reviewed in the next section, all of the analyses assume an absolute limit for aggregate emissions, implemented by tradable emission permits.

Table 7.2 Analyses focussing on the product market

<table>
<tr><th>Reference</th><th>Regulatory target</th><th>Permit / credit market</th><th>Output market</th></tr>
<tr><td>Mauleg (1990)</td><td>Relative</td><td rowspan="2">Perfect competition</td><td>Oligopoly</td></tr>
<tr><td>Sartzetakis (1997a)</td><td rowspan="5">Absolute</td><td rowspan="5">Duopoly</td></tr>
<tr><td>Sartzetakis (1997b)</td><td>Monopoly</td></tr>
<tr><td>von der Fehr (1993)</td><td rowspan="3">Bargaining</td></tr>
<tr><td>Requate (1993)</td></tr>
<tr><td>Fershtman & de Zeuw (1995)</td></tr>
</table>

The analyses are less general than the previous group in that they assume specific functional forms for the firms' cost functions and for demand in the output market. While they all assume that the output market is a homogeneous Cournot duopoly with linear (inverse) demand, the analyses differ in the assumptions that they make about the firms' cost functions (and hence their underlying production and abatement technologies). More importantly, there are also differences in the assumptions that are made regarding the competitive structure of the permit market; reflecting different underlying assumptions about the prevalence of the pollutant, and hence the number of participants in that market.

Sartzetakis (1997a) assumes that the pollutant is emitted by several industries, which have similar intra-industry abatement cost distributions. Consequently, while the market for emission permits comprises many firms, all of which – including the two duopolists – act as price-takers, the equilibrium permit price does not result in a net transfer of permits between industries (i.e. there is intra-industry market clearing). Furthermore, he assumes that the cross-price elasticity of demand between the industries is zero. These rather restrictive assumptions allow Sartzetakis to confine his analysis to the industry in question, without having to consider the "knock-on" impacts on other sectors.[15]

In his model, the two firms have identical production technologies, with constant marginal cost of production (c), and constant emissions rate per unit output (ε). Consequently, in the absence of any regulatory intervention, the market equilibrium is symmetric, with each firm having the same market share and the same level of emissions. The abatement cost for each firm is a quadratic function of its output (y) and permit holding (L) – i.e. $A(y, L) = k(\varepsilon y - L)^2$, where k is a constant that differs between firms.[16] Thus, each firm's marginal cost of output is equal to its marginal cost of production plus the impact of the increase in output on its abatement cost – which is equal to the common emission rate multiplied by its marginal cost of abatement.

[15] Of course, the intra-industry reallocation of permits under trading will affect the outcome in each industry – unless initial permit allocations are efficient.

[16] These functional forms are consistent with both firms having identical Leontief productions technologies with emissions being a linear of production inputs; and having different end-of-pipe abatement technologies with a production function of the form $Z = a\, w^{1/2}$, where w is some input and a is a firm-specific constant.

Sartzetakis assumes that permits are allocated on the basis of pre-regulation emissions (i.e. grandfathered), which implies that the initial allocations are equal. Consequently, if the permits are non-transferable, the firm that is more efficient in abatement has the lower marginal cost of output, and hence increases its market share relative to the pre-regulation equilibrium. If it is very efficient the firm may actually increase its output, although aggregate output will always be lower than the pre-regulation level.

Under the tradable permit system, firms trade permits until marginal abatement costs are equalised – and hence so too are the firms' marginal costs of output and market shares. Under Sartzetakis' assumptions of quadratic abatement cost functions and intra-industry market clearing, the efficient firm reduces output while the inefficient firm increases output. The impact on aggregate output depends on the permit price, with output increasing if this is less than the average of the firms' marginal abetment costs under the fixed emission limits – which Sartzetakis shows is indeed the case. The rise in aggregate output causes the market price to fall. Total surplus increases, with the increase in consumer surplus more than offsetting the decline in aggregate operating profit (i.e. excluding abatement costs).[17]

Permit trading minimizes the aggregate abatement cost for a given level of aggregate output. However, since the minimum cost increases with output, it is possible that the rise in output may result in the aggregate abatement cost being higher under trading. Sartzetakis claims that this is indeed the case, and that the increase in abatement cost is due to the reallocation of market shares from the efficient firm to the inefficient firm. However, it can be shown that neither of these claims is true in general. In particular,

[17] Sartzetakis concludes that the increase in aggregate output leads to a reduction in aggregate operating profit (i.e. excluding abatement costs). However, this implicitly assumes that the combined output under the fixed emissions limits is greater than the monopoly quantity (i.e. $Y^{\#} > (p^0 - c) / 2b$). If the emissions limit is sufficiently restrictive that the aggregate output under trading is less than the monopoly quantity (i.e. $Y^* < (p^0 - c) / 2b$), then aggregate profit is also higher when trading is allowed. In the intermediate case, where $Y^* > (p^0 - c) / 2b > Y^{\#}$, the impact of allowing trading on aggregate profit is ambiguous.

under his assumption of a constant emissions rate, it is straightforward to show that for any given level of aggregate output, the minimum aggregate abatement cost is independent of the firms' market shares.[18] Thus any increase in abatement cost is driven solely by the increase in aggregate output.

If the aggregate abatement cost declines then welfare is unambiguously higher under trading. However, if the abatement cost increases then the overall impact on welfare depends on whether this is outweighed by the increase in total surplus. Sartzetakis shows that, under his assumptions, this is indeed the case.

In a separate analysis, Sartzetakis (1997b) considers a situation in which the pollutant is only emitted by the industry in question, and hence the two duopolists are the only participants in the market for emission permits. He assumes that one of the firms (the "leader") has the power to set the price of emission permits, and assesses whether it can use this power to reduce the existing and potential level of competition in the product market.[19] As such, his analysis complements that of Misiolek & Elder (1989) – evaluating the impact of exclusionary manipulation from the perspective of the product market.

The interaction between the two firms is modelled as a two-stage game. In first stage, the leader chooses the price for emission permits. In the second stage, both firms make their output and abatement decisions (taking the permit price as given), with the leader's choice constrained to clear the permit market. The game structure is analogous to that underlying Misiolek & Elder's analysis of exclusionary manipulation (see previous section), and hence it follows that the leader will set the price of emission permits above

[18] It is also possible to construct a simple numerical example in which the aggregate abatement cost is lower when permits are tradable.

[19] Given the symmetry of the permit market, it is not clear why one firm should be able to exercise such power. In this situation, a bargaining model would appear to be more appropriate, and this is the market institution that is adopted by the other authors in this group (see page 297).

the competitive level.[20] While this has the effect of raising both firms' net production costs, the increase is greater for its rival, and consequently there is a reallocation of output towards the leader. This is reflected in an increase in the leader's profit, and a fall in it rival's. However, the impact of the reallocation on total industry profits and / or overall welfare is unclear. Because of the strategic interaction in the output market, the outcome under the benchmark (i.e. perfect competition in the permit market) is not efficient, and hence it is possible that one, or both, may actually increase.

In order to explore the impacts on industry profits and welfare, Sartzetakis employs a numerical simulation in which the firms have a common, constant emission rate per unit output, and both have quadratic production cost and abatement cost functions. The firms' resultant marginal production and marginal abatement cost functions have common slopes, but different intercepts. Under these assumptions, the trading of emission permits has no impact on aggregate output.[21] Consequently, the impact on welfare is equal to the impact on aggregate profit.

In this simulation, the leader's profits increase regardless of the firms' relative production and abatement efficiencies. However, the impact on aggregate industry profit (and hence social welfare) depends on the relative efficiencies. If the firms have similar production and abatement technologies, or the leader is less efficient, then industry profit declines. Industry profit increases if the leader is more efficient in production, or the follower is more efficient in abatement.

Sartzetakis claims that the increases in the firms' marginal costs of production and abatement can both be expressed as functions of the increase in the permit price versus its competitive benchmark value. Thus, the permit price increase provides a good

[20] As,before, this implies that the leader over-buys / under-sells emission permits in the sense that the marginal factor cost / marginal revenue exceeds its marginal abatement cost.

[21] In general the impact on aggregate output is ambiguous.

indicator of the strength of the exclusionary manipulation. Using the same simulation, he shows that the magnitude of the permit price increase is inversely related to the leader's pre-regulation market share (which is determined by its relative production efficiency); directly related to its share of the initial permit allocation; and directly related to its relative abatement efficiency.

He goes on to consider whether exclusionary manipulation might be used as a barrier to entry to new competitors; augmenting the simulation with a potential market entrant. This firm – which would enter the industry in the absence of regulation – is not given any emission permits in the initial allocation, and so must purchase all of its emission permits from the two incumbents. It is sufficiently efficient to make positive profits when the permit market is perfectly competitive. However, when the permit market is subject to price manipulation, entry is only profitable if it is more efficient in abatement than the leader. If the leader is more efficient, then it is able to set a permit price that deters the new firm from entering the market. This leads to a reduction in aggregate output, compared to the outcome with a perfectly competitive permit market (with all three firms), and hence a reduction in consumer surplus. However, this is outweighed by the increase in industry profit, leading to a rise in overall welfare.

Finally, Sartzetakis considers the case where the permits are auctioned, and shows that this can actually make matters worse. If the firms are all equally efficient then the new firm enters under auctioning (increasing competition in the product market), but social welfare is lower. If the leader is more efficient in abatement, then it buys all of the permits in the auction; forcing Firm 2 out of the market. In this case, both the level of competition and social welfare are lower.

The other three analyses in this group – von der Fehr (1993), Requate (1993), Fershtman & de Zeeuw (1995) – also assume that the pollutant is not generated by any other industry, and hence that the two firms are the only participants in the permit market. However, in these analyses, the authors assume that the number of permits traded, and the terms on which they are traded, are determined by a bargaining process between the firms. Only "cash contracts" are allowed, and no conditions can be attached to the transfer regarding output levels in the product market. Consequently, the firms cannot engage in direct collusion. However, since the final distribution of permits will affect the equilibrium in the product market, the firms can use the transfer of permits to manipulate the outcome.[22]

The analyses all share a common structure; with interaction between the firms being modelled as a two-stage game. In the first stage, the firms bargain over the number of permits to transfer and the size of the cash payment. In the second stage, the firms engage in a Cournot game, in which they determine their respective output and abatement levels, subject to the constraints imposed by their respective permit holdings. There are gains from trading permits if the sum of the Nash equilibrium profits in the second stage Cournot game is greater than the corresponding sum under the original allocations. If this is the case, then there exist payment values such that both firms are better off.

None of the authors analyse the first-stage bargaining game in detail. However, they all note that the equilibrium profits under the initial allocation define the threat point for the game, and that a necessary condition for any Paretian solution is that the resultant allocation of permits maximizes the sum of the firms' equilibrium profits in the second stage. Put another way, the gain from trading must be maximized. The division of the surplus between the two firms (i.e. the value of the transfer payment) is not considered.[23]

[22] As will be seen, the particular assumptions in Requate (1993) mean that the number of permits held by a firm commits it to a particular output level

[23] However, Fershtman & de Zeeuw note that if the firms are risk-neutral (i.e. have linear preferences), then the bargaining frontier is linear, and the Nash bargaining solution is for the surplus to be split equally.

Von der Fehr (1993) considers the case of two identical firms, and assesses whether tradable emission permits are likely to provide an instrument for monopolisation of the product market. Apart from the symmetry of the firms, his analysis is more general than the others in the group, in that he uses general function forms for the conditional production cost function and the revenue function (i.e. $C(y_i, L_i)$ and $R(y_i, y_j)$), with the latter allowing the possibility of product differentiation. For a given total number of permits, he demonstrates that the impact of a reallocation of permits (from Firm 2 to Firm 1) on industry profit can be decomposed into three components: a) a direct cost savings effect; b) an output reallocation effect; and c) an aggregate output change effect.[24]

The direct effect on costs can be in either direction. If products are homogeneous, and the cost function is homogeneous of degree one, then total costs increase. However, in other cases (for example where there are scale economies) total industry costs may decline as a result of the reallocation. If an increase in emission permits reduces the marginal cost of production (i.e. $C_{yL} < 0$), and if the products are strategic substitutes, then the reallocation of permits leads to an increase in output for Firm 1, and lower output for Firm 2. A sufficient condition for the impact of this output reallocation to be non-negative is that $R_{12} \leq 0$ and $R_{22} \geq 0$.[25] These conditions are satisfied for a wide range of revenue functions (e.g. linear inverse demand), and hence the output reallocation is likely to increase industry profits. Furthermore, if aggregate output is reduced – as it will be for most specifications of revenue and cost functions – the output effect will also be positive.

While it is not possible to conclude that industry profits will necessarily increase as the permit allocation becomes more asymmetric, this is likely to be the case for a wide

---

[24] Since the firms are identical in all other respects, the direction of the reallocation does not matter.

[25] R_{22} denotes the second partial derivative of the firm's revenue function with respect to its rival's output.

range of parameter values. In particular, von der Fehr shows that sufficient conditions for industry profits to increase are: a) the products are homogeneous; b) the cost function is sub-additive in outputs and emission permits[26]; and c) emission permits are essential for profitable operation. Since these conditions are not particularly restrictive, he concludes that tradable emission permits can allow the monopolisation of the product market in a wide range of circumstances.

Assuming that aggregate output declines, the reallocation of permits leads to an increase in the market price of the output product, and a reduction in consumer surplus. In order to explore the overall impact on welfare, von der Fehr uses a simple numerical example with specific functional forms for the revenue and cost functions. He finds that there is a threshold value for the cost of abatement, below which welfare is maximized by an equal permit allocation. However, above this threshold value, welfare is maximized by allocating all of the permits to one firm, enabling it to act as a monopoly in the product market.

The same example is used to assess the impact of permit reallocation on industry profits when the sufficient conditions are violated.[27] He finds that if the cost of abatement is sufficiently low, then industry profit is maximized at an equal allocation (i.e. a symmetric equilibrium) if the products are sufficiently differentiated, or if there are sufficient diseconomies of scale. Therefore, it does not follow necessarily that allowing emission permits to be traded will lead to the monopolisation of the product market.

Requate (1993) and Fershtman & de Zeeuw (1995) make specific assumptions about the firms' production technologies and emission functions. In particular, they assume that

[26] A function $F(x, y)$ is sub-additive in x and y if $F(x_1 + x_2, y_1 + y_2) \leq F(x_1, y_1) + F(x_2, y_2)$. Sufficient conditions for a function to be sub-additive are that it is convex and homogeneous of degree $r \geq 1$.

[27] In this example, the cost function takes the form $C(y, L) = c\,y^2 + d\,(y/L)$, and the inverse demand function for firm $i = 1, 2$ is $p_i = 1 - a\,y_i - b\,y_j$ (with $a \geq b$). The sufficient conditions are violated if $c > 0$, or if $a > b$.

the firms' products are undifferentiated and that demand is linear. Furthermore, the firms' marginal production costs (c_i) are constant; as are their emission rates per unit output (ε_i). As Requate notes, this is consistent with the firms having Leontief production technologies, with emissions being a linear function of inputs. For convenience, both analyses assume that $c_1 \leq c_2$ (i.e. that Firm1 is at least as efficient in production as Firm 2). No *a priori* assumptions are made about the values of the respective emission rates, although Fershtman & de Zeeuw assume that $\frac{1}{2} \varepsilon_i < \varepsilon_j < 2 \varepsilon_i$, (i.e. that the emission rates of the two firms are not too different).

Requate assumes that there are no abatement technologies, and hence that the only way for either firm to reduce its emissions is to reduce its output. In contrast, Fershtman & de Zeeuw assume that the firms can reduce emissions either by reducing output, or by engaging in abatement, or both. The unit cost of abatement (a_i) is constant, and hence so too is the marginal abatement cost of output – being equal to equal to the unit abatement cost multiplied by the firm's emission rate. Consequently, if the firm engages in abatement (at any level), it raises its marginal cost of output by a fixed amount.[28]

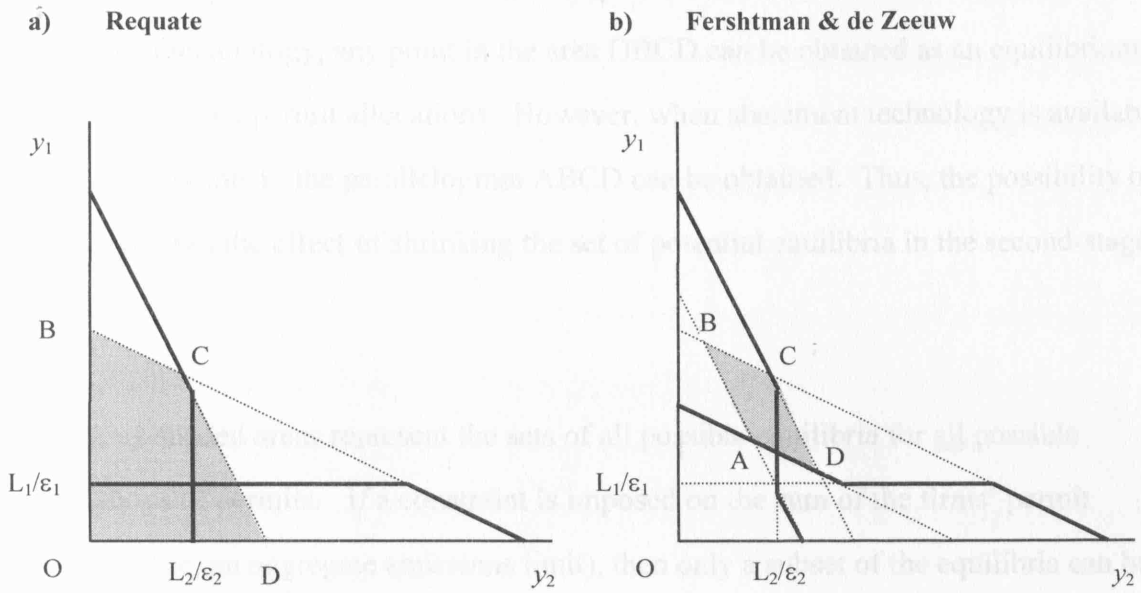
The different assumptions that are made about the availability of abatement technologies means that the shapes of the firms' reaction functions in the second-stage Cournot game differ between the two analyses. Consequently, so too do the sets of potential equilibria. This is illustrated in Figure 7.3.

When there is no abatement technology, the firms' reaction curves in the Cournot game are "kinked". If its rival's output is above a threshold value, then the firm's response is determined by its pre-regulation reaction curve. However, if its rival's output falls

[28] The marginal cost of output is equal to its marginal production cost (c_i) if it does not engage in any abatement; and equal to its marginal production cost plus its marginal abatement cost of output ($c_i + a_i\varepsilon_i$) if it does.

below this threshold, then the firm's response is given by its permit holding divided by its emission rate (i.e. the maximum output level allowed by its holding). It follows immediately from the geometry of the firm's reaction function that the threshold value for its rival's output is equal to twice the difference between the firm's own monopoly output quantity and the output allowed by its permit holding.

Figure 7.3 Reaction curves and feasible equilibria



When an abatement technology is available, each firm has two parallel reaction curves – an upper curve that applies if it does not engage in any abatement, and a lower curve that applies if it does (and hence has a higher constant marginal cost of output). Consequently, each firm's reaction curve is "double-kinked", with two threshold values for its rival's output quantity. The first (upper) threshold is the same as that in Requate's model; with the firm's response being determined by its upper reaction curve if its rival's output exceeds this value. The second (lower) threshold is also equal to twice the difference between the firm's monopoly output quantity and the output allowed by its permit holding – only now the monopoly quantity reflects the higher marginal cost when it engages in abatement. If its rival's output falls below this threshold, then the firm's response is determined by its lower curve. Only if the rival's output lies between the two threshold values, is the firm's response fixed by its permit holding.

In either case, it is clear that the intersection of the firms' reaction curves (and hence the equilibrium) depends on their respective holdings of emission permits. The grey shaded areas in Figure 7.3 represent the sets of equilibria that can be obtained by varying the permit holdings of the two firms without any constraint on the total. When there is no abatement technology, any point in the area OBCD can be obtained as an equilibrium with appropriate permit allocations. However, when abatement technology is available, only those points in the parallelogram ABCD can be obtained. Thus, the possibility of abatement has the effect of shrinking the set of potential equilibria in the second-stage game.

The grey shaded areas represent the sets of all possible equilibria for <u>all</u> possible allocations of permits. If a constraint is imposed on the sum of the firms' permit holdings (i.e. an aggregate emissions limit), then only a subset of the equilibria can be attained by varying the allocations; with the subset changing as the number of permits is reduced. This suggests the use of a two-step procedure to identify the equilibrium permit allocation in the first-stage bargaining game. First, all of the potential equilibria that can be achieved in the second stage for a particular aggregate emissions limit are determined. Then, from this subset the particular equilibrium (i.e. industry structure) that maximizes joint profits is identified.[29]

When there is no abatement technology available, Requate demonstrates that the equilibrium industry profit can be maximized by a number of industry structures, depending on the values of the firms' marginal costs of production (c_1 and c_2) and

[29] This is the procedure that is used by Fershtman & de Zeeuw. Requate does the opposite – first identifying the allocation that maximizes joint profits, and then showing that this is a Nash Equilibrium.

emission rates (ε_1 and ε_2), and hence the social marginal cost of output ($\hat{c}$)[30]; the choke price (p^0); and the stringency of the aggregate emissions limit (i.e. L).

