Capturing the Least Costly Way of Reducing Pollution:

A Shadow Price Approach

Abstract: The production analysis literature is increasingly concerned with estimating marginal abatement costs. Yet, most studies do not emphasize the ways in which pollutants may be reduced and their costs, which makes them unable to identify the least costly compliance strategy. This paper utilizes the materials balance principle to relate pollution to the employment of material inputs. A production model which allows input and output substitution, downscaling of operations, pollution control, and emission permits purchases as compliance strategies is proposed, and the implications of joint and non-joint pollution control for the trade-off between pollutants and desirable outputs are considered. Marginal abatement costs, reflecting the least costly way of compliance, are derived by exploiting the duality between the directional distance function and the profit function.

Keywords: Marginal abatement costs; Materials balance condition; Directional distance function

JEL-codes: D24; Q52

1. Introduction

The production analysis literature is increasingly concerned with environmental issues, in particular with estimating marginal abatement costs. These estimates can play an important role in identifying the costs of environmental regulations which, together with the gains from avoided environmental
damage\(^1\), allow determining the net benefits of environmental legislation. The estimates’ applicability for policy making hinges on their quality and validity. In turn, that also influences whether socially optimal outcomes or welfare increases are achieved. Models that are unable to capture the actual dynamics of pollution generation, as well as producers’ options for complying with environmental regulations, are unlikely to reveal the firms’ actual abatement costs.

The majority of empirical production studies that estimate marginal abatement costs apply the model framework of Färe et al. (1993; 2005) which measures marginal abatement costs by the value of forgone desirable outputs required to reduce pollutants. Yet, there is no clear explanation of how emissions are generated and how they can be reduced. The model framework is therefore not suitable for evaluating the relative costs of different compliance strategies. This is a drawback of the approach, since both common knowledge and economic theory suggest that the producers will evaluate all feasible compliance strategies before selecting the least costly activity.

In the current paper, I explicitly represent the dynamics of pollution generation by the materials balance principle. It allows identifying both uncontrolled (without pollution control) and controlled (with pollution control) emissions. Whenever information on input quantities, output levels, and pollutants is provided, uncontrolled and controlled emissions, as well as pollution control efforts, can be quantified.

My approach responds to Førsund’s (2009) demand for accounting for flexibility in producers’ responses to environmental regulations, by offering them an opportunity to reduce their emissions by input and output substitution, downscaling of operations, or pollution control, and to purchase emission permits. Contrary to the established literature, my approach allows weighing the costs of various approaches for compliance, and further to select the tool or combination of tools that minimize the producers’ costs of complying with environmental regulations. Note that this

\(^1\) This paper follows the literature on polluting technologies by not taking consumer preferences or environmental damage into account. For a discussion on these topics, see Førsund (2009) or Färe et al. (in press).
perspective is in line with the interpretation of abatement costs in the environmental economics literature – the least cost approach to satisfying environmental regulations.

My approach to marginal abatement cost estimation can be considered an extension to the approach of Färe et al. (1993), where abatement costs are derived from distance function derivatives. Contrary to Färe et al. (1993), I consider polluting firms that operate under emission constraints which may be relaxed by pollution control or purchases of emission permits. Thus, the costs of pollution control and emission permits are weighed against the economic benefits of employing polluting inputs. By maximizing profits under emission constraints and applying the duality of the directional distance function to the profit function, optimum conditions can be derived that allow identifying and estimating marginal abatement costs. Profit maximization is considered both when the production of desirable outputs is joint and non-joint with pollution control. I find that a positive trade-off between pollutants and desirable outputs – usually assumed by the literature on polluting technologies – is consistent with joint pollution control, while the trade-off may be both positive and negative in the case of non-joint pollution control. The solutions to the emission constrained profit problems are further shown to rationalize allocative inefficiency for firms that comply with environmental regulations. That is, requirements to reduce emissions increase the effective costs of polluting inputs relative to their market prices, since increases in their employment require additional spending to offset related increases in uncontrolled emissions. This recognition is important for properly understanding the dynamics of environmental regulation.

The paper is organized as follows. I review the production analysis literature on marginal abatement costs estimation in the following section. The method discussed makes up the foundation of the proposed procedures for abatement cost estimation in this paper. Section 3 discusses the materials balance principle, while section 4 incorporates it in an economic model. The derivation of marginal abatement costs is further discussed, both in the case with joint and non-joint
pollution control. Section 5 discusses the extension of the abatement cost method when multiple pollutants are regulated, in addition to focusing on computational approaches. Section 6 concludes.

2. Marginal abatement cost estimation in the literature

The literature usually treats pollutants as inputs or outputs to be included in the technology. In an early attempt to estimate marginal abatement costs, Pittman (1981) incorporates pollutants as inputs in the technology. This treatment is contingent on the assumption that positive marginal productivities of pollutants, enforced by the axiom of free disposability of inputs, characterize transformation of resources from pollution control to intended productions. Pittman defines an environmentally restricted profit function and applies the Lagrangian multiplier on the regulation constraint to obtain estimates of marginal abatement costs for a sample of pulp-and-paper mills. This modeling approach has not been followed up in the literature. However, his restricted profit problem resembles the profit maximization problems found in section 4 of this paper.

Pittman’s dataset was later used by Färe et al. (1993), who introduced a new and innovative method for estimating marginal abatement costs. In their approach, pollutants (or undesirable outputs) are treated as outputs. Let \( x \in \mathbb{R}^N \) denote a vector of inputs and \( y \in \mathbb{R}^M \) denote a vector of desirable outputs. Consider, for simplicity, only one pollutant, \( b \in \mathbb{R}_+ \). An extended output set may then be defined:

\[
P(x) = \{(y,b): x \text{ can produce } (y,b)\} \tag{1}\]

Färe et al. assume that the polluting technology satisfies the standard axioms of inactivity, compact and convex output sets, and free disposability of inputs and desirable outputs. See Färe and Primont (1995) for a discussion on these properties. In addition, two non-standard axioms are imposed to accommodate for the production of bads:
(i) if \((y,b) \in P(x)\) and \(b = 0\), then \(y = 0\)
(ii) if \((y,b) \in P(x)\) and \(0 \leq \theta \leq 1\), then \((\theta y, \theta b) \in P(x)\)

Axiom (i), null-jointness (Shephard and Färe, 1974), imposes unavoidable pollution. Axiom (ii), weak disposability (Shephard, 1970), secures that reductions in the pollutant can be achieved by simultaneously reducing some desirable outputs. According to Färe et al., this is consistent with regulations which require cleanup of pollutants, since resources are diverted from producing desirable outputs to emission reductions.

The directional output distance function is a suitable function representation for the polluting technology from equation 1 (Färe et al., 2005). The directional distance function was introduced in Chambers et al. (1996); Chung et al. (1997); Chambers et al. (1998), and allows defining maximum feasible translation of inputs and outputs in any pre-assigned direction. Here, it seeks the simultaneous maximal reduction of the pollutant and expansions of desirable outputs. Define the direction vector \(g = (g_y, -g_b)\) where \(g_y \in \mathbb{R}^M_+\) and \(g_b \in \mathbb{R}_+\), and the distance function:

\[
D_O(x, y, b; g_y, -g_b) = \sup \{ \beta \in \mathbb{R} : (y + \beta g_y, b - \beta g_b) \in P(x) \} \quad (2)
\]

The directional distance function inherits the properties of the parental technology. Under g-disposability\(^2\), the directional distance function completely characterizes the underlying polluting technology in the sense that:

\[(y, b) \in P(x) \text{ if and only if } D_O(x, y, b; g_y, -g_b) \geq 0 \quad (3)\]

\(^2\)If \((y, b) \in P(x)\) then \((y - g_y, b + g_b) \in P(x)\)
It satisfies the translation property:

\[
\tilde{D}_O(x, y + \alpha g_y, b - \alpha g_b; g_y, -g_b) = \tilde{D}_O(x, y, b; g_y, -g_b) - \alpha, \quad \alpha \in \mathbb{R}
\] (4)

and is homogenous of degree minus one in \((g_y, g_b)\), non-decreasing in \(b\), non-increasing in \(y\), and concave in \((y, b)\).