Equilibrium industry profits are maximized by Firm 1 acquiring all of the permits and acting as a monopolist if: a) it has the lower emission rate ($\varepsilon_1 < \varepsilon_2$); b) if the social marginal cost of output is greater than the choke price ($\hat{c} \geq p^0$); or c) if the emissions limit is greater than an upper threshold value ($L \geq L^+$).[31] In any of these cases, if the total number of permits is greater than the quantity necessary for it to produce its unregulated monopoly output (i.e. point B in Figure 7.3), then it does not use all of the permits, and hence emissions are less than the aggregate limit. Furthermore, a small reduction in the limit will have no impact on the outcome.

If the emissions limit lies between the upper threshold value and a lower threshold value (i.e. $L^+ > L > L^-$), then equilibrium industry profits are maximized by both firms producing. As the limit decreases between the two threshold values, total output remains constant, but production shifts from Firm 1 to Firm 2, implying an increase in the average cost of production.

Finally, if the emissions limit is less than or equal to the lower threshold value, then equilibrium industry profits are maximized by Firm 2 acquiring all of the permits and acting as a monopolist. In contrast to the case where Firm 1 acts as a monopolist, Firm 2 always uses all of the permits. This is because the lower threshold value is less

[30] The minimum social cost of producing a given level of aggregate output is given by $\hat{C}(Y; \delta) = \text{Min } c_1 y_1 + c_2 y_2 + D(\varepsilon_1 y_1 + \varepsilon_2 y_2)$ subject to $y_1 + y_2 \geq Y$, where $D(\cdot)$ is the social damage function. If both firms produce positive quantities in the solution, then the marginal social cost of output is constant, being given by: $\hat{c} = (\varepsilon_1 c_2 - \varepsilon_2 c_1) / (\varepsilon_1 - \varepsilon_2)$.

[31] The upper and lower threshold values are determined by the firms' respective monopoly output quantities in the absence of any regulation and their emission rates. As such, they are both exogenously determined.

than the number of permits necessary for it to produce its unregulated monopoly output (i.e. point D in Figure 7.3).

The outcome also depends on the stringency of the emissions limit when abatement technologies are available. Fershtman & de Zeeuw demonstrate that if the limit is very low (i.e. both firms have to engage in abatement under any allocation of the permits), then there is only one feasible equilibrium – which is determined by the intersection of the firms' respective lower reaction curves (i.e. point A in Figure 7.3). Hence, any reallocation of permits has no impact on equilibrium output levels (individual or aggregate), market price, or operating profits.[32] However, there is an impact on the allocation of the abatement effort between the two firms, with the lower abatement cost firm selling all of its permits to the other. Consequently, there is a reduction in aggregate abatement costs.

If each firms initial allocation is greater than its pre-regulation emissions (which implies that the aggregate emissions limit is not binding), then the subset of feasible equilibria is given by the upper boundary of the parallelogram (i.e. BCD in Figure 7.3). In this case, the imposition of the emissions limit has no impact on the equilibrium if the trading of permits is prohibited. However, if trading is allowed, then the joint profits can be increased by moving away from the pre-regulation equilibrium along one of the firm's upper reaction curves.

The direction and extent of the movement depends on the respective values of the firms' marginal production costs and marginal abatement costs of output.[33] If Firm 1 also has

[32] The reallocation of permits shifts the kink in each firm's reaction curve (i.e. the jump from its upper curve to its lower curve). However, it has no impact on the equilibrium, which is always given by the intersection of the firms' respective lower reaction curves.

[33] Neither firm engages in any abatement in the equilibrium. Hence, the reallocation of permits has no impact on abatement costs, which remain equal to zero. However, the value of each firm's (potential)

the lower marginal abatement cost of output, then the "optimal" equilibrium is given by the intersection of its upper reaction curve with Firm 2's lower curve (i.e. point B in Figure 7.3). Consequently, there is a reallocation of output to Firm 1. However, if this is not the case, then – depending on parameter values – it is possible that joint profits may be maximized by moving in the opposite direction (i.e. along CD); in which case Firm 2 gains market share, and the average cost of production rises. In either case, aggregate output declines, and hence the market price rises.

Changing the initial permit allocations (keeping the aggregate limit constant) in either of the above cases has no impact on the equilibrium. However, in the second case one can no longer draw any conclusions about the impact of trading on aggregate output. Depending on the initial allocation of permits, output may decline, rise, or stay the same.

Between these two extreme cases, there are a whole range of different intermediate cases that can be considered. Fershtman & de Zeeuw analyse a number of these, and they show that the position of the equilibrium under trading depends on the respective values of the firms' parameters. In these cases, it is not possible to draw any general conclusions about the impact of trading on industry output. Depending on the initial allocation of permits, it may decline, rise, or stay the same. However, in the symmetric case, where firms are identical, the trading of permits will always lead to a reduction in industry output unless the initial permit allocation is optimal (i.e. it already maximizes joint equilibrium profits).

Neither Requate nor Fershtman & de Zeeuw consider how the trading of permits affects welfare. However, it is clear that the impact on gross welfare may be in either direction, depending on the stringency of the emissions limit and the initial allocation of the

---

marginal abatement cost of output determines how far it is possible to move the equilibrium along the other firm's upper reaction curve by reallocating permits.

permits. For example, in the first case considered by Fershtman & de Zeeuw, welfare clearly increases as the only impact is to reduce aggregate abatement costs. However in the second case, gross welfare declines if there is a reallocation of output to the firm with the higher cost of production, or if the firms are symmetric. In the intermediate cases, the impact on welfare will depend on the initial allocations and on the relative values of the firms' marginal costs of production. If the trading of emission permits causes aggregate output to rise, then gross welfare is unambiguously higher. However, if it causes it to fall, then the impact depends on whether the increase in aggregate profits outweighs the reduction in consumer surplus.

Furthermore, both Requate and Fershtman & de Zeeuw show that in certain cases aggregate emissions fall when permit trading is allowed, as one of the firms acquires all (or most) of the permits but does not use them all. In these cases, if the resultant reduction in social damages is sufficiently large, then it is possible that net welfare may improve even if gross welfare does not.

7.3 Market power with a relative performance standard

Mauleg (1990) is the only author to have considered the issue of market power in the context of a relative performance standard; assessing the implications of using a performance-based credit trading scheme to implement a target rate for emissions per unit output when the product market is imperfectly competitive. In particular, he considers the impact on the equilibrium for an oligopolistic product market; using the outcome under a common performance standard as the benchmark for the comparison.

Like Sartzetakis (1997a), Mauleg assumes that the pollutant is emitted by several industries, and hence that the there are many participants in the market for emission credits – all of whom act as price-takers. Furthermore, by confining his analysis to a single industry, he implicitly shares Sartzetakis' assumption that the cross-price elasticities with the products produced by other industries are all zero, and that there are no net-transfers of emission credits to or from other industries when trading is allowed.[34]

While Mauleg assumes that the regulatory intervention takes the form of a relative performance standard for emissions per unit output, it would appear that the underlying environmental objective is defined in absolute terms. For example, in the introduction to his analysis he states that "... *as long as the introduction of the emission credit trading programme maintains the aggregate emissions level, health consequence associated with the pollutant are unchanged*". This inference is supported by the constraint that he imposes on the relative values of the parameters in his analysis (see below) to ensure that the trading of emission credits does not affect the aggregate emissions for the industry.

In Mauleg's model the industry is characterised as a homogeneous Cournot oligopoly with a fixed number of firms (N); each firm having a constant marginal cost of production (c_i) and a constant emission rates per unit output (ε_i).[35] For any given level of output, each firm's abatement cost is given by $A^i(Y_i, \rho_i) \equiv Y_i\, a^i(\rho_i)$, where ρ_i is its abatement per unit output and $a^i(\rho_i)$ is its abatement cost per unit output.[36] Consequently, for a given level of unit abatement (i.e. abatement effort), each firm's marginal abatement cost of output $a_i(\rho_i)$ is constant, and hence so too is its gross marginal cost of output.

[34] Unlike Sartzetakis however, Mauleg does not state these assumptions explicitly.

[35] As noted previously, this is consistent with the firms having Leontief production technologies and linear emission functions.

[36] This implies that the firm's abatement cost function $a(R)$ is linearly homogeneous. In which case, for any given level of output Y, $a(R) = Y\, a(R/Y)$. Mauleg further assumes that the function is non-negative and convex.

Mauleg's analysis is founded on the assumption that the trading of emission credits reduces the net marginal production costs of all firms – irrespective of whether they are buyers or sellers. At first sight, this assumption does not appear to be justified. In general one would expect that while the net marginal cost of production would fall for a buyer of credits, it would rise for a seller (although the total net costs of production would fall for both). This is certainly the case when the regulatory intervention takes the form of an absolute limit. However, when the intervention takes the form a relative performance standard for emissions per unit output, it is straightforward to show that the introduction of trading will reduce the net marginal production costs of all firms provided that their production and emission functions are both homogeneous of degree one.[37]

On the basis of this fundamental assumption, Mauleg evaluates the impact of the introduction of the trading scheme using the following standard solutions for a homogeneous Cournot oligopoly with linear inverse demand ($P(Y) = p^0 - bY$):

$$Y^{\#} = \frac{1}{(N+1)\,b}\left[N p^0 + \sum_{i \in I} \hat{c}_i \right] \qquad \ldots (7.1.a)$$

$$y_i^{\#} = \frac{1}{(N+1)\,b}\left[p^0 + \sum_{j \in I} c_j - (N+1)\,\hat{c}_i \right] \qquad \ldots (7.1.b)$$

$$\Pi_i^{\#} = (y_i^{\#})^2\, b \qquad \ldots (7.1.c)$$

where $\hat{c}_i$ is the net marginal production cost of firm $i \in I$.

Since the introduction of the trading scheme lowers the marginal cost $\hat{c}_i$ for all firms, it is clear form (7.1.a) that the aggregate output $Y^{\#}$ increases, which in turn implies that the market price declines and the consumer surplus increases. However, the impact on

[37] See Appendix A7.1. The linear homogeneity of the two functions is consistent with Mauleg's assumption that the firms' marginal production costs and (uncontrolled) emission rates are constant.

the outputs of individual firms, and hence their profits, is not so obvious. Taking the total differential of (7.1.b) gives:

$$dy_i^{\#} \quad = \quad \frac{1}{(N+1)\,b}\left[\sum_{j\in I} d\hat{c}_j - (N+1)\,d\hat{c}_i\right] \tag{7.2}$$

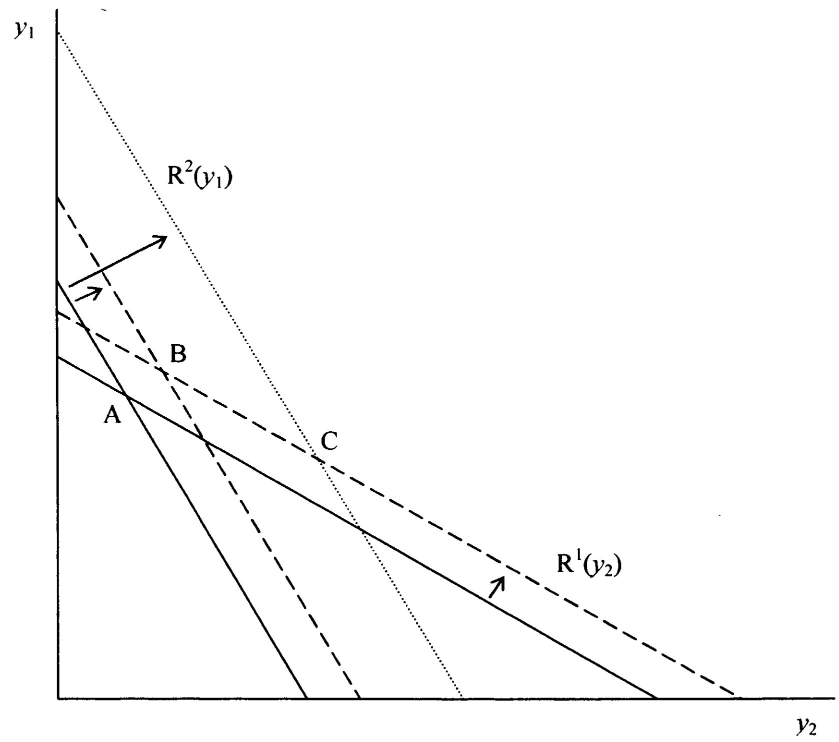
Thus the output and profit of firm $i \in$ I increases if and only if:

$$|d\hat{c}_i| \quad > \quad \left(\frac{N}{N+1}\right)\frac{\sum_{j\in I}|d\hat{c}_j|}{N}$$

Since the right hand side of this condition is strictly less than the average cost reduction across all firms, it will always be satisfied for at least one firm. It is also clear that if there is no variation in the cost reductions (i.e. $d\hat{c}_i = d\hat{c} < 0$ for all $i \in$ I) then the outputs and profits of all firms will increase. By continuity, this will also be the case when the firms' cost reductions are approximately equal. However, if there is a significant variation in the cost reductions then there will be at least one firm that suffers a reduction in its output and profit. The extent to which the cost reductions can vary before this occurs depends on the number of firms (N). As this increases, the ratio $N/(N+1) \rightarrow 1$, and it therefore becomes increasingly likely that there will be some losers from the introduction of the trading scheme.

The relative impacts on firms' output levels are illustrated in Figure 7.4 for the case of a duopoly. When trading is not allowed the firms reaction curves are shown as solid lines, and the resultant equilibrium is represented by the point A. The dashed reaction curves represent the situation where the impact of the introduction of trading on the marginal production cost is the same for both firms. In this case, the equilibrium moves to point B, with both firms enjoying increases in output. However, if the reduction in marginal cost is greater for Firm 2, then its reaction shifts further out (to the dotted line), with C representing the resultant equilibrium. In this case the output (and profit) of Firm 2 increases, but that of Firm 1 declines.

Figure 7.4: Impact of trading on reaction curves



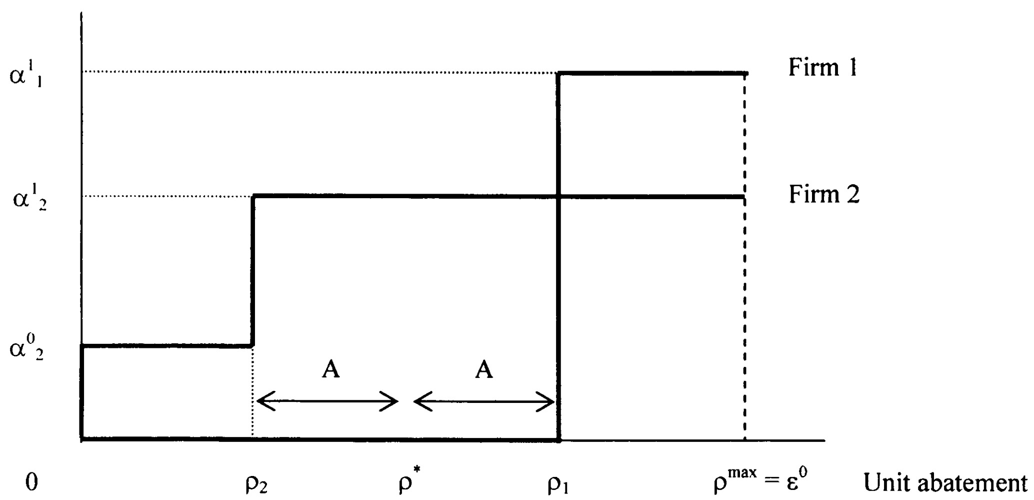
The impact on aggregate profit depends on which (if any) firms experience a reduction in their output. To see this, Mauleg notes that for small changes in output levels, the change in aggregate profit is approximately

$$d\Pi^{\#} \quad = \quad 2b \sum_{i \in N} y_i^{\#} \, dy_i^{\#}$$

Thus, aggregate profit will decline if it is the firms with relatively high market share that suffer the reduction in output. This will be the case if the reductions in net marginal production costs resulting from the introduction of the trading scheme are greatest for those firms with the higher net marginal costs under the common emissions standard. If aggregate profit increases then welfare unambiguously increases. However, if aggregate profit declines, then the impact on gross welfare depends on the relative magnitudes of the changes in aggregate profit and consumer surplus.

In order to explore the impact on welfare, Mauleg employs a simple numerical example in which the industry comprises two firms with identical marginal costs of production (c^0) and pre-regulation emission rates (ε^0), and whose marginal costs of unit abatement take the form of step-functions shown in Figure 7.5. The firms' unit abatement costs $a^i(\rho_i)$ are given by the areas under their respective step-functions. Thus, while Firm 2 incurs a cost for any level of abatement, Firm 1 can reduce its emissions rate by up to ρ_1 without incurring any cost. As noted above, Mauleg imposes a constraint on the relative values of the various parameters to ensure that the aggregate emissions of the two firms when trading of emission credits is allowed is the same as under the common emissions standard.

Figure 7.5 Marginal costs of unit abatement



Mauleg does not justify the use of step functions for the firms' marginal costs of unit abatement, other than on the grounds of simplicity. However, the functions would take the forms shown in Figure 7.5 if each firm has three alternative plants that it can use to produce output – with each plant having different Leontief technologies whose (constant) marginal costs of production and emission rates satisfy the following conditions:

$$c^0 = c_1^0 = c_1^1 < c_1^2 = c^2 \text{ and } \varepsilon^0 = \varepsilon_1^0 > \varepsilon_1^1 = r - A > \varepsilon_1^2 = \varepsilon^2 \qquad \text{for Firm 1}$$

$$c^0 = c_2^0 < c_2^1 < c_2^2 = c^2 \text{ and } \varepsilon^0 = \varepsilon_2^0 > \varepsilon_2^1 = r + A > \varepsilon_2^2 = \varepsilon^2 \qquad \text{for Firm 2}$$

where r is the common target emission rate and A is an arbitrary constant. Thus, the "low cost / high emission" and "high cost / low emission" plants for each firm are identical, and only the intermediate plants differ; with Firm 1 having the lower emissions rate and the lower unit cost. In the absence of any regulation each firm produces all of its output in its "low cost / high emission" plant with unit production cost c^0 and emission rate ε^0. However, in order to reduce their emission rates below ε^0, the firms must switch production to the alternative plants, and the areas under the step-functions represent the resultant increases in their respective unit production costs.[38]

Under the common performance standard, each firm has to reduce its emissions per unit by the same amount ρ^*. While Firm 1 can achieve the necessary reduction in its emissions rate without incurring any cost, Firm 2 incurs a unit abatement cost equal to $\alpha_2^0\, \rho^* + (\alpha_2^1 - \alpha_2^0)\,(\rho^* - \rho_2)$. Consequently, the sales and profits of Firm 1 increase relative to the pre-regulation equilibrium, while those of Firm 2 fall.[39]

Mauleg does not model the market for emissions credits explicitly when trading is allowed. Rather he makes the arbitrary assumption that the equilibrium price of credits

[38] Under the stipulated conditions for the marginal production costs and emission rates, neither firm will ever use more than two of its plants, nor simultaneously use its "low cost / high emission" and "high cost / low emission" plants. Consequently, $\alpha_i^m = (c_i^{m+1} - c_i^m) / (\varepsilon_i^m - \varepsilon_i^{m+1})$.

[39] While the firm with lower cost increase always gains market share, it is not generally the case that is experiences an increase in absolute sales. It follows directly from (7.1.b) that the absolute sales of both firms will decline if and only if $2\Delta\hat{c}_2 > \Delta\hat{c}_1 > \Delta\hat{c}_2 / 2$.

$q^{\#} = \alpha^0{}_2$ in Figure 7.5, and that this induces Firm 1 to increase its abatement per unit output to ρ_1, while Firm 2 reduces its abatement effort to ρ_2. Unfortunately, by not considering the market for emission credits, Mauleg fails to realize that this symmetric response in terms of unit abatement does not imply symmetry in terms of credit transfers. Because of the differences in the two firms' output levels, the number of credits generated by the Firm 1 is greater than the number required by Firm 2. Since the market price of credits is strictly positive, these excess credits must be purchased by firms in other industries. This has two implications. First, the aggregate emissions of these industries will be higher than under the common performance standard. Second, unless the marginal abatement costs of all firms in the other sectors are constant, there will be an increase in the aggregate outputs of these industries with corresponding welfare impacts.

In order to ensure this does not occur, it is necessary to impose an intra-industry market clearing constraint on the transfers of emission credits. Unfortunately, since the x-axis in Figure 7.5 represents unit abatement rather than absolute abatement, it is not possible to interpret the step-functions as inverse supply and demand curves for emission credits, and hence to derive the market equilibrium directly. However, it is clear that there are only four possible candidates:

(i)	$0 < q^{\#} < \alpha_2^0$	$\rho_1^{\#} = \rho_1^1$	$\rho_2^{\#} = 0$
(ii)	$q^{\#} = \alpha_2^0$	$\rho_1^{\#} = \rho_1^1$	$0 \leq \rho_2^{\#} \leq \rho_2^1$
(iii)	$\alpha_2^0 < q^{\#} < \alpha_2^1$	$\rho_1^{\#} = \rho_1^1$	$\rho_2^{\#} = \rho_2^1$
(iv)	$q^{\#} = \alpha_2^1$	$\rho_1^{\#} = \rho_1^1$	$\rho_2^1 \leq \rho_2^{\#} \leq \rho^{max}$

It is straightforward to show that the market cannot clear under candidates (i), (iii) and (iv), and – as has already been noted – neither does it clear under candidate (ii) with

$\rho_2^{\#} = \rho_2^{1}$ (i.e. the equilibrium assumed by Mauleg). For the market to clear it must be the case that:

$$q^{\#} = \alpha_2^{0} \qquad \rho_1^{\#} = \rho_1^{1} \qquad 0 < \rho_2^{\#} = \left(\frac{y_1^{\#}}{y_2^{\#}}\right)\rho_2^{1} < \rho_2^{1}$$

where $y_1^{\#}$ and $y_2^{\#}$ are the equilibrium output quantities of the two firms.