Equation 3 allows defining the revenue function in terms of the distance function. Let \(r \in \mathbb{R}_+^M\) and \(q \in \mathbb{R}_+\) be vectors of (shadow) prices and define the revenue function:

\[
R(x, r, q) = \max_{y, b} \left\{ ry - qb : \tilde{D}_O(x, y, b; g_y, -g_b) \geq 0 \right\}
\]

\[
= \max_{y, b} \left\{ ry - qb + (rg_y + qg_b) \tilde{D}_O(x, y, b; g_y, -g_b) \right\}
\] (5)

where the last equality is due to Chambers et al. (1998). The first order conditions for revenue maximization are:

\[
(rg_y + qg_b) \nabla_y \tilde{D}_O(x, y, b; g_y, -g_b) = -r
\] (6)

\[
(rg_y + qg_b) \partial \tilde{D}_O(x, y, b; g_y, -g_b) / \partial b = q
\] (7)

For the output \(m\) and the pollutant \(b\), it follows that their relative price equals the corresponding ratio of distance function derivatives. Hence:
The shadow price $q$ can now be obtained from equation 8, by assuming that the observed sales price of the output $y_m$ equals its shadow price (Färe et al., 1993). The shadow price is here interpreted as the value of desirable output that must be forgone in order to marginally reduce the pollutant. In other words, it defines the marginal abatement costs.

Färe et al.’s approach to abatement cost estimation benefits from the use of distance functions. They do not rely on price information and are therefore suitable in cases with missing prices for pollutants. Consequentially, the procedure is very popular and has been employed in several studies on polluting industries, e.g. electricity generation (Coggins and Swinton, 1996; Färe et al., 2005), agriculture (Färe et al., 2006), ceramic pavement industry (Reig-Martínez et al., 2001), and aquaculture (Liu and Sumaila, 2010).

Recently, several authors have started questioning Färe et al.’s approach to pollution modeling. Coelli et al. (2007) argue that Färe et al.’s model is inconsistent with the materials balance condition, a law of physics to be treated in the subsequent section. The approach of Färe et al. also imposes severe constraints on producers’ responses to environmental regulations, since the axiom of weak disposability of desirable and undesirable outputs only considers reductions in desirable outputs to be a feasible compliance strategy (Førsund, 2009). This is a clear weakness of the approach since other compliance strategies are preferred when their costs are less. In such cases, the approach of Färe et al. is likely to overstate the costs of compliance. The purpose of the current paper is therefore to propose some new procedures for estimating marginal abatement costs that maintain the desirable features of Färe et al.’s method, but which overcome the critique of Coelli and Førsund.
3. **Controlled and uncontrolled emissions**

Most production processes involve transformations of materials which have low economic values into final products which have higher economic values. Usually, the energy required to perform the transformations is supplied by material fuels. Baumgärtner et al. (2001) and Baumgärtner and Arons (2003) show that byproducts are inevitable in such production processes. This is a consequence of physical limits to production, imposed by the first and second law of thermodynamics. The first law of thermodynamics, often called the materials balance condition, secures that materials can neither be created nor destroyed, but may only change their form. Its implication for pollution generation is evident: the share of material inputs that is not recuperated in intended products ends up as (undesirable) byproducts. Whenever such byproducts affect the welfare of external agents, they are dubbed externalities or pollutants. See Ayres and Kneese (1969) for a brilliant discussion on the subject.

The implications of materials balance condition can for example be illustrated by sulfur dioxide emissions from fossil fuel fired power plants, where sulfur dioxide emissions depend on the sulfur contents of fossil fuels. Sulfur is not recuperated by the electricity output, and its release during fossil fuel based electricity generation is unavoidable.