Mauleg adopts the following parameter values for his simulation, with the values being chosen to yield a wide variation in the cost reductions of the two firms when trading of emission credits is allowed, while maintaining the same aggregate level of emissions.

p^0	=	2	b	=	1	r	=	1
ε^0	=	36	ε^2	=	0	A	=	0.135
c^0	=	1	c_2^1	=	1.3487	c^2	=	2.5450

From which it follows that:

ρ^*	=	35	ρ_1	=	35.135	ρ_2	=	34.865
α_1^0	=	0	α_2^1	=	1.78613			
α_2^0	=	0.01	α_2^1	=	1.05405			

Table 7.3 gives the simulation outcomes under the common performance standard and under the trading scheme. Three different outcomes are shown for emission credit trading. The first is the outcome derived by Mauleg. The reduction in marginal net cost of production arising from the introduction of trading is greatest for Firm 2 – which has the higher cost under the common performance standard. Consequently, aggregate profit declines, and this more than offsets the increase in consumer surplus. However, as has already been noted, in this outcome the market for performance credits does not clear; the supply of credits exceeding the demand. As a result, the actual aggregate emission rate of 0.914 (i.e. 0.5030 / 0.5505) is less than the target rate.

Table 7.3 Market equilibrium outcomes

	Common SEC standard	Emission credit trading			Diff. vs. common std.	
		(1)	(2)	(3)	(2)	(3)
c_1	1.0000	0.9987	0.9987	0.9995	(0.00135)	(0.00050)
c_2	1.4909	1.3500	1.3500	1.3509	(0.14095)	(0.14009)
y_1	0.4970	0.4509	0.4509	0.4506	(0.04608)	(0.04637)
y_2	0.0060	0.0996	0.0996	0.0993	0.09351	0.09323
Y	0.5030	0.5505	0.5505	0.5499	0.04743	0.04686
π_1	0.2470	0.2033	0.2033	0.2031	(0.04368)	(0.04394)
π_2	0.0000	0.0099	0.0099	0.0099	0.00987	0.00982
Π	0.2470	0.2132	0.2132	0.2129	(0.03381)	(0.03412)
CS	0.1265	0.1515	0.1515	0.1512	0.02498	0.02467
GW	**0.3735**	**0.3647**	**0.3647**	**0.3641**	**(0.00882)**	**(0.00945)**
E	0.5030	0.5030	0.5505	0.5030	0.04743	0.00000
D(E)	0.0503	0.0503	0.0551	0.0503	0.00474	0.00000
NW	**0.3232**	**0.3144**	**0.3097**	**0.3138**	**(0.01357)**	**(0.00945)**
q	-	0.01	0.01	0.01		
v_1	-	0.0609	0.0609	0.0224	-	-
v_2	-	(0.0134)	(0.0609)	(0.0224)	-	-
	-	0.0474	0.0000	0.0000	-	-

(1) Mauleg (1990)

(2) Market clearing for performance credits

(3) Market clearing for performance credits and same aggregate energy use

Imposing the market clearing condition for performance credits (i.e. outcome (2)) has no impact on the marginal net production cost of either firm. This is because the reduction in the gross unit abatement cost for firm 2, arising from the change in the mix of its production processes, is exactly offset by the increased cost of performance credits. Consequently, aggregate profit (Π) and gross economic benefit (GW) are both unaffected. However, the resultant increase in emissions means that environmental damages are higher than in outcome (1), and hence net economic benefit (NW) is lower.

In outcome (3) the target SEC rate has been reduced to 0.915, so that aggregate emissions under performance credit trading is the same as under the common performance standard, and hence there is no change in environmental damages. This has the effect of raising the marginal net production costs of both firms, although they are still lower than under the common performance standard. Consequently, aggregate output and consumer surplus are lower than in outcome (2), as is aggregate profit. This is reflected in a reduction in gross welfare. However, the lower level of environmental damages means that net welfare is higher.

7.4 Summary

A number of authors have considered how market power can affect the outcome under a market-based implementation mechanism. In all but one of these analyses, the authors have assumed that the regulatory target takes the form of an absolute limit for the aggregate emissions of some pollutant, and hence that the mechanism is a "cap and trade" permit trading system. Only Mauleg (1990) has considered the issue of market power in the context of a relative performance standard – assessing the implications of using a performance-based credit trading scheme to implement a target rate for emissions per unit output when the product market is a homogeneous Cournot oligopoly.

The analyses cover a number of different forms of strategic interaction, with varying degrees of market power. Essentially however, they can be divided into two broad groups: those that focus on the implications of market power in the permit market for the outcome in that market; and those that focus on the implications of using tradable emission permits / credits for the outcome in the output market, when the latter is imperfectly competitive.

In the first group, Hahn (1984) shows that when one firm has price-setting power in the permit market, it will set the price so as to equate its marginal abatement cost to marginal permit revenue (if it is a seller) or marginal permit cost (if it is a buyer). Consequently, unless the initial allocation of permits to the price-setter is efficient (in which case it does not buy or sell any permits), the total cost of achieving the aggregate emissions limit will exceed the minimum level. The key insight provided by Hahn's analysis is that the scale of the efficiency loss depends on the initial permit allocation; increasing as the price-setter's allocation diverges from the efficient level. The analysis is extended to an inter-temporal framework by Hagem & Westskog (199x) under the assumption that there are no restrictions on the banking and borrowing of permits between time periods. They show that while permit trading leads achieves inter-temporal efficiency within firms, it does not usually achieve intra-temporal efficiency between firms.

Misiolek & Elder (1989) consider the situation in which the price-setting firm in the permit market is also dominant in the product market, and can set the price in that market as well. By raising the price of permits, the dominant firm (or predator) increases the marginal production costs of the price-taking Fringe, to its advantage. If the predator would be a seller of permits if it did not engage in this "exclusionary manipulation", then it sets the permit price such that its marginal permit revenue exceeds its marginal abatement cost (i.e. it "under-sells"). If it would be a buyer, then it sets the price such that its marginal permit cost exceeds its marginal abatement cost (i.e. it "over-buys"). In the first case, this exacerbates the efficiency loss arising from its market power in the permit market. However, in the second case it can lead to an improvement in abatement efficiency – even to the attainment of the optimal allocation of abatement effort (given the strategic interaction in the output market).

Finally in the first group, von der Fehr (1993) considers the case where two firms that compete as Cournot duopolists in the product market can also influence the price of emission permits through their respective purchase quantity choices. He concludes that strategic considerations typically strengthen the incentive to purchase permits, with the result that the firms "over-buy" emission permits.[40] By increasing its purchases of permits a firm reduces its own marginal cost of production, making it more aggressive in the output market (a "Top Dog" strategy). Furthermore, if one of the firms can act as a Stackelberg leader in the permit market it has an additional incentive to increase its purchases. As with the case of a dominant firm, by increasing its purchases the leader drives up the price of permits, and hence raises the marginal production cost of its rival. In either case, the "over-buying" of emission permits is likely to lead to an improvement in abatement efficiency,

While all of these analyses were undertaken in the context of an absolute limit for emissions, most of the conclusions carry over to the case of a relative performance standard for emissions per unit output. Since Hahn's analysis relies only on the strict convexity of the conditional production cost function $C(y, L)$, the results apply to the case of a relative standard implemented by a performance-based credit trading scheme with only minor amendments.[41] In particular, the scale of the efficiency loss depends on the values that are set of the performance adjustment factors in the firms' individual performance rules, rather than on the initial allocation of permits. Similarly, von der

[40] Again, the firms "over-buy" in the sense that their respective marginal permit costs exceed their marginal abatement costs.

[41] In the case of relative standard implemented by a performance-based credit trading scheme, the firm's conditional production cost function is given by $C(y, L: r) = \text{Min } \mathbf{c'w}$ s.t. $f(\boldsymbol{w}) \geq y$ and $g(\boldsymbol{w}) - r\, f(\boldsymbol{w}) \leq L$, where $L = v + \delta$ is its holding of credits. The only difference to the definition under the absolute limit arises from the inclusion of $f(\boldsymbol{w})$ in the second constraint., which does not affect the convexity of the cost function. .

Fehr's findings in relation to the impact on the firms' own costs would also be expected to hold in the case of a relative standard.

The translation of the "exclusionary manipulation" findings derived by Misiolek & Elder is not quite so straightforward. Underlying this analyses is the assumption that a firm's marginal cost of production declines as its holding of emission permit increases (i.e. $C_{yL}(y, L) < 0$). Under an absolute emissions limit, sufficient conditions for this to be true are that the firm's production function is homogeneous of degree $k \leq 1$ and its emissions function is linearly homogeneous. However, as was demonstrated in Chapter 5 (section 5.3), these conditions are not sufficient if the regulatory target takes the form of a relative performance standard. In particular, if a firm is already a large seller of credits then a further increase in the number that it sells may lead to a reduction in its marginal cost of production. Consequently, while the findings should carry over to the case of a relative standard if the dominant firm is a seller of credits, they may not do so if it is a buyer. If the predator buys a lot of credits in the absence of exclusionary manipulation, any further over-buying may lead to a reduction in the marginal costs of the Fringe making it more competitive in the product market.

The second group of analyses consider the implications of using tradable emission permits to implement an aggregate emissions limit for the outcome in the product market – assessing the impact on output levels, abatement costs, profits and welfare. The analyses are less general than the first group, in that they assume specific functional forms and (in some cases) particular parameter values. As such, it is not possible to draw any general conclusions from the analyses – other than that the impacts depend on the particular assumptions that are made.

While all of the analyses assume that the product market is a homogeneous Cournot duopoly with linear (inverse) demand, they differ in the assumptions that they make about the firms' cost functions and about the competitive structure of the permit market.

Sartzetakis (1997a) assumes that the regulated pollutant is emitted by several industries, which have similar intra-industry abatement cost distributions. Consequently, while the market for emission permits comprises many firms – all of which act as price-takers, there is intra-industry market clearing. He shows that if the firms differ only in terms of their respective abatement efficiencies, then the trading of emission permits will lead to an increase in aggregate output. While this may lead to an increase in aggregate abatement costs (despite the improvement in abatement efficiency), any such increase is outweighed by the gain in total surplus in the product market, and hence there is an improvement in welfare.

In a separate analysis, Sartzetakis (1997b) considers a situation in which the two firms are the only participants in the market for emission permits, with one of the firm (the leader) having the power to set the price of emission permits. As in the case of the dominant firm analysed by Misiolek & Elder, the firm raises the permit price above the competitive level in order to raise the marginal production cost of its rival. Under the particular functional forms and parameter values adopted by Sartzetakis, this exclusionary manipulation of the permit market has no impact on aggregate output in the product market. While it always benefits the leader, the impact on aggregate profit and welfare may be in either direction, depending on the relative production and abatement efficiencies of the two firms. The leader may also be able to deter a new firm from entering the product market by setting a permit price that is sufficiently high to make entry unprofitable, although this requires that it is more efficient in abatement than the potential entrant. However, while this leads to a reduction in aggregate output and consumer surplus, the rise aggregate profits of the incumbent firms means that overall welfare is higher than it would have been if the new firm had entered the market.

The other three analyses in this group – von der Fehr (1993), Requate (1993), Fershtman & de Zeeuw (1995) – also assume that the two firms are the only participants in the permit market. However, in these analyses, the authors assume that the number of permits traded, and the terms on which they are traded, are determined by a bargaining process between the firms. Only "cash contracts" are allowed, and hence the firms cannot engage in direct collusion. However, since the final distribution of permits affects the equilibrium in the product market, the firms can use the transfer of permits to manipulate the outcome to their mutual advantage.

The analyses differ in the assumptions that they make about the firms' production technologies, and the availability and nature of abatement technologies; with von der Fehr's analysis being the most general. However, they all find that equilibrium industry profit can be maximized by a number of industry structures – including complete monopolisation, depending on the values of the firms' parameters and the stringency of the emissions limit. The impact of permit trading on aggregate output and gross welfare depends on the initial allocations to the two firms, and hence the equilibrium when trading is not allowed. Either, or both, may increase, decrease, or stay the same. Furthermore, it is possible that trading may lead to a reduction in emissions as one the firms acquires all of the permits but does not use them all.

The findings arising from this second group of analyses rely on the adoption of specific functional forms for the firms' cost functions. Since these in turn depend on the specification of the firms' individual performance rules, one would not necessarily expect the results to carry over to the case of a relative performance standard. This is clearly illustrated if one compares the findings of Sartzetakis (1997a) and Mauleg (1990), who both assume that the product market is a homogeneous Cournot duopoly, and that the firms have identical, constant marginal costs and emission rates. While they differ in the assumptions that they make about the firms' abatement cost functions, the main difference between the two analyses relates to the form of the regulatory target – with Sartzetakis assuming an absolute emissions limit, and Mauleg a relative emissions standard. In both cases, aggregate output is higher when the firms are allowed to trade emission permits / credits, causing an increase in consumer surplus but a decline in aggregate profits (after abatement costs). However, whereas the increase in consumer surplus outweighs the reduction in profits under the absolute limit (and hence gross welfare improves), the opposite is true under the relative emissions standard. Furthermore, while trading has no impact on aggregate emissions under the absolute limit, the increase in output causes emissions to rise under the relative standard, exacerbating the reduction in net welfare.

Appendix A7.1 Impact of credit trading on net marginal cost of production

The net cost of production for a firm facing a relative performance standard for emissions per unit output is defined as:

$$C(Y{:}\,r) \quad = \quad \min_{w,v} \quad \sum_{k \in K} c_k\, w_k \; + \; q\, v$$

$$\text{s.t.} \quad f(\boldsymbol{w}) \; - \; Y \; \geq \; 0$$

$$r\,Y \; + \; v \; - \; g(\boldsymbol{w}) \; \geq \; 0$$

$$\boldsymbol{w} \; \geq \; \mathbf{0}$$

where $f(\boldsymbol{w})$ is its production function and $g(\boldsymbol{w})$ is its emissions function. It is assumed that the production function is concave and that the emissions function is convex, and hence that the necessary conditions for a solution are also sufficient.

Denoting the Lagrange multipliers for the production constraint and the emissions constraint by μ and λ respectively, the Kuhn-Tucker conditions for the solution are:

$$c_k \; + \; \lambda\, g_k \; - \; \mu\, f_k \; \leq \; 0 \qquad w_k \; [\; c_k \; + \; \lambda\, g_k \; - \; \mu\, f_k \;] \; = \; 0 \qquad \text{for all } k \in K$$

$$q \; - \; \lambda \; = \; 0$$

From which it follows directly that:

$$\sum_{k \in K} c_k\, w_k \; + \; q \sum_{k \in K} w_k\, g_k \; = \; \mu \sum_{k \in K} w_k\, f_k$$

However, if the production and emission functions are both homogeneous of degree one, then $\sum_{k \in K} w_k\, f_k \; = \; f(\boldsymbol{w})$ and $\sum_{k \in K} w_k\, g_k \; = \; g(\boldsymbol{w})$. In which case:

$$\sum_{k \in K} c_k w_k \quad + \quad q\, g(\boldsymbol{w}) \qquad = \qquad \mu\, f(\boldsymbol{w})$$

$$\sum_{k \in K} c_k w_k \quad + \quad q\,(rY + v) \qquad = \qquad \mu\, Y$$

$$[\sum_{k \in K} c_k w_k \quad + \quad q\, v\,]\ /\ Y \qquad = \qquad \mu \ - \ r\, q$$

But by the envelope theorem, $C_Y(Y; r) = \mu - r\lambda = \mu - r\,q$. Therefore,

$$C(Y; r) / Y \quad = \quad C_Y(Y; r)$$

That is, the firm's net marginal cost of production is equal to its net average cost of production. This has two implications. First, it implies that the net marginal cost of production is constant. Second, since a necessary condition for a firm to trade credits (either as a buyer or as a seller) is that the trade reduces its net cost of production, it implies that for any given level of output, the trading of credits reduces the firm's net marginal cost of production.

Chapter 8 Monopoly power and packaging recovery targets

In Chapter 6 it was shown that performance-based credit trading provides a cost-efficient implementation mechanism for a packaging recovery target under extended producer responsibility (EPR) provided that households are charged the full marginal cost of waste collection. Furthermore, the outcome is independent of the assignment of the initial property rights to the credits that are generated as a result of the transactions between waste collectors and waste reprocessors. However, this conclusion relies on the assumption that all markets in the packaging system are perfectly competitive. In this chapter the implications of relaxing this assumption are investigated, and it is assumed that there is only a single firm in the waste reprocessing sector, which can exercise monopoly (or monopsony) power in any market in which it operates.

The focus of the analysis is rather different to that of the analyses reviewed in the previous chapter, in that it considers the implications of the market power for the design of the mechanism – in particular the assignment of initial property rights that are generated when waste packaging is diverted. To this end, the economic and environmental outcomes are compared under two alternative property rights regimes. Under the first regime, the property rights are assigned to the (price-setting) waste reprocessor. Under the second, they are assigned to the (price-taking) waste collectors.

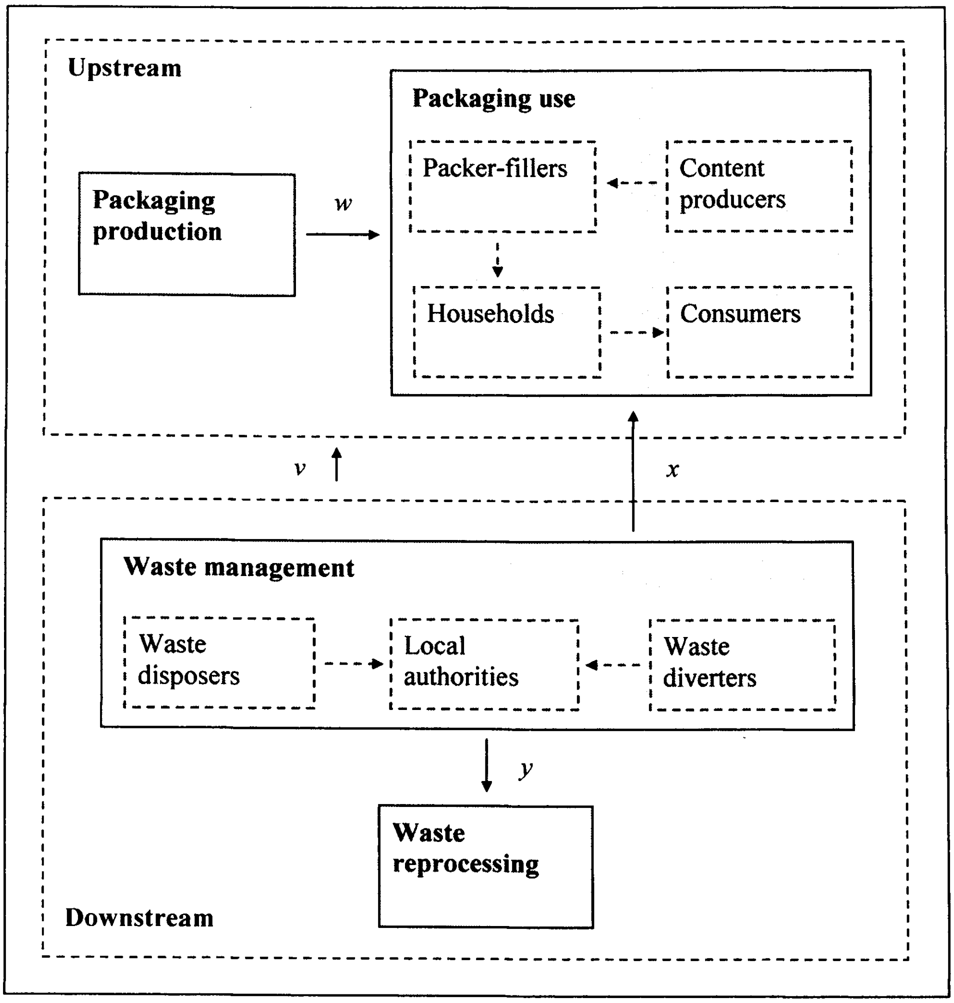
The issue is analysed using a simplified version of the model defined in Chapter 6, in which a number of the component sectors in the packaging system have been amalgamated. As was noted in the introduction to that chapter, the objective of the model is to explore the properties of the trading mechanism, and it is not intended to provide a faithful representation of the actual system. The chapter starts by providing a detailed description of the model. In the next section, the market equilibrium conditions are derived under the alternative property rights regimes, and these are then used to draw some conclusions regarding the relative values of the system variables. Finally, a specific numerical example is used to illustrate the results, and to investigate some issues that cannot be determined analytically.

8.1 Model

The model of the packaging system that underpins the analysis in this chapter has the same basic structure as that used in the analysis of the market equilibrium under perfect competition (see section 6.3.2 in Chapter 6). As can be seen in Figure 8.1 – which provides a schematic representation of the system – a number of the component sectors have been amalgamated, so that the model comprises only four sectors: packaging production, packaging use, waste management, and waste reprocessing.

There is a single reprocessor, which acts as a monopsonist in the market for diverted waste packaging, and as a monopolist in market for performance credits if it is granted the initial property rights. In order to simplify notation it is assumed that there is also only one agent in each of the other three sectors. However, these should be interpreted as representative, or aggregate agents, that are price-takers in any market in which they operate.

Figure 8.1 Simplified packaging system [1]



The quantities of packaging used, waste collection services and diverted waste packaging are denoted by the variables w, x and y respectively; while the quantity of performance credits transferred between the downstream and upstream sectors is denoted by v. All of the variables are non-negative.[2] In order to simplify notation, it is assumed that there is only one packaging material, which has a constant marginal cost

[1] The arrows in Figure 8.1 represent the economic flows of commodities (i.e. products and services) rather than the physical flows of packaging material.

[2] It should be noted that – unlike in previous chapters – the variables w and y represent (positive) quantities of two different commodities, rather than the (negative) input quantity and (positive) output quantity of the same commodity.

of production.[3] Consequently the inverse supply curve for packaging is horizontal, and the market price of packaging is constant (denoted by c).

As in Chapter 6, it is assumed that the marginal cost of waste disposal is constant, and hence so too is the price of the waste collection service.[4] Consequently, the demand for packaging is independent of the price of diverted waste packaging, and the supply of diverted waste packaging is independent of the price of packaging. This allows the two halves of the system to be decoupled, and means that there is no need to consider the market for the waste collection service explicitly. It also implies that in the absence of any regulatory intervention, the market for packaging is completely unaffected by the strategic behaviour in the market for diverted waste packaging.

Again, reduced form inverse supply and demand functions are used for the analysis rather than the underlying production functions. The packaging user's marginal willingness-to-pay (WTP) for packaging is denoted by $b(w)$. The reprocessor's marginal WTP for diverted waste packaging is denoted by $\beta(y)$; while the waste manager's net marginal cost (MC) of diversion is denoted by $\chi(y)$.[5] It is assumed that:

(i)	$b(w)$	$>$	0	for all $w \geq 0$
	$b'(w)$	$<$	0	for all $w \geq 0$
	$b''(w)$	$<$	$-2b'(w) / w$	for all $w \geq 0$

[3] The assumption of a single packaging material is not necessary for the analysis. However, with multiple materials one would have to assume that the same firm reprocesses all of the different materials in order to achieve the desired market structure for performance credits (i.e. a monopoly). Not only is this unlikely to be the case in practice, it adds unnecessary complication to the notation.

[4] Strictly speaking, this is only true if some of the collected waste packaging is sent for disposal. However, unless the diversion rate is 100%, this will be the case by default.

[5] The net marginal cost of diversion is equal to the gross marginal cost, less the avoided marginal cost of waste disposal, which is assumed to be constant.

$$\text{(ii)} \quad \beta(y) > 0 \quad \text{for all } y \geq 0$$

$$\beta'(y) < 0 \quad \text{for all } y \geq 0$$

$$\text{(iii)} \quad \chi(y) \begin{cases} < 0 & \text{for all } y < \hat{y} \\ = 0 & \text{for } y = \hat{y} \\ > 0 & \text{for all } y > \hat{y} \end{cases}$$

$$\chi'(y) > 0 \quad \text{for all } y \geq 0$$

$$\chi''(y) > -2\chi'(y)\,/\,y \quad \text{for all } y \geq 0$$

where $\hat{y} > 0$ is the threshold quantity of diverted waste packaging, below which the gross marginal cost of diversion is less than the (avoided) cost of disposal to landfill. The assumptions regarding the second derivatives of $b(w)$ and $\chi(y)$ ensure that the reprocessor's marginal revenue curve for performance credits is downward sloping, and that it's marginal acquisition cost curve for diverted waste packaging is upward sloping.

The target diversion rate for waste packaging is denoted by $r \in (0,1)$. The performance adjustment factors are all set to zero, and the assignment parameter for packaging is set equal to one so that the obligation to acquire performance credits is placed on the packaging user. Consequently, the packaging producer is inactive in the market for performance credits. The individual performance rules of the other three agents are:

Waste reprocessor: $\theta y - v \geq 0$ ... (8.1.a)

Waste manager: $(1-\theta)y - v \geq 0$ ... (8.1.b)

Packaging user: $v - r w \geq 0$ ... (8.1.c)

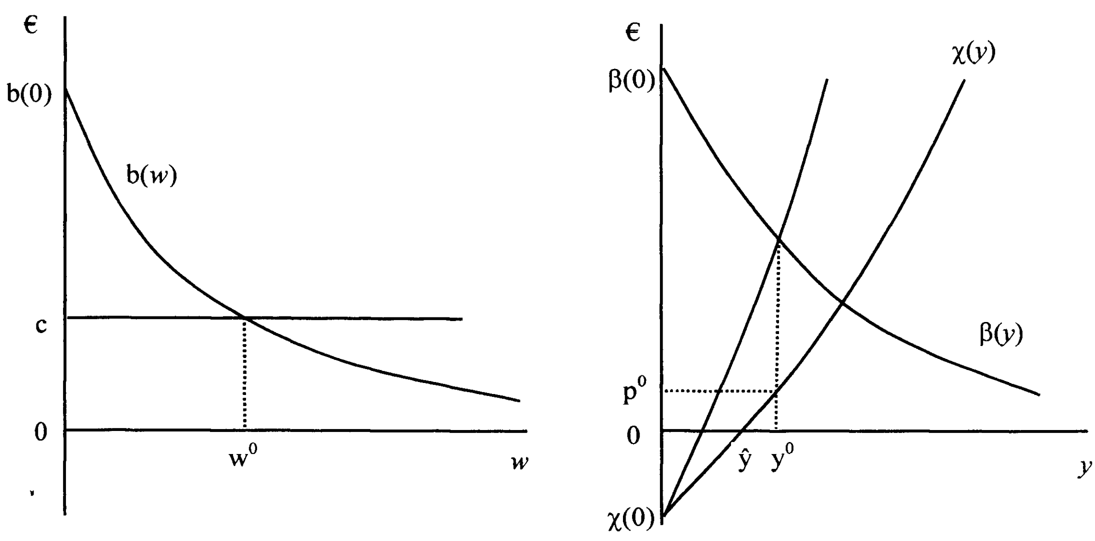
where the assignment parameter θ takes the value 1 or 0, depending on whether the rights to the performance credits that are created when waste packaging is diverted are granted to the reprocessor, or to the waste manager.

As has already been noted, the prices of packaging, and of waste collection services, are both constant (and per-determined). The price of diverted waste packaging (p) is set by reprocessor; as is the price of performance credits (q) if it has been granted the property rights. However, if the rights are assigned to the waste manager, then the price of performance credits is determined as the market clearing price. Under the assumptions that have been made about the net marginal cost of diversion, the price of diverted waste packaging can be negative. This will occur if the reprocessor's WTP curve crosses the marginal acquisition cost curve to the left of the threshold value $\hat{y}$. In this case, the waste manager pays the reprocessor to take the diverted packaging.

It is assumed that prior to the regulatory intervention, the price of diverted waste packaging is strictly positive, and the diversion rate is less than the target rate r (i.e. $p^0 > 0$ and $r\,w^0 > y^0 > 0$). As Figure 8.2 illustrates, this implies that the choke price for packaging is greater than the constant marginal cost of production (i.e. $b(0) > c$), and that the reprocessor's inverse demand curve for diverted waste packaging crosses the marginal acquisition cost to the right of the threshold quantity $\hat{y}$.

Figure 8.2 Pre-regulation market equilibria

a) Packaging ***b) Diverted waste packaging***



8.2 Market equilibrium

Using the model outlined in the previous section, the necessary and sufficient conditions for a joint market equilibrium are derived for the alternative property right regimes (i.e. for $\theta = 1$ and $\theta = 0$). A comparison of these conditions allows conclusions to be drawn regarding the relative values of the real variables and endogenous market prices under the two regimes.

8.2.1 Property rights assigned to the reprocessing sector ($\theta = 1$)

The joint market equilibrium is modelled as a two-stage game. In the first stage the reprocessor sets the prices of diverted waste packaging and performance credits, taking account of the waste manager's supply function and the packaging user's demand function respectively. In the second stage, the waste manager chooses the quantity of waste packaging to divert, and the packaging user chooses the quantity of performance credits to acquire; each taking the respective prices as given.

Stage 2: Derivation of supply and demand functions

As usual, the game is solved by backwards induction starting with the second stage; which comprises the following independent decision problems for the waste manager and the packaging user respectively.

$$\max_{y} \quad \mathrm{p}\,y \;-\; \int_0^y \chi(u)\,\mathrm{d}u \qquad \ldots (8.2.\mathrm{a})$$

$$\max_{w,v} \quad \int_0^w \mathrm{b}(u)\,\mathrm{d}u \;-\; \mathrm{c}\,w \;-\; \mathrm{q}\,v \qquad \text{s.t.} \qquad \mathrm{r}\,w \;-\; v \;\leq\; 0 \qquad \ldots (8.2.\mathrm{b})$$

The Kuhn-Tucker first order conditions for the two problems are:

$$p - \chi(y) \leq 0 \qquad y \geq 0 \qquad y\left[p - \chi(y)\right] = 0$$

$$b(w) - c - r\lambda \leq 0 \qquad w \geq 0 \qquad w\left[b(w) - c - r\lambda\right] = 0$$

$$\lambda - q \leq 0 \qquad v \geq 0 \qquad v\left[\lambda - q\right] = 0$$

$$rw - v \leq 0 \qquad \lambda \geq 0 \qquad \lambda\left[rw - v\right] = 0$$

where λ is the Lagrange multiplier for the performance rule constraint in the packaging user's optimization problem. Under the assumptions that have been made about the properties of $\chi(y)$ and $b(w)$, the objective functions in (8.2.a) and (8.2.b) are both strictly concave, and hence these conditions are both necessary and sufficient for the solutions to the respective problems. Assuming non-zero solutions[6], they can be solved explicitly to yield the following three identities:

$$y(p) \equiv \chi^{-1}(p) \qquad \text{... (8.3.a)}$$

$$v(q; c, r) \equiv r\, b^{-1}(rq + c) \qquad \text{... (8.3.b)}$$

$$w(c; q, r) \equiv b^{-1}(rq + c) \qquad \text{... (8.3.c)}$$

These represent respectively the waste manager's supply function for diverted waste packaging, and the packaging user's demand functions for performance credits and packaging. Thus a change in the value of the diversion target (r) affects the two demand curves, but has no impact on the supply curve for diverted waste packaging. It should be noted that $v(0; c, r) \equiv r\, b^{-1}(c) = r\, w^0$. That is, the packaging user's marginal willingness to pay for performance credits falls to zero once it has acquired sufficient to satisfy its obligation without reducing its packaging use from the pre-regulation level.

[6] Necessary and sufficient conditions for non-zero solutions are that $p > \chi(0)$, and $q < (\, b(0) - c\,) / r$.

Furthermore, the maximum price that the packaging user is willing to pay for performance credits is $q^{max} = (b(0) - c) / r$.

Since (8.3.a) and (8.3.b) are identities, they can be differentiated to give:

$$1 / y' \quad = \quad \chi'(y(p)) \quad > \quad 0 \qquad \ldots (8.4.a)$$

$$1 / v' \quad = \quad b'(v(q)/r) / r^2 \quad < \quad 0 \qquad \ldots (8.4.b)$$

Thus, the supply curve for diverted waste packaging is upward sloping, and the demand curve for packaging is downward sloping, as one would expect.

Stage 1: Setting the prices

In the first stage, it is instructive (but not necessary) to break down the optimization problem for the reprocessor into two steps – with the first step relating to the determination of the price of performance credits, and the second step to the price of diverted waste packaging. It is assumed that the waste reprocessor knows the functional forms of the respective demand and supply curves (i.e. (8.3.a) and (8.3.b)).

step 1: $$\max_{q} \quad \pi(q) \;+\; q\,v(q) \qquad \ldots (8.5.a)$$

step 2: $$\pi(q) \quad = \quad \max_{p} \quad \int_0^{y(p)} \beta(u)\,du \;-\; p\,y(p) \qquad \ldots (8.5.b)$$

$$\text{s.t.} \quad v(q) \;-\; y(p) \;\leq\; 0$$

The necessary first order conditions for step 2 are:

$$\beta(y(p))\,y'(p) \;-\; y(p) \;-\; p\,y'(p) \;+\; \mu\,y'(p) \;=\; 0$$

$$v(q) \;-\; y(p) \;\leq\; 0 \qquad \mu \geq 0 \qquad \mu\,[\,v(q) \;-\; y(p)\,] \;=\; 0$$

where μ is the Lagrange multiplier for the constraint. Denoting the solution values by $p = p(q)$ and $\mu = \mu(q)$, and noting that by the envelope theorem $\pi'(q) = -\mu\, v'(q)$, it follows that:

$$\left[p + \frac{y(p)}{y'(p)} \right] - \beta(y(p)) = \mu = -\frac{\pi'(q)}{v'(q)} = -\pi'(q)\, q'(v(q)) \qquad \ldots (8.6)$$

$$\text{where} \quad \mu = 0 \quad \text{if} \quad v(q) \leq v^0 = y^0 = y(p^0)$$
$$> 0 \quad \text{if} \quad v(q) > v^0 = y^0 = y(p^0)$$

Thus, for a given choice of price for performance credits (q), the reprocessor sets the price for diverted waste packaging (p) such that the marginal acquisition cost exceeds the marginal benefit by an amount equal to the shadow value of its individual performance rule constraint. This value can be interpreted as the marginal cost of producing the quantity performance credits demanded by the packaging user at this price (i.e. $v \equiv v(q)$). If the price of performance credits is sufficiently high that the quantity demanded is less than the maximum number that can be created at the pre-regulation level of diversion (i.e $v^0 = y^0$), then the shadow value is zero and the price of diverted waste packaging is unchanged from its pre-regulation value. It should be noted that while there is no restriction on the sign of p, it must be the case that $p > \chi(0)$.[7]

The necessary first order condition for step 1 of the reprocessor's problem is:

$$\pi'(q) + q\, v'(q) + v(q) = 0$$

Denoting the solution value by q^*, and recalling the envelope condition from step 2, it follows that:

[7] If $v(q) = 0$, then the constraint becomes redundant, and the solution is given by the pre-regulation price (i.e. $p(q) = p^0$), which is strictly positive by assumption. If $v(q) > 0$, then the value of p(q) must be such that $y(p(q)) > 0$, which requires that $p(q) > \chi(0)$.

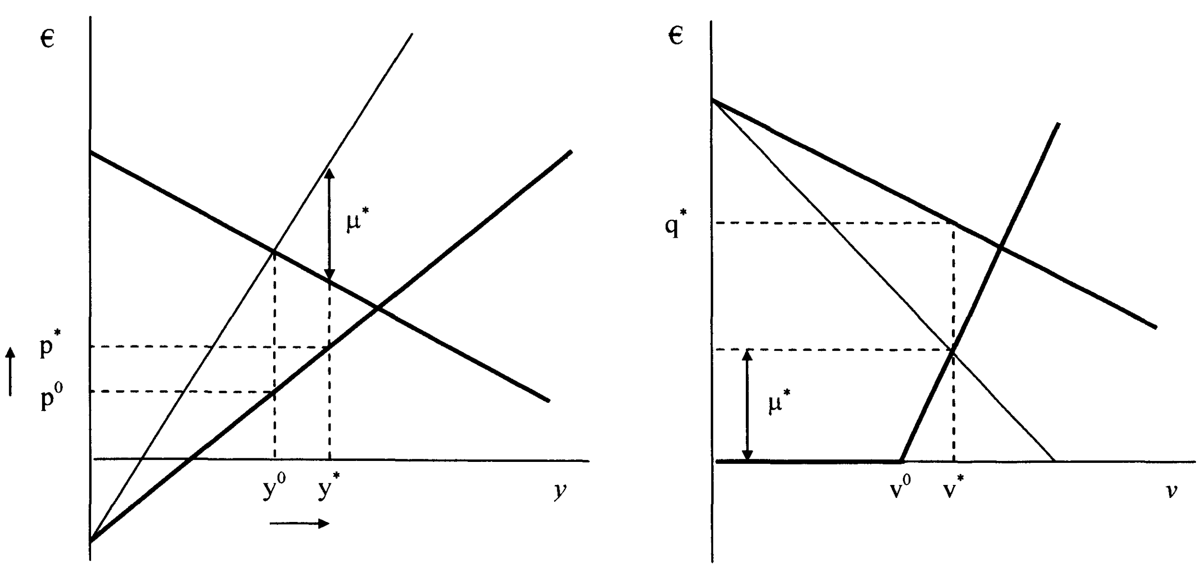
$$\left[q^* + \frac{v(q^*)}{v'(q^*)} \right] \quad = \quad -\frac{\pi'(q^*)}{v'(q^*)} \quad = \quad \mu^* \quad = \quad \mu(q^*) \qquad \ldots (8.7)$$

That is, the marginal cost of producing performance credits is equal to the marginal revenue resulting from their sale. This of course is the usual profit maximization condition for a monopolist.

Figure 8.3 illustrates the market equilibria for diverted waste packaging and for performance credits, where $v^* = v(q^*)$, $p^* = p(q^*)$ and $y^* = y(p^*)$. From this, it is clear that the price and quantity of diverted waste packaging cannot be lower than their respective pre-regulation values p^0 and y^0, and that both are strictly greater whenever the reprocessor's performance rule constraint is binding (i.e. $\mu^* > 0$).

Figure 8.3 Marginal cost of producing performance credits

a) Diverted waste packaging **b) Performance credits**



In Figure 8.3 the shadow value of the constraint is greater than zero, reflecting the fact that the reprocessor's marginal revenue curve for performance credits crosses its marginal cost curve to the right of the threshold quantity v^0. However, this need not be the case. In particular, it can be shown that there is a threshold value for the target diversion rate; below which the curves cross to the left of v^0, and hence the shadow value is zero. This value will depend on the functional form of $v(q; c, r)$, which in turn depends on the functional form of $b(w)$, but it is always strictly greater than the pre-regulation diversion rate (r^0).

<u>Proposition 8.1</u>

For any functional form of $b(w)$ that satisfies the assumptions set out in section 8.1, there exists a threshold value $\check{r} > r^0$ such that $v^* < v^0 = y^0 = y^*$ for all values of $r < \check{r}$.

Proof: see Appendix A8.1

<u>Corollary 8.1</u>

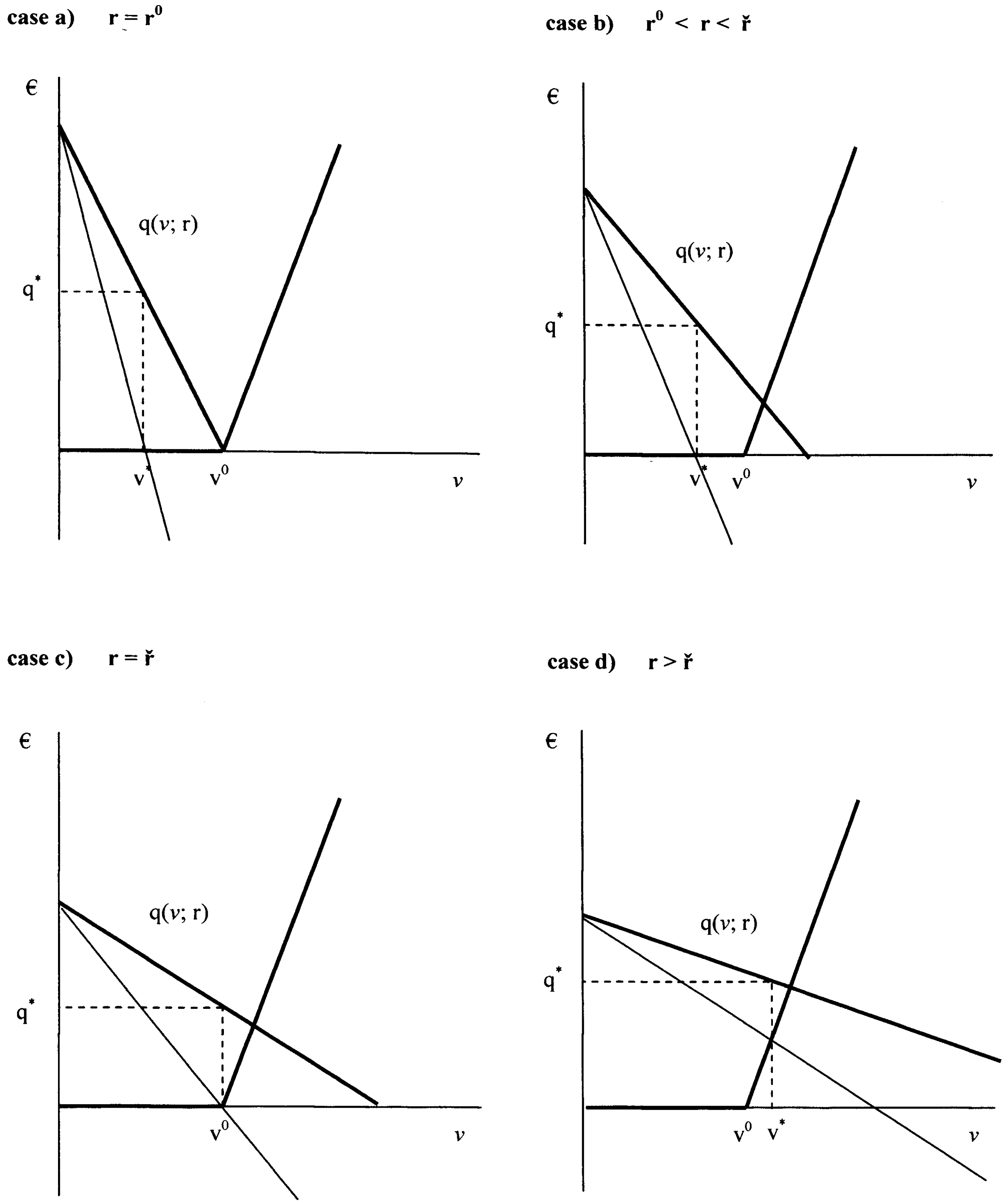
If the target diversion rate is less than the threshold rate ($\check{r}$), then the actual diversion rate is equal to the threshold rate, i.e.

$$\check{r} = \frac{w^*}{y^*} > \frac{v^*}{y^*} = r \qquad \text{for all } r < \check{r}$$

Proof: see appendix A8.2

Proposition 8.1 is illustrated in Figure 8.4 for the case of a linear inverse demand function for performance credits (denoted by $q(v{:}r)$). By definition, when $r = r^0$ the packaging user's willingness-to-pay for credits is equal to zero for any quantity of credits greater than or equal to v^0. Since the marginal revenue curve lies strictly below the inverse demand curve, it follows that it must intersect the marginal cost curve to the left of v^0. As the target rate increases, the intercept and the slope both decline (in magnitude); causing the inverse demand and marginal revenue curves to rotate anti-clockwise, and hence the equilibrium quantity of performance credits transferred to increase. However, initially this has no impact on the marginal cost, which remains equal to zero. It only when the target rate rises above the threshold $\check{r}$ that the curves intersect to the right of v^0, and hence that the marginal cost is positive.