The materials balance condition implies that the weight of inputs, including “non-economic” inputs such as oxygen, must amount to the weight of the desirable outputs and byproducts. Microeconomic production analysis is, however, only concerned with inputs that are in some sense economically scarce and over which the entrepreneur exercises effective control (Chambers, 1988). In this setting, it is convenient to apply emission factors and recuperation factors to represent the materials balance condition. The factors approximate the amount of byproducts released per unit of inputs used and the amount recuperated per unit of desirable outputs produced. This approach to representing the materials balance condition, which is reproduced by equation 9, is increasingly popular in the production analysis literature (Coelli et al., 2007; Lauwers, 2009). Let $u \in \mathbb{R}_+^N$ be a
vector of emission factors and \( v \in \mathbb{R}^M_v \) be a vector of recuperation factors for the desirable outputs, and define:

\[
b = ux - vy
\]  

Equation 9 is a representation of the producer’s uncontrolled emissions, i.e. emissions prior to pollution control. In the case of electricity production and sulfur dioxide emissions, the emission factors report unit emissions of sulfur dioxide for each type of fossil fuels. Non-polluting inputs such as labor and capital inputs receive emission factors of zero since they do not contribute to the generation of sulfur dioxide emissions. Similarly, \( v \) is the zero-vector since there is no sulfur dioxide embedded in the final product, i.e. electricity.

Since equation 9 is a “production function” for the pollutant it allows me to consider feasible producer responses that limit pollution. It is clear that uncontrolled emissions can be reduced by decreasing the overall scale of operations, by substituting high-polluting inputs, i.e. inputs which emission factors are relatively high, with low-polluting inputs, and by choosing output mixes that favor “high-recuperating” desirable outputs. Reductions in uncontrolled emissions can also be achieved by efficiency improvements or by technical changes that reduce the amounts of inputs required to produce a certain amount of desirable outputs. Efficiency improvements are not emphasized by the current paper since marginal abatement costs are evaluated for technically efficient firms. Also, the paper does not treat technical changes that take place over time. My model framework may, however, be extended to take such effects into account\(^3\).

So far I have only considered the uncontrolled byproduct. Firms do, however, often engage in pollution control activities to reduce pollutants. Common examples are end-of-pipe abatement

\(^3\) See for example Lee et al. (2002) for a discussion on how the estimates of marginal abatement costs can be adjusted to take inefficiency into account.
technologies or dust collection systems. Pollution control does not diminish uncontrolled emissions but transforms them into different byproducts. For example, in the case of electricity production the sulfur byproducts can be applied to produce gypsum. It does not reduce the power plants’ uncontrolled sulfur dioxide emissions, but their release to air is reduced because they are partly absorbed by end-of-pipe scrubbers. Emissions remaining after pollution control are called the controlled emissions (which is the sulfur dioxide that is emitted to air). Let \( a_c \in \mathbb{R}^+ \) denote the amount of the byproduct \( b \) that is absorbed or transformed by pollution controls. Controlled emissions are then defined by:

\[
b = ux - vy - a_c
\]  

Pollution control offers an alternative to input and output substitution, downscaling of operations, or productivity improvements that allows reducing regulated pollutants. Along with purchases of emission permits (in cases with cap-and-trade regulations), this leaves the producers with a wide range of tools for complying with environmental regulations. Unlike the modeling approaches that were reviewed in section 2, I take the relative costs of the different compliance strategies into account when I now propose a new approach to marginal abatement cost estimation.

4. Marginal abatement costs

My starting point is the firm’s technical possibilities at a given point in time, summarized by the technology set. In the case where the firm is only concerned with its technical possibilities to convert inputs into desirable or intended outputs and neglects its impact on the environment, the technology can formally be defined by:
\[ T_1 = \{(x, y) : x \text{ can produce } y\} \quad (11) \]

1. \( T_1 \) is a representation of the firm’s possible input-output choices in absence of environmental regulation. It is assumed to satisfy the standard axioms of inactivity and no free lunch, free disposability of inputs and outputs, and that \( T_1 \) is a non-empty, closed, and convex set. See Chambers (1988) for a discussion of these properties. They guarantee the existence of cost functions, revenue functions, and profit functions that describe the firm’s optimal allocations in absence of environmental regulation, i.e. they describe the Business as Usual scenario.