Figure 8.4 Market for performance credits



Thus, when the target diversion rate is less than the threshold rate, the reprocessor can divorce its two price decisions. It can continue to set the pre-regulation monopsony price for diverted waste packaging, while setting a monopoly price for performance credits based on a zero marginal cost of production. This has two interesting consequences. First, since there is no increase in the quantity of waste packaging diverted, the target diversion rate is achieved solely through a reduction in the amount of packaging used. Second, Corollary 8.1 implies that the aggregate performance rule is not binding in the equilibrium, despite the positive price of performance credits. This, of course, could not occur if the markets were perfectly competitive.

In contrast, when the target diversion rate exceeds the threshold value, the reprocessor must make a joint decision; balancing the loss of profit from an (incremental) increase in the quantity of diverted waste packaging that it accepts against the additional revenue that it receives from selling the resultant performance credit. In this case, the target rate is achieved by a combination of source reduction and increased diversion, and the aggregate performance rule is binding.

Combining conditions (8.3), (8.4), (8.6) and (8.7) yields the following sets of necessary market equilibrium conditions:

a) target diversion rate below the threshold value:

$$\beta(y^*) = p^* + y^* \chi'(y^*) \qquad \ldots (8.8.a)$$

$$0 = q^* + v^* b'(v^*/r) / r^2 \qquad \ldots (8.8.b)$$

$$y^* = \chi^{-1}(p^*) \qquad \ldots (8.8.c)$$

$$v^* = r\, b^{-1}(r\, q^* + c) \qquad \ldots (8.8.d)$$

$$r' w^* = v^* < y^* \qquad \ldots (8.8.e)$$

b) target diversion rate greater than or equal to the threshold value:

$$\beta(y^*) = [p^* + y^* \chi'(y^*)] - [q^* + v^* b'(v^*/r) / r^2] \quad \text{... (8.9.a)}$$

$$y^* = \chi^{-1}(p^*) \quad \text{... (8.9.b)}$$

$$v^* = r\, b^{-1}(r\, q^* + c) \quad \text{... (8.9.c)}$$

$$r\, w^* = v^* = y^* \quad \text{... (8.9.d)}$$

8.2.2 Property rights assigned to the waste management sector ($\theta = 0$)

Again, the market equilibrium is modelled as a two stage game. In the first stage the reprocessor sets the price of diverted waste packaging, taking account of the waste manager's supply function. In the second stage, the waste manager chooses the quantity level of waste packaging to divert and the quantity of performance credits to sell, and the packaging user chooses the quantity of performance credits to acquire; each taking the respective price as given.

Stage 2: Derivation of supply and demand functions

In the second stage, the optimization problem for packaging user remains unchanged, and hence so too do identities (8.3.b) and (8.3.c). However, the problem for the waste manager is now:

$$\underset{y,v}{\text{Max}} \quad p\,y - \int_0^y \chi(u)\,du + q\,v \qquad \text{s.t.} \qquad v - y \le 0 \quad \text{... (8.10)}$$

The Kuhn-Tucker first order conditions are for this problem are:

$$p + \lambda - \chi(y) \le 0 \qquad y \ge 0 \qquad y\,[p + \lambda - \chi(y)] = 0$$

$$q - \lambda \le 0 \qquad \nu \ge 0 \qquad \nu\left[q - \lambda\right] = 0$$

$$\nu - y \le 0 \qquad \lambda \ge 0 \qquad \lambda\left[\nu - y\right] = 0$$

where λ is the Lagrange multiplier for the performance rule constraint. As in the previous scenario, these conditions are both necessary and sufficient. Assuming a non-zero solution, they can be solved explicitly to yield the following two identities:

$$\tilde{y}(p, q) \equiv \chi^{-1}(p + q) \qquad \ldots (8.11.a)$$

$$\tilde{v}(p, q) \equiv \chi^{-1}(p + q) \qquad \ldots (8.11.b)$$

These represent respectively the waste manager's supply functions for diverted waste packaging and performance credits. Since these are identities, it follows that:

$$1 / \tilde{y}_p(p, q) = 1 / \tilde{y}_q(p, q) = \chi'(\tilde{y}(p, q)) > 0 \qquad \ldots (8.12.a)$$

$$1 / \tilde{v}_p(p, q) = 1 / \tilde{v}_q(p, q) = \chi'(\tilde{y}(p, q)) > 0 \qquad \ldots (8.12.b)$$

Hence, for both supply functions, the impact of a change to the price of diverted waste packaging, and of a change to the price of performance credits, are symmetric.

The market clearing condition for performance credits is:

$$\tilde{v}(p, q) - v(q; c, r) \ge 0 \qquad q \ge 0 \qquad q\,[\tilde{v}(p, q) - v(q; c, r)] = 0$$

Assuming that the equilibrium price of performance credits is strictly positive, it is therefore defined implicitly by the equation:

$$g(p, q) \equiv r\, b^{-1}(r\, q + c) - \chi^{-1}(p + q) = 0$$

from which it is clear that the price is affected by the reprocessor's choice of price for diverted waste packaging, as well as by the stringency of the diversion target.

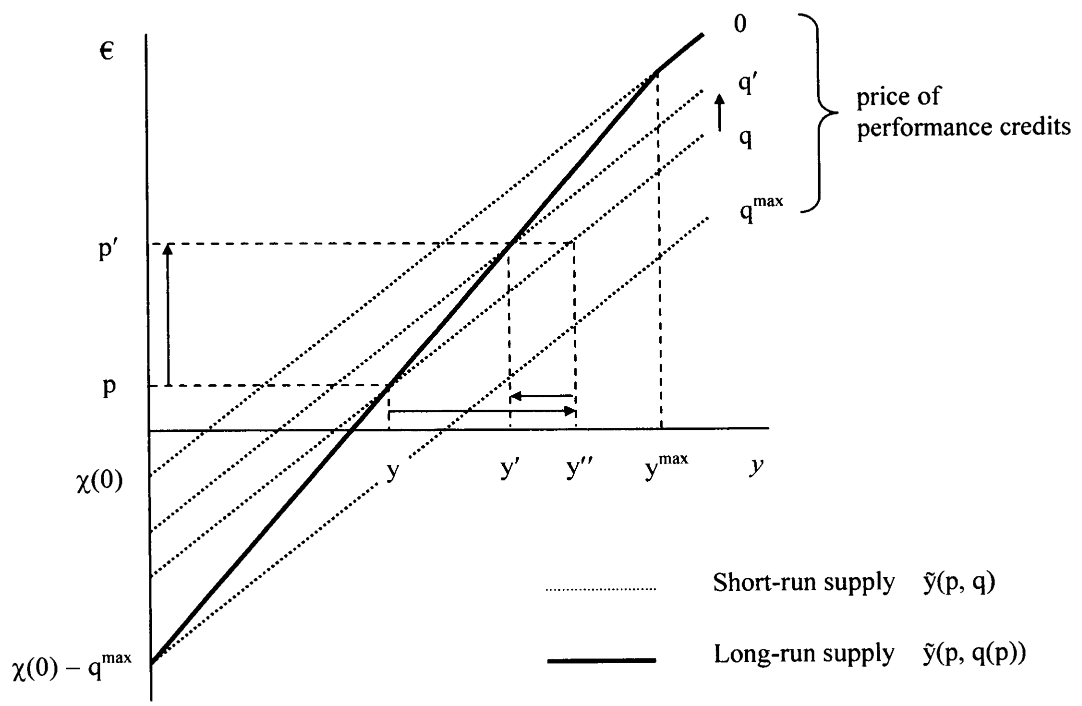
By the implicit function theorem:

$$q'(p) \quad = \quad -\frac{\partial g}{\partial p}\bigg/\frac{\partial g}{\partial q} \quad = \quad \frac{\chi^{-1/}}{r^2\, b^{-1/} \; - \; \chi^{-1/}}$$

$$= \quad \frac{1}{r^2\left(\dfrac{\chi'(\cdot)}{b'(\cdot)}\right) - 1} \quad < \quad 0 \qquad \ldots (8.13)$$

Thus an increase in the price of diverted waste packaging causes the equilibrium price of performance credits to fall. Since the supply of diverted waste packaging is a function of both prices (see 8.11.a), it follows that an increase in the price of diverted waste packaging (p) affects the quantity supplied in two ways: directly, through the own-price effect; and indirectly, via the impact on the price of performance credits. This is illustrated in Figure 8.5

Figure 8.5 Supply of diverted waste packaging



The direct (or short-run) effect of a rise in the price of diverted waste packaging (from p to p′) is to increase the quantity supplied from y to y″. However, the resultant increase in the supply of performance credits causes the price to fall from q to q′. This shifts the supply curve upwards, reducing the quantity supplied from y″ to y′. Thus, the long-run inverse supply curve (shown in bold) is steeper than the short-run curves (which are all parallel).

Stage 1: Setting the price

The reprocessor faces the unconstrained optimization problem:

$$\max_{p} \int_{0}^{\tilde{y}(p,q)} \beta(u)\, du \; - \; p\, \tilde{y}(p, q) \qquad \ldots (8.14)$$

As in the previous case, it is assumed that the reprocessor knows the functional form of the short-run supply function for diverted waste packaging (8.11.a). That is, it understands that the inverse supply curve is shifted down (from its pre-regulation position) by an amount equal to the equilibrium price of performance credits. However, it is also necessary to make an assumption about the reprocessor's expectation (or belief) regarding the price of credits; in particular, how this is affected by the price that it sets for diverted waste packaging.

Two alternative cases are considered. In the first case (termed *complete knowledge*), it is assumed that the reprocessor correctly forecasts the relationship between the price of diverted waste packaging and the price of performance credits (i.e. q = q(p)). That is, the reprocessor realises that even though the waste manager cannot affect the price of performance credits (since it is a price-taker), it can do so by manipulating the quantity of waste packaging that is diverted, and hence the total quantity of credits supplied.

In the second case (termed *partial knowledge*), it is assumed that the reprocessor correctly forecasts the equilibrium market price of performance credits (i.e. $q = q^{\Delta}$), but it does not realise that it can influence this price through its actions in the diverted waste packaging market.

The relative merits of the alternative assumptions are discussed in section 8.4. However, it should be noted that in both cases, the reprocessor's expectation is rational in the sense that the resultant outcome is consistent with its prior belief – i.e. there is no disconfirming evidence to contradict the hypothesis.

Denoting the solution when the reprocessor has complete knowledge by $p^{\#}$, and the solution when it has partial knowledge by p^{Δ}, the respective necessary first-order conditions are:

$$\left[p^{\#} + \frac{\tilde{y}(p^{\#},q^{\#})}{\tilde{y}_p(p^{\#},q^{\#}) + \tilde{y}_q(p^{\#},q^{\#})\, q'(p^{\#})} \right] = \beta(\tilde{y}\,(p^{\#}, q^{\#})) \qquad \ldots (8.15.a)$$

$$\left[p^{\Delta} + \frac{\tilde{y}(p^{\Delta},q^{\Delta}))}{\tilde{y}_p(p^{\Delta},q^{\Delta})} \right] = \beta(\tilde{y}\,(p^{\Delta}, q^{\Delta})) \qquad \ldots (8.15.b)$$

where $q^{\#} = q(p^{\#})$ and $q^{\Delta} = q(p^{\Delta})$

In each case, the reprocessor equates the marginal acquisition cost of waste packaging with the marginal benefit. However, in the first case the marginal acquisition cost is calculated using the long-run supply curve, while in the second case it is derived from the short-run supply curve (at the equilibrium price of credits q^{Δ}).

Combining conditions (8.3.b), (8.11.a), (8.12), (8.13) and (8.15) yields the following sets of necessary conditions for the market equilibrium:

a) complete knowledge

$$\beta(y^{\#}) = p^{\#} + y^{\#} [\chi'(y^{\#}) - b'(v^{\#}/r) / r^2] \qquad \ldots (8.16.a)$$

$$y^{\#} = \chi^{-1}(p^{\#} + q^{\#}) \qquad \ldots (8.16.b)$$

$$v^{\#} = r\, b^{-1}(r\, q^{\#} + c) \qquad \ldots (8.16.c)$$

$$r\, w^{\#} = v^{\#} = y^{\#} \qquad \ldots (8.16.d)$$

b) partial knowledge

$$\beta(y^{\Delta}) = p^{\Delta} + y^{\Delta}\, \chi'(y^{\Delta}) \qquad \ldots (8.17.a)$$

$$y^{\Delta} = \chi^{-1}(p^{\Delta} + q^{\Delta}) \qquad \ldots (8.17.b)$$

$$v^{\Delta} = r\, b^{-1}(r\, q^{\Delta} + c) \qquad \ldots (8.17.c)$$

$$r\, w^{\Delta} = v^{\Delta} = y^{\Delta} \qquad \ldots (8.17.d)$$

8.2.3 Comparison of market equilibria

Comparison of the alternative market equilibrium conditions (8.8), (8.9), (8.16) and (8.17) yields the following two propositions.

Proposition 8.2

If the reprocessor has complete knowledge, and the target diversion rate is greater than the threshold value (ř), then the equilibrium quantities of all real system variables are unaffected by the allocation of the initial property rights to the performance credits; as is the price of performance credits. However, if the target rate is less than the threshold value, then the outcomes differ. In this case, if the initial property rights are granted to the waste manager (i.e. $\theta = 0$), then the:

- price of performance credits is lower (i.e. $q^{\#} < q^{*}$)
- quantity of diverted waste packaging is lower (i.e. $y^{\#} < y^{*}$)
- quantity of performance credits is higher (i.e. $v^{\#} > v^{*}$)
- quantity of packaging used is higher (i.e. $w^{\#} > w^{*}$)
- quantity of waste packaging sent to landfill is higher (i.e. $w^{\#} - y^{\#} > w^{*} - y^{*}$)

Proof: see Appendix A8.3

Proposition 8.3:

If the initial property rights to the performance credits are granted to the waste manager then, for all values of the target diversion rate, the outcomes will differ depending on the assumption that is made regarding the knowledge of the reprocessor regarding the price of performance credits. If the reprocessor is has partial knowledge (i.e. it assumes that $q = q^{\Delta}$), then the:

- price of performance credits is lower (i.e. $q^{\Delta} < q^{\#}$)
- price of diverted waste packaging is higher (i.e. $p^{\Delta} > p^{\#}$)
- quantity of diverted waste packaging is higher (i.e. $y^{\Delta} > y^{\#}$)
- quantity of performance credits is higher (i.e. $v^{\Delta} > v^{\#}$)
- quantity of packaging used is higher (i.e. $w^{\Delta} > w^{\#}$)
- quantity of waste packaging sent to landfill is higher (i.e. $w^{\Delta} - y^{\Delta} > w^{\#} - y^{\#}$)

Proof: see appendix A8.4

Thus, it is clear that the differences between the outcomes under the two alternative property rights regimes depends critically on the assumption that is made regarding the knowledge of the reprocessor, and on the value of the target diversion rate. If the reprocessor has complete knowledge, and the target rate is above the threshold value ($\check{r}$), then it makes no difference whether the initial property rights are granted to the reprocessor or to the waste manager. The outcome, and the distribution of costs and benefits, is exactly the same in each case. However, if the reprocessor has partial knowledge, or if target rate is below the threshold level then the outcomes will depend on which sector is granted the initial property rights.

Table 8.1 Relative values of system variables

	Below threshold (i.e. $r < \check{r}$)	Above threshold ($r \geq \check{r}$)
Price of:		
- performance credits	$q^{\Delta} < q^{\#} < q^{*}$	$q^{\Delta} < q^{\#} = q^{*}$
- diverted waste packaging	$p^{\Delta} > p^{\#} < p^{*}$	$p^{\Delta} > p^{\#} = p^{*}$
Quantity of:		
- diverted waste packaging	$y^{\Delta} > y^{*} > y^{\#}$	$y^{\Delta} > y^{*} = y^{\#}$
- performance credits	$v^{\Delta} > v^{\#} > v^{*}$	$v^{\Delta} > v^{\#} = v^{*}$
- packaging used	$w^{\Delta} > w^{\#} > w^{*}$	$w^{\Delta} > w^{\#} = w^{*}$
- landfilled waste packaging [(a)]	$l^{\Delta} > l^{\#} > l^{*}$	$l^{\Delta} > l^{\#} = l^{*}$

(a) By definition $l \equiv w - y$

By combining propositions 8.2 and 8.3 it is possible to compare the equilibrium values of the various model variables for the alternative combinations of property right assignments and knowledge assumptions. These rankings are shown in Table 8.1, from which it can be seen that the environmental outcome is always worse (or at best no better) when the initial property rights are granted to the waste manager. Under either knowledge assumption, and for any diversion rate, the amount of packaging used by the producer, and the amount of waste packaging sent to landfill, is greatest when the waste manager has the property rights.

8.3 Illustrative example

In order to gain a clearer understanding of the relative outcomes under the different property rights regimes, the equilibrium values of the endogenous system variables (prices and quantities) will be derived for an illustrative example, using the specific functional forms and parameter values given in Table 8.2.

Table 8.2 Parameter values and functional forms

Exogenous price of packaging	c	=	2
Inverse demand for packaging	$b(w)$	=	$7 - w$
Inverse supply for diverted waste packaging	$\chi(y)$	=	$2y - 1$
Inverse demand for diverted waste packaging	$\beta(y)$	=	4

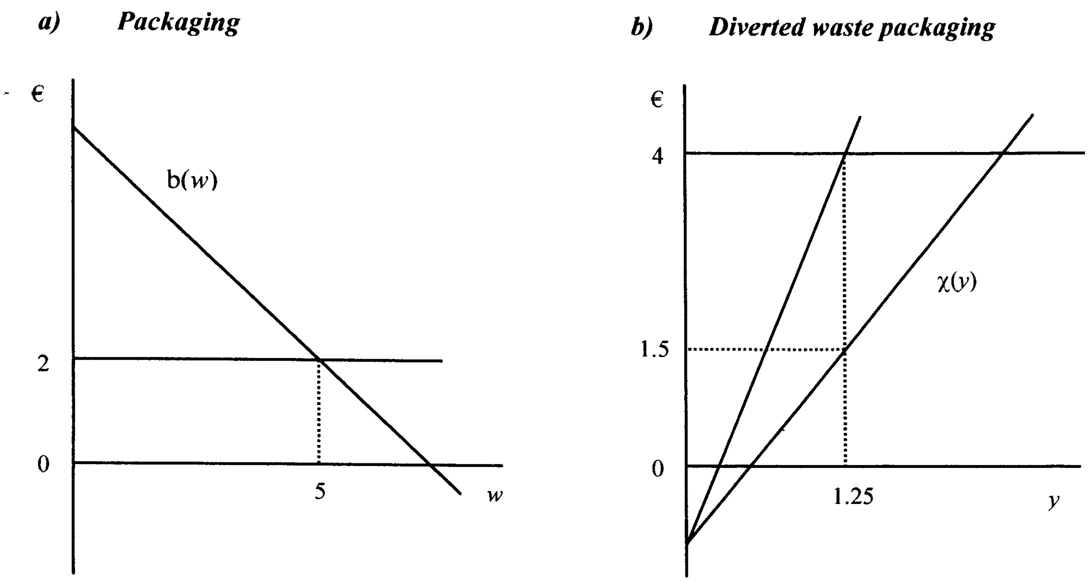
Thus the inverse demand for packaging and the inverse supply of diverted waste packaging are both linear, while the reprocessor's willingness to pay for diverted waste packaging is horizontal. The resultant pre-regulation market equilibria are shown in Figure 8.6. The price of diverted waste packaging is € 1.5 per tonne, and the diversion rate is 25% (i.e. 1.25 / 5).

It follows from the assumption of a linear demand curve for packaging material that the inverse demand curve for performance credits and the marginal revenue curve are both linear in v, respectively taking the forms:

$$q(v; r) = \frac{5}{r} - \frac{v}{r^2} \quad \text{and} \quad m(v; r) = \frac{5}{r} - \frac{2v}{r^2}$$

Thus, as the target rate r increases, the curves rotate anti-clockwise, becoming flatter at higher target diversion rates. Setting m(1.25; r) = 0 yields a value for the threshold diversion rate (ř) of 50%, below which the waste reprocessor can divorce its decisions when it has the property rights.

Figure 8.6 Pre-regulation market equilibria



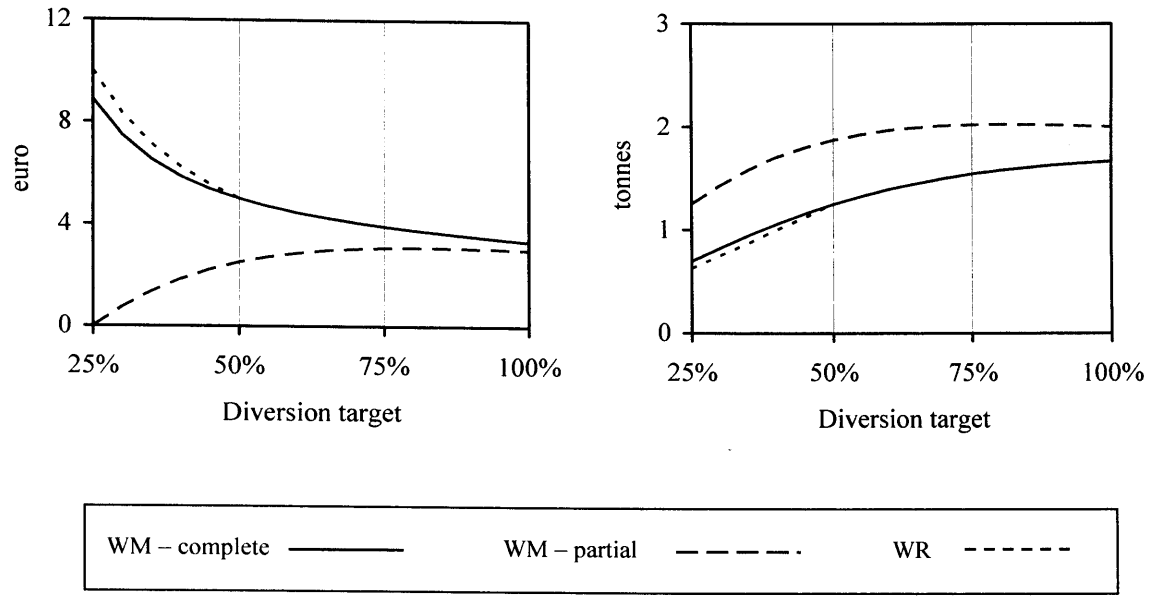
It is straightforward to derive algebraic expressions for the real variables (w, y and v) and the endogenous prices (p and q) in terms of the target level of diversion (r).[8] Figures 8.7 – 8.10 show the market equilibrium values of these variables for all values of the target diversion rate above the pre-regulation value. As can be seen, for any given diversion rate, the relative values of the variables under the different property rights regimes and knowledge assumptions are the same as those predicted in Table 8.1. While this is reassuring, it is not of particular interest. Of greater interest is the impact

[8] These are given in Appendix A8.5.

of increases in the target diversion rate, and therefore the discussion in this section focuses on this aspect.[9]

Figure 8.7 Performance credits

a) price (q) ***b) quantity (v)***



If the property rights are granted to the waste manager and the reprocessor has partial knowledge, then the price rises. In contrast, if the rights are granted to the reprocessor (or if the waste manager has the rights and the reprocessor has complete knowledge), then the price of performance credits declines as the target diversion rate increases. The negative relationship between the price of credits and the target diversion rate is not a consequence of market power *per se*. As was noted in Chapter 6, it can arise when the inverse supply curve is steep relative to the inverse demand. However, since the inverse

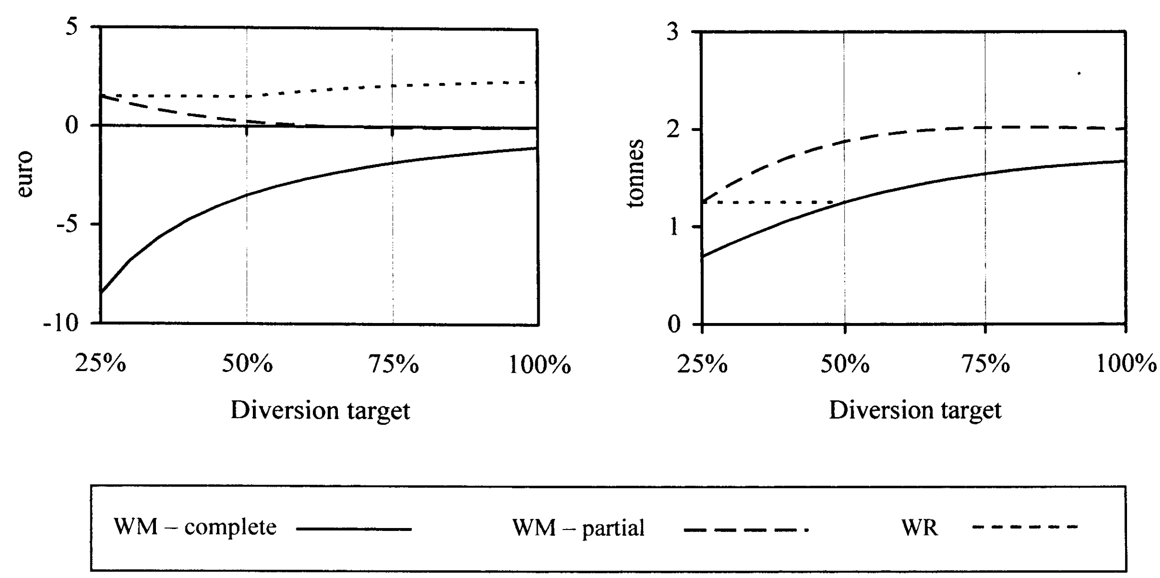
[9] It should be noted that the functional forms have been chosen so as to facilitate the derivation of an analytical solution to of the market equilibrium conditions. Consequently, the results should be taken as illustrative of what can happen, rather than as general predictions of what will happen.

supply curve for performance credits is steeper when the property rights are granted to the price-setting reprocessor, it is more likely that the price will decline.[10]

When the reprocessor is granted the property rights, the introduction of the regulation has no impact on the market for diverted waste packaging if the target diversion rate is below the 50% threshold. This reflects the fact that the reprocessor can divorce its decisions in the two markets, and has no reason to deviate from its pre-regulation intake of diverted waste. However, as the target rate increases above the threshold, the reprocessor loses this freedom of action, and must increase the price of diverted waste packaging (and hence the quantity supplied) if it wishes to increase the number of performance credits that it can sell.

Figure 8.8 Diverted waste packaging

a) price (p) *b) quantity (y)*



[10] It is steeper because it is equal to the difference between the reprocessor's marginal cost of acquisition for diverted waste packaging and its inverse demand, rather than the difference between the inverse supply and demand (see Figure 8.3 and Figure 8.6).

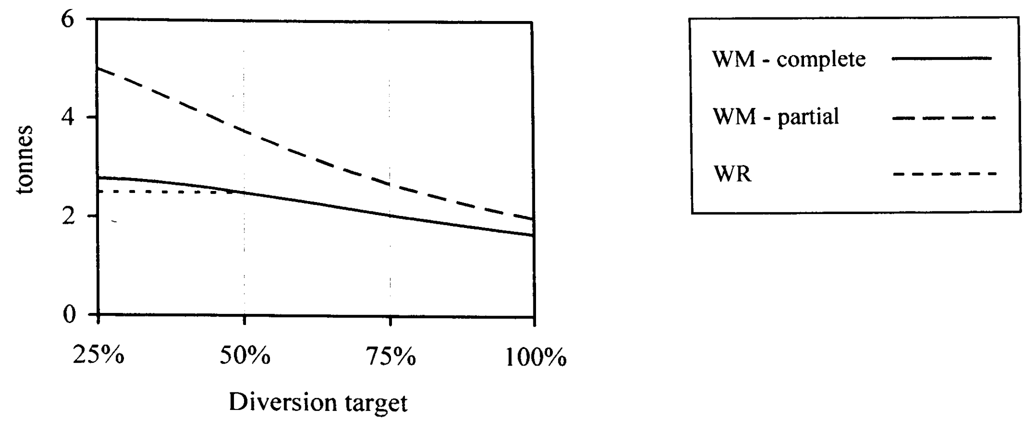
When the rights are granted to the waste manager, the impact is very different under the alternative assumptions regarding the reprocessor's knowledge. If the reprocessor has partial knowledge, then as the target becomes more stringent, the quantity of diverted waste rises from its pre-regulation level, while the price falls. This is in marked contrast to the impact if the reprocessor has complete knowledge. In this case it realises that, rather than pay the waste manager for diverted waste packaging, it can charge to receive it (i.e. set a negative price). The waste manager is willing to pay this price in order to sell the resultant performance credits (at a higher price). As the target diversion rate increases the price that the reprocessor can charge declines, reflecting the decline in the price of performance credits (see Figure 8.7).

As was noted in Chapter 6, the imposition of the obligation on the packaging user has the effect of increasing the price of packaging – by an amount equal to the price of performance credits multiplied by the target diversion rate. Consequently, for a given target rate, the higher the price of credits, the smaller the quantity of packaging that is used (i.e. the greater the level of source reduction). When the reprocessor has the property rights and the target diversion rate is below the 50% threshold, the price of credits is inversely proportional to the diversion rate.[11] Consequently, the value of the "packaging tax" is constant; as is the amount of packaging used (albeit at half the pre-regulation level). As predicted in Corollary 8.1, the actual diversion rate is equal to the threshold value of 50%. However, when the target diversion rate is above the threshold, or when the rights are granted to the waste manager, the value of the "packaging tax" is positively related to the target diversion rate, and hence the amount of packaging used declines as the target rate rises.[12]

[11] This follows directly from the expressions for $q(v; r)$ and $m(v; r)$, noting that $m(v^*; r) = 0$.

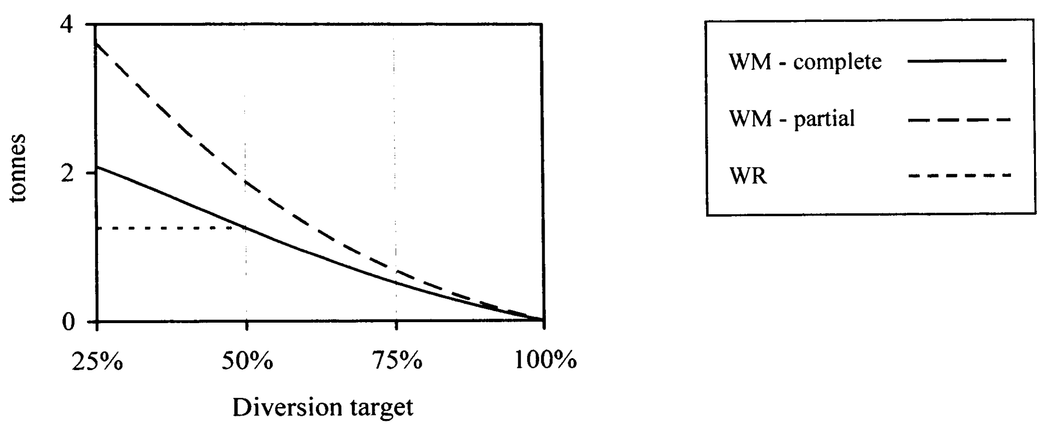
[12] If the reprocessor has the property rights, then the price of performance credits declines. However, this is more than offset by the increase in the diversion target.

Figure 8.9 Packaging used (*w*)



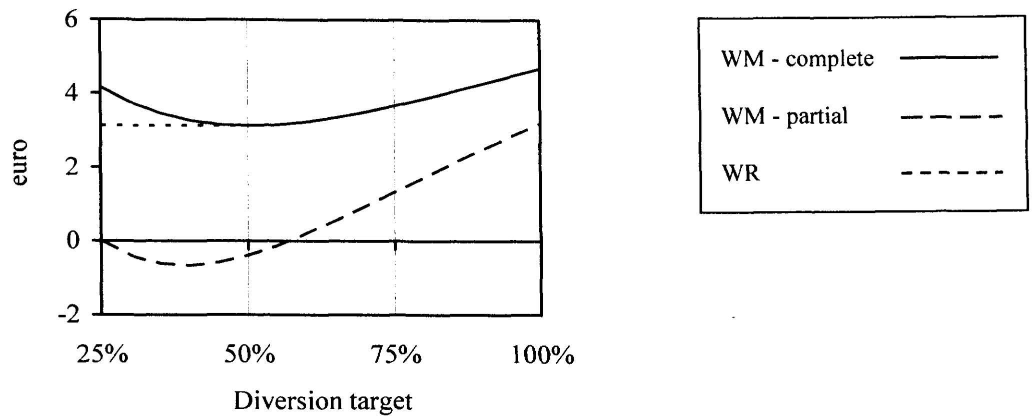
A very similar picture can be seen with regard to the quantity of waste packaging that is sent to landfill. When the target diversion rate is below the threshold and the reprocessor has the property rights, then both the amount of packaging used and the amount diverted are constant, and hence so to is the amount sent to landfill. In all of the other cases the amount of diverted waste packaging is equal to the number of credits purchased by the producer, and consequently landfilled waste is equal to (1 – r)% of the amount used. Consequently, as the amount used declines, so too does the amount going to landfill.

Figure 8.10 Waste packaging sent to landfill (*l*)



Turning to a comparison of the financial impacts, Figure 8.11 shows the total cost under the alternative property rights regimes (i.e. the reduction in aggregate benefits).[13] It is clear from this that the relative efficiency of the two property rights regimes depends critically on the assumption that is made about the knowledge of the reprocessor regarding the price of performance credits.

Figure 8.11 Total cost



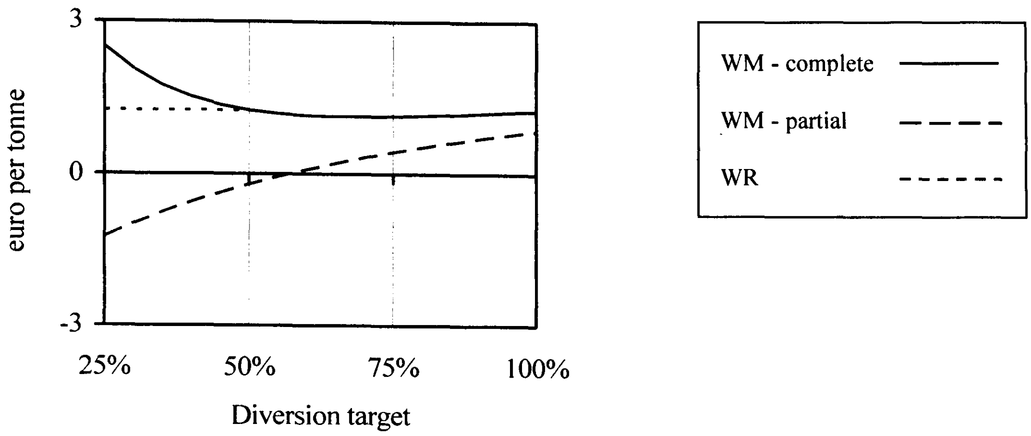
If the reprocessor has complete knowledge then, for target diversion rates below the 50% threshold, the aggregate cost is higher when the property rights are granted to the waste manager. However, if the reprocessor has partial knowledge then the aggregate cost is always lower if the waste manager has the rights. Indeed, for diversion rates below 60%, aggregate profits actually increase (i.e. the aggregate cost is negative). This is because at these rates, the scheme is also effectively acting as an instrument of market regulation – reducing the deadweight loss in the market for diverted waste packaging arising from the monsopony power of the reprocessor (with the packaging user rather than the government providing the subsidy). For relatively small increases in the target diversion rate (versus the pre-regulation value of 25%), and hence relatively small

[13] This has been calculated under the assumption that households are charged the full marginal cost of waste collection, and hence that source reduction by packaging user has no financial impact on the waste manager (see Chapter 6, section 6.3.2).

increases in the quantity of diverted waste packaging, this gain is sufficiently large to outweigh the cost imposed on the packaging user.

Given that the environmental outcomes can differ under the alternative property rights regimes (see Figure 8.10), it may be more appropriate to compare efficiency in terms of the cost per tonne of abatement (i.e. the reduction in waste packaging that is sent to landfill). However, as can be seen from Figure 8.12, this makes no difference to the conclusions regarding the relative costs under the two property rights regimes.

Figure 8.12 Cost per tonne of reduction of landfilled waste packaging



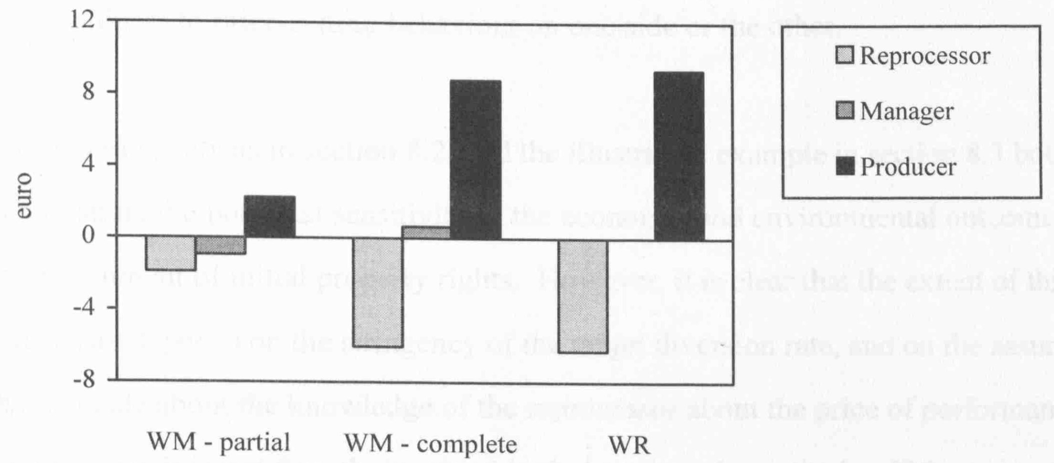
Finally, Figure 8.13 considers the distribution of (positive) costs and (negative) benefits between the three sectors for two particular target diversion rates – one below the 50% threshold, and one above.[14] There are two main points to note. First, in all cases there is a financial transfer from the packaging user to the downstream sectors. This, of course, is to be expected given the discussion in Chapter 6. However, it is true even when there is a net benefit (i.e. with the 35% target and partial knowledge). Second, while the reprocessor always benefits from the introduction of the scheme, the impact on the waste management sector depends on the level of the target, the allocation of the

[14] The horizontal supply curve for packaging implies that the packaging producers are unaffected by the regulatory intervention – making zero profits in all cases.

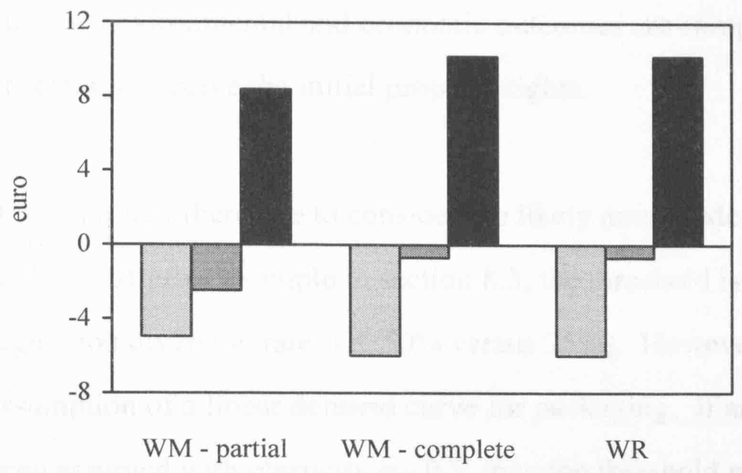
property rights, and the assumption that is made about the reprocessor's knowledge. Interestingly, when the target diversion rate is 35% and the reprocessor has complete knowledge, the waste collector is actually worse off if it has the property rights.

Figure 8.13 Distribution of costs and benefits

a) Diversion target = 35%



b) Diversion target = 70%



8.4 Summary

This chapter has considered how the assignment of initial property rights in a performance-based credit trading scheme for a waste diversion target under extended producer responsibility (EPR) can affect the economic and environmental outcomes when market for the diverted waste is subject to monopsony power. However, the insights that it provides are relevant to other applications where the supply of performance credits is (largely) determined by the sales of a single market commodity that is subject to price-setting behaviour on one side or the other.

The generic analysis in section 8.2, and the illustrative example in section 8.3 both demonstrate the potential sensitivity of the economic and environmental outcomes to the assignment of initial property rights. However, it is clear that the extent of this sensitivity depends on the stringency of the target diversion rate, and on the assumption that is made about the knowledge of the reprocessor about the price of performance credits when it is not directly involved in that market. In particular, if the reprocessor fully understands the relationship between the price of diverted waste packaging and the price of performance credits, and the target diversion rate is above the threshold level, then the environmental and economic outcomes are completely unaffected by the choice of sector to receive the initial property rights.

It is important therefore to consider the likely magnitude of the threshold diversion rate. In the illustrative example in section 8.3, the threshold is equal to twice the pre-regulation diversion rate (i.e. 50% versus 25%). However, this is an artefact of the assumption of a linear demand curve for packaging. If an iso-elastic demand curve had been assumed with elasticity = - 0.5, then the threshold rate would have been four times greater than the pre-regulation rate (i.e. 100%). Consequently, it is quite possible that the threshold rate will exceed any feasible target diversion rate, and therefore that the outcomes will depend on which sector is granted the initial property rights.

This leads to the question of which assumption is most realistic regarding the knowledge of the reprocessor. Unfortunately, neither assumption seems entirely plausible. The assumption of complete knowledge implies that the reprocessor fully understands the demand for credits, so that it can correctly forecast how its actions in the diverted waste packaging market will affect the price of credits. However, this seems unlikely, given that it does not itself participate in the market.

On the other hand, the partial knowledge assumption implies that the reprocessor can correctly predict the equilibrium price in the market for credits, but does not have any information about the shape of the demand curve. One possible explanation for this apparent anomaly would be that the price is correctly forecast by outside experts such as market analysts, who then make the information public without revealing their underlying assumptions. A more realistic assumption might be that the reprocessor is partially rational – i.e. it understands that it can influence the price of credits, but it has imperfect information about the shape of the demand curve. This would provide a useful extension to the relatively simple model considered here.

If there is reason to believe that the reprocessor has complete knowledge, then it would be better to grant the property rights to the reprocessor. The environmental outcome is better (i.e. less packaging is produced, and less waste packaging is sent to landfill), and the aggregate cost is lower.[15] The only caveat is that the lower cost hides a greater financial transfer to the reprocessor, and the gross cost to the producer is higher. If on the other hand, it is felt that the partial knowledge assumption is more appropriate, then the regulator is faced with a trade-off. The aggregate cost will be lower if the property rights are granted to the waste collectors, but the environmental outcome will be worse. One approach to reconciling this trade-off is to base the decision on the average cost per

[15] The conclusions regarding the relative costs of the alternative property rights regimes are based on the simulation results in section 5, and have not been proved analytically.

tonne of abatement (i.e. reduction in landfilled waste packaging). On this basis it is preferable to assign the property rights to the waste collectors. However, in order to get the same environmental outcome it will be necessary to set a higher target diversion rate.

Of course, this analysis has been based on a highly stylised model of the packaging system, and the simplest case of market power – i.e. monopoly power. In a more realistic representation there would be more than one packaging material, with the varying degrees of strategic behaviour in the respective markets for diverted waste. This would open up a range of potential market power configurations across the various markets for diverted waste packaging and the market for performance credits; each of which would need to be investigated individually.

Appendix A8.1 Proof of Proposition 8.1

For any functional form of $b(w)$ that satisfies the assumptions set out in section 8.1, there exists a threshold value $\check{r} > r^0$ such that $v^* < v^0 = y^0 = y^*$ for all values of $r < \check{r}$.

Proof

Denote the pre-regulation market equilibrium values by p^0, y^0 and w^0. The pre-regulation diversion rate is $r^0 = y^0 / w^0$, and the producers marginal willingness to pay for packaging is equal to the exogenous price, i.e. $b(w^0) = c$

By construction, if $v^0 = y^0$ then $q^0 = q(v^0; r^0) \equiv \frac{1}{r^0}\left[b\left(\frac{v^0}{r^0}\right) - c \right] = 0$

Denote the reprocessor's marginal revenue function in for performance credits by $m(v; r)$, where:

$$m(v; r) \equiv \frac{1}{r}\left[b\left(\frac{v}{r}\right) + \frac{v}{r} b'\left(\frac{v}{r}\right) - c \right]$$

Then: $m(v^0; r^0) < 0$ since $b' < 0$

$m(0; r^0) > 0$ since $b(0) > c$

Under the assumptions that have been made for the properties of the marginal benefit function $b(w)$, it follows that $m(v; r)$ is continuous and that $m'(v; r) < 0$ for all v. Therefore, by the intermediate value theorem, there exists a unique value $\hat{v}^0 \in (0, v^0)$ such that $m(\hat{v}^0; r^0) = 0$.

For $r \geq r^0$, let $\hat{v}(r)$ solve the implicit function $m(v; r) = 0$. By the implicit function theorem:

$$\hat{v}'(r) = -\frac{\partial m}{\partial r} \bigg/ \frac{\partial m}{\partial v} = -\left(\frac{-\hat{v}}{r^2}\right) \bigg/ \left(\frac{1}{r}\right) = \frac{\hat{v}}{r}$$

which implies that $\hat{v}(r) = k\,r$ (where $k = \hat{v}^0 / r^0$)

Therefore, there exists a finite value $\check{r} = \left(\frac{v^0}{\hat{v}^0}\right) r^0 > r^0$, such that

$$\hat{v}(\check{r}) = v^0 = y^0$$

$$\hat{v}(r) < v^0 = y^0 \quad \text{for all } r \in [r^0, \check{r})$$

That is, for all $r \in [r^0, \check{r})$ marginal revenue equals zero for a quantity of performance credits that is strictly less than $v^0 = y^0$. However, the marginal cost of creating credits is zero for all $v \leq y^0$ (i.e. $\mu = 0$). Thus, for all $r \in [r^0, \check{r})$, the reprocessor will maximise profits by setting $v^* = \hat{v}(r) < v^0 < y^0$ and $y^* = y^0$.

Q.E.D.

Appendix A8.2 Proof of Corollary 8.1

If the target diversion rate is less than the threshold rate (ř), then the actual diversion rate is equal to the threshold rate, i.e.

$$\check{r} = \frac{w^*}{y^*} > \frac{v^*}{y^*} = r \quad \text{for all } r < \check{r}$$

Proof

From the proof of Proposition 8.1

$$k\,r = v^* < y^* = y^0 = v^0 = k\,\check{r} \quad \text{for all } r \in [r^0, \check{r})$$

Therefore, since $w^* > 0$

$$\frac{k}{w^*}r = \frac{v^*}{w^*} < \frac{y^*}{w^*} = \frac{y^0}{w^*} = \frac{v^0}{w^*} = \frac{k}{w^*}\check{r}$$

But $\frac{k}{w^*} = 1$ since $\frac{v^*}{w^*} = r$ by condition (8.8.c)

Hence, $\check{r} = \frac{w^*}{y^*} > \frac{v^*}{y^*} = r$ for all $r \in [r^0, \check{r})$

Q.E.D

Appendix A8.3 Proof of Proposition 8.2

If the reprocessor has complete knowledge, and the target diversion rate is greater than the threshold value (ř), then the equilibrium price of performance credits and the quantities of all real system variables are unaffected by the allocation of the initial property rights to the performance credits. However, if the target rate is less than the threshold value, then the outcomes differ. In this case, if the initial property rights are granted to the waste manager (i.e. $\theta = 0$), then the:

- price of performance credits is lower (i.e. $q^{\#} < q^{*}$)
- quantity of performance credits is higher (i.e. $v^{\#} > v^{*}$)
- quantity of diverted waste packaging is lower (i.e. $y^{\#} < y^{*}$)
- quantity of packaging used is higher (i.e. $w^{\#} > w^{*}$)
- quantity of waste packaging sent to landfill is higher (i.e. $w^{\#} - y^{\#} > w^{*} - y^{*}$)

The price of diverted waste packaging is lower when the rights are granted to the waste manager (i.e. $p^{\#} < p^{*}$) for <u>all</u> values of the target diversion rate.

Proof

a) Target diversion rate greater than or equal to the threshold value (i.e. $r \geq ř$)

Proof is by contradiction.[16] Assume that $q^{\#} + p^{\#} > p^{*}$

$\Rightarrow \quad \chi^{-1}(q^{\#} + p^{\#}) \quad > \quad \chi^{-1}(p^{*}) \qquad$ since $\chi'(y) > 0$ by assumption

$\Rightarrow \quad y^{\#} \quad > \quad y^{*}$

$\Rightarrow \quad v^{\#} \quad > \quad v^{*}$

[16] Unless explicitly stated, the justification for each step of the proof is provided by the necessary conditions for the respective market equilibria (i.e. conditions (8.8)/(8.9) and (8.16), or by earlier steps in the proof.

and $\quad w^{\#} \quad > \quad w^{*}$

$\Rightarrow \quad q^{\#} \quad < \quad q^{*} \qquad$ since $b'(w) < 0$ by assumption

$$\text{and} \quad q^{\#} + \frac{v^{\#}}{r^2} b'\left(\frac{v^{\#}}{r}\right) \quad = \quad \frac{1}{r}\left[b\left(\frac{v^{\#}}{r}\right) + \frac{v^{\#}}{r} b'\left(\frac{v^{\#}}{r}\right) \right]$$

$$< \quad \frac{1}{r}\left[b\left(\frac{v^{*}}{r}\right) + \frac{v^{*}}{r} b'\left(\frac{v^{*}}{r}\right) \right]$$

$$= \quad q^{*} + \frac{v^{*}}{r^2} b'\left(\frac{v^{*}}{r}\right)$$

since $2b'(w) + w\, b''(w) \; < \; 0$ by assumption.

Consequently:

$$p^{\#} + y^{\#}\left[\chi'(y^{\#}) - \frac{1}{r^2} b'\left(\frac{v^{\#}}{r}\right) \right] \quad = \quad \left[p^{\#} + q^{\#} + y^{\#} \chi'(y^{\#}) \right] - \left[q^{\#} + \frac{v^{\#}}{r^2} b'\left(\frac{v^{\#}}{r}\right) \right]$$

$$> \quad \left[p^{*} + y^{*} \chi'(y^{*}) \right] - \left[q^{*} + \frac{v^{*}}{r^2} b'\left(\frac{v^{*}}{r}\right) \right]$$

$$= \quad \beta(y^{*})$$

$$> \quad \beta(y^{\#}) \qquad \text{since } \beta'(y) < 0 \text{ by assumption.}$$

But this contradicts the necessary condition (8.16.a). Consequently, the assumption cannot be true. Assume instead that $q^{\#} + p^{\#} \; < \; p^{*}$. Following the same logic (but with the inequality signs reversed), this implies that:

$$p^{\#} + y^{\#}\left[\chi'(y^{\#}) - \frac{1}{r^2} b'\left(\frac{v^{\#}}{r}\right) \right] \quad < \quad \beta(y^{\#})$$

which again contradicts condition (8.16.a).

Therefore, it must be the case that $q^{\#} + p^{\#} = p^{*}$, which implies that $p^{\#} < p^{*}$. It follows from the above logic that $q^{\#} = q^{*}$, $y^{\#} = y^{*}$, $w^{\#} = w^{*}$ and $v^{\#} = v^{*}$, and hence that $w^{\#} - y^{\#} = w^{*} - y^{*}$.

Q.E.D.

b) Target diversion rate less than the threshold (i.e. $r < \check{r}$)

Proof is by contradiction. Assume that $q^{\#} + p^{\#} \geq p^{*}$

$$\Rightarrow \quad \chi^{-1}(q^{\#} + p^{\#}) \quad \geq \quad \chi^{-1}(p^{*}) \qquad \text{since } \chi'(y) > 0 \text{ by assumption}$$

$$\Rightarrow \quad y^{\#} \quad \geq \quad y^{*}$$

$$\Rightarrow \quad v^{\#} \quad > \quad v^{*}$$

and $\quad w^{\#} \quad > \quad w^{*}$

$$\Rightarrow \quad q^{\#} \quad < \quad q^{*} \qquad \text{since } b'(w) < 0 \text{ by assumption}$$

and
$$q^{\#} + \frac{v^{\#}}{r^2} b'\left(\frac{v^{\#}}{r}\right) = \frac{1}{r}\left[b\left(\frac{v^{\#}}{r}\right) + \frac{v^{\#}}{r} b'\left(\frac{v^{\#}}{r}\right) \right]$$

$$< \frac{1}{r}\left[b\left(\frac{v^{*}}{r}\right) + \frac{v^{*}}{r} b'\left(\frac{v^{*}}{r}\right) \right]$$

$$= q^{*} + \frac{v^{*}}{r^2} b'\left(\frac{v^{*}}{r}\right)$$

since $2b'(w) + w\, b''(w) < 0$ by assumption.

Consequently:

$$p^{\#} + y^{\#}\left[\chi'(y^{\#}) - \frac{1}{r^2}b'\left(\frac{v^{\#}}{r}\right)\right] \quad = \quad \left[p^{\#} + q^{\#} + y^{\#}\chi'(y^{\#})\right] - \left[q^{\#} + \frac{v^{\#}}{r^2}b'\left(\frac{v^{\#}}{r}\right)\right]$$

$$> \quad \left[p^{*} + y^{*}\chi'(y^{*})\right] - \left[q^{*} + \frac{v^{*}}{r^2}b'\left(\frac{v^{*}}{r}\right)\right]$$

$$= \quad \beta(y^{*})$$

$$> \quad \beta(y^{\#}) \qquad \text{since } \beta'(y) < 0 \text{ by assumption.}$$

But this contradicts the necessary condition (8.16.a). Consequently, the assumption cannot be true, and it must be the case that $q^{\#} + p^{\#} < p^{*}$. It follows directly that $p^{\#} < p^{*}$ and that $y^{\#} < y^{*}$. Unfortunately, because of the strict inequality $y^{*} > v^{*}$, it is not possible to make any inferences about the relative values of other variables using the foregoing logic (i.e. one cannot infer the relative magnitudes of $v^{\#}$ and v^{*} from $y^{\#}$ and y^{*}). However, the fact that $y^{\#} < y^{*}$,

$$\Rightarrow \quad \beta(y^{\#}) \quad > \quad \beta(y^{*}) \qquad \text{since } \beta'(y) < 0 \text{ by assumption}$$

$$\Rightarrow \quad p^{\#} + y^{\#}\left[\chi'(y^{\#}) - \frac{1}{r^2}b'\left(\frac{v^{\#}}{r}\right)\right] \quad > \quad p^{*} + y^{*}\chi'(y^{*})$$

$$\Rightarrow \quad y^{\#}\chi'(y^{\#}) - \frac{y^{\#}}{r^2}b'\left(\frac{v^{\#}}{r}\right) - q^{\#} \quad > \quad y^{*}\chi'(y^{*}) - \frac{v^{*}}{r^2}b'\left(\frac{v^{*}}{r}\right) - q^{*}$$

$$\text{since} \quad p^{\#} < p^{*} - q^{\#} \quad \text{and} \quad \left[q^{*} + \frac{v^{*}}{r^2}b'\left(\frac{v^{*}}{r}\right)\right] = 0$$

$$\Rightarrow \quad q^{\#} + \frac{v^{\#}}{r^2}b'\left(\frac{v^{\#}}{r}\right) \quad > \quad q^{*} + \frac{v^{*}}{r^2}b'\left(\frac{v^{*}}{r}\right)$$

$$\Rightarrow \quad b\left(\frac{v^{\#}}{r}\right) + \frac{v^{\#}}{r}b'\left(\frac{v^{\#}}{r}\right) \quad > \quad b\left(\frac{v^{*}}{r}\right) + \frac{v^{*}}{r}b'\left(\frac{v^{*}}{r}\right)$$

$\Rightarrow \quad v^{\#} \quad < \quad v^{*} \qquad$ since $2b'(w) + w\,b''(w) < 0$ by assumption

$\Rightarrow \quad w^{\#} \quad > \quad w^{*}$

and $\quad q^{\#} \quad < \quad q^{*} \qquad$ since $b'(w) < 0$ by assumption

Finally, since $w^{\#} > w^{*}$ and $y^{\#} < y^{*}$ it follows that $w^{\#} - y^{\#} > w^{*} - y^{*}$.

Q.E.D.

Appendix A8.4 Proof of Proposition 8.3

If the initial property rights to the performance credits are granted to the waste manager then, for all values of the target diversion rate, the equilibrium outcomes will differ depending on the assumption that is made regarding the knowledge of the reprocessor regarding the price of performance credits. If the reprocessor is has partial knowledge (i.e. it assumes that $q = q^{\Delta}$), then the:

- price of performance credits is lower (i.e. $q^{\Delta} < q^{\#}$)
- price of diverted waste packaging is higher (i.e. $p^{\Delta} > p^{\#}$)
- quantity of performance credits is higher (i.e. $v^{\Delta} > v^{\#}$)
- quantity of diverted waste packaging is higher (i.e. $y^{\Delta} > y^{\#}$)
- quantity of packaging used is higher (i.e. $w^{\Delta} > w^{\#}$)
- quantity of waste packaging sent to landfill is higher (i.e. $w^{\Delta} - y^{\Delta} > w^{\#} - y^{\#}$)

Proof

Proof is by contradiction.[17] Assume that $q^{\#} + p^{\#} \geq q^{\Delta} + p^{\Delta}$

$\Rightarrow \quad \chi^{-1}(q^{\#} + p^{\#}) \geq \chi^{-1}(q^{\Delta} + p^{\Delta})$ since $\chi'(y) > 0$ by assumption

$\Rightarrow \quad y^{\#} \geq y^{\Delta}$

$\Rightarrow \quad v^{\#} \geq v^{\Delta}$

and $\quad w^{\#} \geq w^{\Delta}$

$\Rightarrow \quad q^{\#} \leq q^{\Delta}$ since $b'(w) < 0$ by assumption

[17] Unless explicitly stated, the justification for each step of the proof is provided by the necessary conditions for the respective market equilibria (i.e. conditions (8.16) and (8.17), or by earlier steps in the proof.

$$\Rightarrow \quad p^{\#} \quad \geq \quad p^{\Delta} \qquad \text{since } q^{\#} - p^{\Delta} \geq q^{\Delta} - p^{\#} \text{ by hypothesis.}$$

Therefore

$$\begin{aligned}
p^{\Delta} + y^{\Delta}\chi'(y^{\Delta}) &= \chi(y^{\Delta}) + y^{\Delta}\chi'(y^{\Delta}) - q^{\Delta} \\
&\leq \chi(y^{\#}) + y^{\#}\chi'(y^{\#}) - q^{\#} \\
&\quad \text{since } 2\chi'(y) + y\,\chi''(y) > 0 \text{ by assumption} \\
&< p^{\#} + y^{\#}\chi'(y^{\#}) - \frac{y^{\#}}{r^{2}} b'\left(\frac{v^{\#}}{r}\right) \qquad \text{since } b' < 0 \text{ by assumption} \\
&= \beta(y^{\#}) \\
&< \beta(y^{\Delta}) \qquad \text{since } \beta' < 0 \text{ by assumption}
\end{aligned}$$

But this contradicts the necessary condition (8.17.a). Consequently, the assumption cannot be true, and it must be that $q^{\#} + p^{\#} < q^{\Delta} + p^{\Delta}$. In which case it follows from the above logic that $y^{\#} < y^{\Delta}$, $v^{\#} < v^{\Delta}$, $w^{\#} < w^{\Delta}$, $q^{\#} > q^{\Delta}$ and $p^{\#} < p^{\Delta}$. By conditions (8.16.d) and (8.17.d), it follows directly that $w^{\#} - y^{\#} < w^{\Delta} - y^{\Delta}$.

Q.E.D.

Appendix A8.5 Analytic solutions for illustrative example

a) Waste reprocessor has initial property rights (WR)

Diversion rate below threshold

$$y^* = 1.25$$

$$w^* = 2.5$$

$$v^* = 2.5\,r$$

$$p^* = 1.5$$

$$q^* = \frac{2.5}{r}$$

Diversion rate above threshold

$$y^* = 2.5\left(\frac{r\,(1+r)}{1+2r^2}\right)$$

$$w^* = 2.5\left(\frac{1+r}{1+2r^2}\right)$$

$$v^* = 2.5\left(\frac{r\,(1+r)}{1+2r^2}\right)$$

$$p^* = 5\left(\frac{r\,(1+r)}{1+2r^2}\right) - 1$$

$$q^* = 2.5\left(\frac{1-r+4r^2}{r\,(1+2r^2)}\right)$$

b) Waste manager has initial property rights (WM)

Reprocessor has complete knowledge

$$y^\# = 2.5\left(\frac{r\,(1+r)}{1+2r^2}\right)$$

$$w^\# = 2.5\left(\frac{1+r}{1+2r^2}\right)$$

$$v^\# = 2.5\left(\frac{r\,(1+r)}{1+2r^2}\right)$$

$$p^\# = 4 - 2.5\left(\frac{1+r}{r}\right)$$

$$q^\# = 2.5\left(\frac{1-r+4r^2}{r\,(1+2r^2)}\right)$$

Reprocessor has partial knowledge

$$y^\Delta = 5\left(\frac{r\,(1+r)}{1+4r^2}\right)$$

$$w^\Delta = 5\left(\frac{1+r}{1+4r^2}\right)$$

$$v^\Delta = 5\left(\frac{r\,(1+r)}{1+4r^2}\right)$$

$$y^\Delta = 4 - 10\left(\frac{r\,(1+r)}{1+4r^2}\right)$$

$$q^\Delta = 5\left(\frac{4r-1}{1+4r^2}\right)$$

Part Four

Conclusion

Chapter 9 Conclusion

In an ideal, first-best world there are clear theoretical arguments for preferring a regulatory target that takes the form of an (explicit or implicit) absolute aggregate limit over one that takes the from of a relative aggregate performance standard. When the respective implementation mechanisms are both cost-efficient, the gross economic cost of achieving any given environmental objective is higher under the performance standard. Consequently, while it is possible to achieve the socially optimal outcome (i.e. maximize net welfare) under the aggregate limit, it is not possible to do so under the aggregate performance standard.

However, when one moves away from this rather unrealistic view of the world, to one in which there is uncertainty about the economic conditions that will apply when the regulatory intervention is implemented, regulated firms are able to act strategically in their output markets, and the government uses distortionary labour taxes to raise revenue, then the picture becomes more complex. In particular, in the presence of labour taxes net welfare is higher under the absolute limit if the implementation mechanism raises revenue which is used to reduce the tax rates. However, if revenues are returned in the form of lump sum payments, or if the mechanism used to implement the absolute limit is revenue-neutral, then "potential" net welfare can be higher under the performance standard. Therefore, if revenue-neutrality is a necessary condition for the political acceptability of an intervention, it may be preferable that the target takes the form of an aggregate performance standard – provided that this can be implemented efficiently.

Notwithstanding the theoretical arguments, relative performance standards are used in a wide range of environmental policy areas – including waste / resource management, fuel / energy efficiency, climate change and air / water quality. This provides a strong practical argument for the development of a cost-efficient implementation mechanism, as it is likely to be easier to change the existing mechanism than to move to a completely new form of regulatory intervention based on an absolute limit.

To date the majority of relative performance standards have been implemented either through a system of fixed individual standards, or by negotiated agreements with industry associations that do not specify individual targets. However, there are a small number of examples where market-based, trading mechanisms have been used. The earliest example, and probably the most well known, was the lead credit trading programme that operated in the USA in the mid-1980s. More recently, trading mechanisms have been introduced in both the United Kingdom and the Netherlands.

Thus, there is clearly an interest on the part of practitioners in using a trading mechanism to implement relative performance standards. Unfortunately, little is known about the theoretical properties of such a mechanism. In particular, the following questions were posed at the end of Chapter 1:

- Can all of the many different types of relative performance standard be implemented by a trading mechanism?
- Does a trading mechanism achieve the aggregate performance standard at the lowest possible gross economic cost?

- What impact does the trading-mechanism have on the prices and quantities of commodities used and produced by the regulated agents?
- What are the distributional impacts of the mechanism, and is it possible to manipulate these impacts?
- What information does the trading mechanism provide about the cost of the regulatory intervention?
- Is there an equivalence (symmetry) between the trading mechanism and a price-based mechanism?
- What are the implications of market imperfections for the environmental effectiveness and economic efficiency of the mechanism?

Based on the findings of the various analyses in Chapters 2 – 8, and the insights provided by the illustrative examples, it is now possible to provide some answers to these seven questions.

Existence

Performance-based credit trading can be used to implement any regulatory target that can be expressed in terms of a linear *aggregate performance rule*. The general formulation of the rule is very flexible, and by choosing appropriate values for the rule parameters, it is possible to represent a variety of different forms of target with a range of potential policy applications. In particular, the rule can represent any relative performance standard that is expressed as a *proportion*, a *weighted average*, or a *rate*. These include, for example, targets for industrial energy efficiency; CO_2 emissions per unit output; the percentage of end-of-life products or materials recovered from the waste stream; the average fuel efficiency of vehicles.

The detailed design of the scheme can be tailored to meet other (social and political) policy constraints, or to ameliorate the effect of market distortions, by varying the values of two sets of *design parameters* that affect the distribution of *obligations* and / or *property rights* between agents. However, provided that – after trading – each regulated agent satisfies its own *individual performance rule*, the aggregate performance is guaranteed to be satisfied irrespective of the values that are chosen for the design variables. Hence, the mechanism is environmentally effective in the "narrow" sense of the term – i.e. that the regulatory target is achieved.

Efficiency

If all markets are perfectly competitive, and there is symmetry between buying prices and selling prices (i.e. there are no product taxes or subsidies), then the joint market equilibrium under performance-based credit trading minimizes the cost of satisfying the aggregate performance rule, irrespective of the values that are chosen for the design parameters. As is the case with a "cap and trade" scheme for an absolute limit, the magnitude of the cost saving versus a common performance standard will depend on the heterogeneity of the agents' compliance costs. However, the simulation in the energy efficiency application (Chapter 5) suggests that – at least in some cases – the saving may be significant.

If either of the two conditions are contravened, then the absolute cost-efficiency of the mechanism is no longer guaranteed, although it may still be relatively cost-efficient compared to alternative implementation mechanisms. In particular, as the packaging recovery application (Chapter 6) demonstrates, the aggregate cost of meeting the standard may not be minimized if a commodity is subject to a tax or subsidy – even if that commodity is not itself included in the performance rule.

Cost efficiency requires that a wedge be driven between the input shadow price and output shadow price of each commodity that is included (either directly or indirectly) in the aggregate performance rule. That is, there must be a divergence between the marginal private benefit derived from the use of the commodity and the marginal private cost of its production. Depending on the specific values of the performance rule parameters the input price of a particular commodity may be higher than the output price, or it may be lower. For example, if the aggregate performance rule represents a relative standard for vehicle fuel efficiency then the output price is higher than the input price for models with better fuel efficiency than the standard, and lower models with worse fuel efficiency.

Impacts

Without making specific assumptions about the nature of agents' production technologies it is not possible to say how the shadow input and output prices compare to the pre-regulation situation (in which they are equal). Furthermore, since the actual market prices of the commodities depend on the values that are set for the *assignment parameters*, the market price of a particular commodity included in the performance rule may be higher than its pre-regulation level, lower, or the same. This is clearly illustrated in the packaging recovery application (Chapter 6), in which the regulatory intervention causes the market price of packaging to rise if the obligation to purchase performance credits is placed on the packaging producers, and to fall if it is placed on the packer-fillers.

While a change to the value of the assignment parameter for a particular commodity affects its market price, and hence the gross economic benefit of any agent that produces or uses it, this impact is offset by the corresponding change in the total value of the performance credits that the agent must acquire, or that it can sell. Consequently, the net unit benefit received by a producer, and the net unit cost paid by a user, are both unaffected by the change.

The changes to the shadow prices of the commodities included in the performance rule induce changes in the relative quantities that are produced / used compared to the pre-regulation situation. For example, if the shadow price of one input commodity falls relative to another, then there will be a rise in the relative quantity used. However, as with the prices themselves, it is not possible to draw any conclusions about changes to absolute quantities without placing further restrictions on the agents' production technologies.

Neither is it possible to draw any general conclusions about the outcome under performance-based credit trading compared to that under a common performance standard, as the analysis of the energy efficiency application (Chapter 5) clearly demonstrates. In this case, while output is unambiguously higher for a buyer of performance credits, the output of a seller may be higher or lower. The overall impact of trading on the aggregate output of the regulated sector may also be in either direction; and hence so too may the impact on aggregate energy consumption. Thus, unlike the trading of permits in a "cap and trade" scheme for an absolute limit, the trading of performance credits is likely to affect the environmental effectiveness of the regulatory intervention. However, it does not necessarily follow that the impact will be detrimental. As the numerical simulation in Chapter 5 illustrates, the environmental outcome may well be better under performance-based credit trading than under a common performance standard.

Information

At the aggregate level, the market price of performance credits provides a measure of the increase in aggregate gross economic benefit that would arise from a (small) relaxation of the performance rule – via marginal changes to the regulatory control variables. Without being more specific about the form of the regulatory target – and hence the way in which the control variables enter into the rule, it is not possible to provide a more intuitive interpretation of the price. In the case of a relative performance standard for emissions per unit output, or a hybrid emissions target, it is equal to the marginal aggregate cost of (absolute) emissions reduction multiplied by a factor that reflects the relative responsiveness of aggregate output and emissions to the changes in the target rate.

For an individual agent, the price represents the increase in gross economic benefit that would arise from a marginal increase in the number of performance credits that it holds. Again in the case of a relative performance standard for emissions, this is equal to its marginal cost of abatement multiplied by a factor that reflects the relative responsiveness of its individual output and emissions. Thus, while the trading of performance credits results in the marginal costs of satisfying the individual performance rules being equalized across agents, the marginal costs of (absolute) emissions reduction will only be equalized if the value of the factor is the same for all agents – and there is no *a priori* reason why this should be the case.

The impact of an increase in the stringency of the regulatory target on the market price of performance credits depends on the particular form of the target. For both an absolute emissions limit, and a relative emissions standard, an increase in the stringency of the regulatory target leads to a rise in the equilibrium price of performance credits. This is in line with intuition and, in the first case, is consistent with the impact of a reduction in the total number of permits that are issued in a "cap and trade" permit system.

In other applications however, the relationship is not so clear-cut. For example, in the case of a hybrid emissions target, the price of performance credits unambiguously increases if the value of the "limit control variable" is reduced. However, if the value of the "rate control variable" is reduced, then the impact depends the on the shape of the aggregate production function, and the current value of the variable. In particular, if returns to scale for the aggregate production are close to being constant, and the current value of the variable is relatively low (i.e. the target is already relatively stringent), then it is possible that a further decrease may cause the price of performance credits to fall.

A similar response can also occur in the case of a recycled content standard. In this case, the impact of an increase in the value of the target rate depends on the elasticity of substitution between the recycled and virgin input commodities. If this is less than or equal to one, then the price of performance credits unambiguously increases. However, if it significantly greater than one (i.e. the commodities are very close substitutes), and the target rate is relatively high, then it is possible that the price of credits may fall.

The market for performance credits also provides information on the aggregate cost of the regulatory intervention; this being equal to the area under the lower envelope of the aggregate inverse demand and supply curves for performance credits. A linear approximation of this area can be calculated simply from the agents' pre-regulation quantities of the commodities included in the aggregate performance rule and the equilibrium price of credits. If the inverse demand and supply functions are convex, then this approximation provides an upper bound on the actual cost. An important point to note is that the total value of the financial transfer resulting from the performance credit transactions is <u>not</u> generally equal to the aggregate cost of the intervention. Unfortunately, it is not possible to draw any unambiguous conclusion regarding the efficiency application (Chapter 5) illustrates, care needs to be taken when using this approach as apparently egalitarian rules may actually make the problem worse. Alternatively, as the packaging recovery simulation illustrates, the factors may be determined *ex post* once the scheme is operational, using information on the agents' sales / purchases of performance credits.

Equivalence

Performance-based credit trading is not the only cost-efficient implementation mechanism for a relative performance standard. The aggregate cost minimum can also be attained by a price-based mechanism in which taxes are applied to the input prices and / or output prices of the commodities included in the aggregate performance rule; with negative taxes representing subsidies. Indeed, there is a range of different tax-subsidy schemes that will support the cost minimum; with each scheme corresponding to a specific set of values for the assignment parameters under performance-based credit trading.

While the values of the taxes / subsidies for particular commodities vary between the different schemes, the net revenue received by the government is same in all cases; being equal to the market equilibrium price of performance credits multiplied by the value of the constant term in the aggregate performance rule. Thus, for any "pure" relative performance standard, where the constant term is set equal to zero, any supporting tax-subsidy scheme is revenue neutral. This is illustrated in the packaging recovery application (Chapter 6), in which the revenue raised from the tax on packaging products exactly matches the cost of providing the subsidy on diverted waste packaging.

Market power

The preceding conclusions rely on the assumption that the market for performance credits, and the markets for all commodities in the production system, are perfectly competitive, and hence that all agents are price-takers. If there is strategic interaction between agents, with one or more of the agents being able to exercise market power, then some of these conclusions will be altered.

Previous analyses of the implications of strategic interaction for market-based implementation mechanisms have – with one exception – assumed that the regulatory target takes the form of an absolute limit for emissions, and that the mechanism is a "cap and trade" permit system. The only analysis to assume a relative emissions standard has shown that if the output market is oligopolistic, then overall welfare may be higher under a common performance standard even if the market for performance credits is perfectly competitive, and hence that the trading mechanism is cost-efficient in terms of achieving the standard.

While some of the conclusions of the other analyses apply to the case of a relative emissions standard – implemented by a performance-based credit trading scheme, others do not. For example, in the situation where one firm has the power to set the price of performance credits but all other markets are competitive, the conclusions carry over directly. The only difference being that the scale of the efficiency loss depends on the values that are set of the performance adjustment factors rather than on the initial allocation of permits. However, in the case where the firm is also able to set the price in the output market, the conclusion that it will engage in "exclusionary manipulation" of the market for performance credits in order to enhance its position in the output market may not apply if the firm is a buyer of credits.

relative magnitudes of the financial transfer and the aggregate cost. Depending on the shapes of the inverse supply and demand curves, the financial transfer may overstate, or understate the cost.

Distributional flexibility

While the trading of performance credits will minimize the aggregate cost of meeting the performance target, it may have significant distributional impacts. As the simulations in relation to the energy efficiency application (Chapter 5) and the packaging recovery application (Chapter 6) illustrate, the scale of the financial transfers arising from the trading of credits may be substantially higher than the aggregate cost, and may exacerbate an already inequitable cost distribution under the common performance standard.

Interestingly, the choice of values for the assignment parameters in the individual performance rules (i.e. the assignment of obligations and property rights) has no impact on the distribution of the cost between agents; the change in the costs / revenues of agents' credit transfers being exactly offset by changes to the prices of the commodities that they produce or use. However, it is possible to alter the distribution of the aggregate cost between the agents by setting non-zero values for the *performance adjustment factors* in the individual performance rules. Varying the values of these factors results in a reallocation of property rights between agents; changing the number of performance credits that each agent can sell, or that it must acquire. Since this has no impact on the market price of performance credits, it has the effect of making lump-sum transfers between the agents.

The values of the adjustment factors can be set *ex ante* using some mechanistic burden sharing rule – for example, that the percentage improvement versus pre-regulation performance is the same for all agents. However, as the simulation for the energy

Given the many different potential manifestations of strategic interaction, it is not possible to draw any general conclusions about the implications for the effectiveness, or the efficiency of performance-based credit trading. Each case must be considered individually. However, if it affects the market for any of the commodities included in the performance rule, then it is likely to have implications for the design of the mechanism – in particular the assignment of obligations and / or initial property rights. For example, as the analysis of the packaging recovery application (Chapter 8) illustrates, if the market for the commodity associated with the generation of performance credits is monopolistic (monopsonistic), then both the environmental effectiveness and the economic efficiency of the mechanism may depend on whether the initial property rights are granted to the single seller (buyer) of the commodity, or the many sellers (buyers).

References

ACR (1999) *The Application of Local Taxes and Fees for the Collection of Household Waste: Local Authority Jurisdiction and Practice in Europe*, Technical Report, Brussels, Association of Cities for Recycling

Adar, Z. & Griffin, J. M. (1976) "Uncertainty and the Choice of Pollution Control Instruments, *Journal of Environmental Planning and Management*, 3:178-188

Bailey, I. (1999) "Competition, Sustainability and Packaging Policy in the UK", *Journal of Environmental Planning and Management*, 42(1), 83-102

Barnett, A. H. (1980) "The Pigouvian Tax Rule under Monopoly", *American Economic Review,* LXX: 1037-1041

Baumol, W. J. & Oates, W. E. (1988) *The Theory of Environmental Policy*, Cambridge, Cambridge University Press

Buchanan, J. M. (1969) "External Diseconomies, Corrective Taxes, and Market Structure", *American Economic Review*, LIX: 174-177

Burtraw, D. & Cannon, M. (2000) *Heterogeneity in Costs and Second-Best Policies for Environmental Protection*, Discussion Paper 00-20, Washington, DC, Resources for the Future

DEFRA (2001) *Framework for the UK Emissions Trading Scheme*, London, Department for the Environment, Farming and Rural Affairs

Dekkers, C. (2003) "Netherlands to Launch Legislation on NOx and CO2 Emissions Trading, *The Emissions Trader*, 7 (1), Emissions Marketing Association

DETR (1998) *Review of the Producer Responsibility Obligations (Packaging Waste) Regulations 1997: A Consultation Paper*, London, UK Department of the Environment, Transport and the Regions

Dinan, T.M. (1992) "Implementation Issues for Marketable Permits: A Case Study of Newsprint", *Journal of Regulatory Economics*, 4(1), pp 71-87

Dinan, T.M. (1993) "Economic Efficiency Effects of Alternative Policies for Reducing Waste Disposal", *Journal of Environmental Economics and Management*, 25: 242-256

ENDS (2001) "Scramble to meet EC packaging recovery target reaches last lap", *ENDS Report*, No. 317

Fershtman, C. & de Zeeuw, A. (1995) *Tradable Emission Permits in Oligopoly*, CentER Discussion Paper 9630, Tilburg University

FES (2003) *Climate Change Agreements – Results of the First Target Period Assessment*, Report to Department of the Environment, Farming and Rural Affairs, Future Energy Solutions, April 2003 (Revised July 2004)

Fischer, C. (2001) *Rebating Environmental Policy Revenues: Output-Based Allocations and Tradable Performance Standards*, Discussion Paper 01-22, Resources for the Future, Washington DC

Fullerton, D & Wu, W. (1998) "Policies for Green Design", *Journal of Environmental Economics and Management*, 36: 131-148

Gielen, A. M., Koutstaal, P. R. & Vollebergh, H. R. J. (2002) *Comparing Emission Trading with Absolute and Relative Targets*, Paper presented at the 2nd CATEP Workshop on the Design and Integration of National Tradable Permit Schemes for Environmental Protection, University College London, 25-26 March 2002

Goulder, L. H., Parry, I. W. H., Williams III, R. C. & Burtraw, D. (2000) "The cost-effectiveness of alternative instruments for environmental protection in a second best-setting", *Journal of Public Economics*, 72: 329-360

Hagem, C. & Westskog, H. (1998) “The Design of a Dynamic Tradable Quota System under Market Imperfections”, *Journal of Environmental Economics and Management*, 36: 89-107

Hahn, R. W. & Hester, G. L. (1989) “Marketable Permits: Lessons for Theory and Practice”, *Ecology Law Quarterly*, 16: 361-406

Hahn, R. W. (1984) "Market Power and Transferable Property Rights." *Quarterly Journal of Economics* **99**(4): 753-765.

Mas-Colell, A, Whinston, M. D. & Green, J. R. (1995) *Microeconomic Theory*, Oxford, Oxford University Press

Mauleg, D.A. (1990) “Welfare Consequences of Emission Credit Trading Programs”, *Journal of Environmental Economics and Management*, 18 (1): 66-77

Macauley, M.K. & Walls, M.A. (2000) “Solid Waste Policy”, in: Portney, P.R. & Stavins, R.N. (eds.) *Public Policies for Environmental Protection*, Second Edition, Washington DC, Resources for the Future, Washington, 261-286

Miranda, M.L., LaPalme, S. & Bynum, D.Z. (1998) *Unit-based pricing in the United States: a tally of communities*, Report to the USEPA, July 1998

Misiolek, W. S. & Elder, H. W. (1989) “Exclusionary Manipulation of Markets for Pollution Rights”, *Journal of Environmental Economics and Management*, 16: 156-166

Montgomery, W. D. (1972) “Markets in Licenses and Efficient Pollution Control Programs”, *Journal of Economic Theory*, 5: 395-418

O’Doherty, R. & Bailey, I. (2001) *Trading in the Environment? Tradable Permits and the UK Packaging Regulations*, Mimeo, University of the West of England

OECD (1991) *Environmental Policy: How to Apply Economic Instruments*, Paris, Organisation for Economic Cooperation and Development

Palmer, K. & Walls, M. (1999) *Extended Product Responsibility: An Economic Assessment of Alternative Policies*, Discussion Paper 99-12, Resources for the Future, Washington DC

Parry, I. W. H. & Williams III, R. C. (1999) "A second-best evaluation of eight policy instruments to reduce carbon emissions", *Resource and Energy Economics*, 21: 347-373

Pezzey, J. (1992) "The symmetry between controlling pollution by price and controlling it by quantity", *Canadian Journal of Economics*, XXV, No. 4: 983-991

Requate, T. (1993) "Pollution Control in a Cournot Duopoly via Taxes or Permits", *Journal of Economics*, 58: 255-291

Rose, A., Stevens, B, Edmonds, J. & Wise, M. (1998) "International Equity and Differentiation in Global Warming Policy", *Environmental and Resource Economics*, 12: 25-51

Salmons, R. (2002) "UK Energy Efficiency Agreement", in De Clercq, M. and Ameels, B. (eds.), *Negotiated Environmental Agreements in Europe: Critical Factors for Success*, Cheltenham, Edward Elgar

Salop, S. C. & Scheffman, D. T. (1983) "Raising Rivals' Costs", *American Economic Review*, May: 267-271

Salop, S. C. & Scheffman, D. T. (1987) "Cost Raising Strategies", *The Journal of Industrial Economics*, 26: 19-34

Sartzetakis, E. S. (1997a) "Tradeable Emission Permits Regulations in the Presences of Imperfectly Competitive Product Markets: Welfare Implications", *Environmental and Resource Economics* **9**(1): 65-81.

Sartzetakis, E. S. (1997b) "Raising Rivals' Costs Strategies via Emission Permits Markets", *Review of Industrial Organization* **12**(5-6): 751-765.

Silberberg, E. & Suen, W. (2001) *The Structure of Economics: A Mathematical Analysis*, Boston, McGraw Hill

Simon, C. P. & Blume, L. (1994) *Mathematics for Economists*, New York, W. W. Norton & Co.,

Spulber, D. F. (1985) "Effluent Regulation and Long-Run Optimality", *Journal of Environmental Economics and Management*, 12: 103-116

Tirole, J. (1988) *The Theory of Industrial Organisation*, Cambridge, MIT Press

USEPA (2001) *The United States Experience with Economic Incentives for Protecting the Environment*, Washington, National Center for Environmental Economics, US Environmental Protection Agency (EPA-240-R-01-001)

von der Fehr, N-H, M. (1993) "Tradable Emission Rights and Strategic Interaction", *Environmental and Resource Economics* 3: 129-151.

Weitzman, M. L. (1974) "Prices vs. Quantities", *The Review of Economic Studies*, XLI: 477-491