Following Rødseth (2011), environmental regulations are introduced in the form of restrictions on the firm’s access to \( T_1 \). For example, if an electricity plant is forced to reduce its sulfur dioxide emissions it may be prevented from using high-sulfur fuels. Clearly, if the profit maximizing input mix in \( T_1 \) comprises high-sulfur fuels, the plant will experience an economic loss because of environmental legislation. In order to capture these dynamics, the “environmentally regulated technology” is modeled as the intersection of two sub-technologies (Krysiak and Krysiak, 2003; Murty et al., 2012) – \( T_1 \) and the materials balance condition -, and can be viewed as an extension to Førsund’s (2009) approach to polluting technologies. Førsund’s approach is based on the production theory of Frisch (1965), which shows that the producer’s freedom to choose the output mix for given inputs is related to the number of sub-technologies (or relationships between inputs and/or outputs) that make up the production model. A great deal of flexibility in selecting the output mix for given inputs is not consistent with equation 9, especially in the case where recuperation factors are zero and uncontrolled emissions are contingent on input use. A model consisting of multiple sub-technologies exhibits reduced degree of freedom in selecting the output mix, purposely making pollutants undesirable byproducts with limited substitutability.

Let \( T_2(b) \) define the set of inputs and desirable outputs that by the materials balance condition are feasible given \( b \). When \( b \) reflects the legal constraints the producer is facing on emissions, \( T_2(b) \)
is the set of inputs and desirable outputs that comply with existing environmental regulations. $T2(b)$ must be bounded for finite $b$ (unless the vector of emission factors is the zero vector), which implies that environmental regulations restrict the access to technology $T1$. That is, $T(b) \subseteq T1$, where $T(b)$ is the polluting technology:

$$T(b) = T1 \cap T2(b),$$

$$T1 = \{(x, y) : x \text{ can produce } y\}$$

(12)

$$T2(b) = \{(x, y) : ux - vy \leq b\}$$

My model approach builds on indirect production theory (Shephard, 1974), which considers restricted access to $T1$ due to cost or revenue constraints. Lee and Chambers (1986) and Färe et al. (1990) extended this theory to consider profit maximization when the producer faces expenditure constraints that prevent him from operating economically optimal. In my setting, this translates to economic losses due to legal constraints on emissions.

Notice that equation 12 constraints the producer’s uncontrolled emissions to be less or equal to the emission constraint $b$. However, it is usually the controlled emissions, and not the uncontrolled emissions, that are under legislation. Pollution control thus allows relaxing the emission constraint that the firm is operating under, since $T2(b) \subseteq T2(b + a_c)$, where $T2(b + a_c) = \{(x, y) : ux - vy \leq b + a_c\}$, and consequentially, $T(b) \subseteq T(b + a_c) \subseteq T1$. Second, if a market for emission permits exists, the firm may relax the emission constraint by purchasing permits. That is, $T(b) \subseteq T(b + a_c) \subseteq T1$, where $a_c \in \mathbb{R}_+$ is the increase in legal emission due to emission permits. Pollution control or permits are profitable if the relaxation of the emission constraint contributes to increases in profits that exceed the costs of controls or permits.

Figure 1 provides a graphical representation of the production possibilities of a firm under environmental regulation. It considers the case of coal inputs, $x$, used to produce electricity, $y$, along with sulfur dioxide emissions, $b$. The materials balance condition for sulfur dioxide is represented
pictorially by the lines \( ux \) (uncontrolled emissions) and \( ux-a_c \) (controlled emissions) in the figure’s lower panel.

The firm, which faces legal restrictions on its sulfur emissions, is implicitly restricted in terms of coal consumption. When complying with the emission target \( b' \) it can only consume input bundles which lie in the shaded area of figure 1, below or equal to the input quantity \( x'^{uc} \), as long as the firm is not involved in pollution control activities or buys emission permits. The firm is thereby prevented from maximizing its profits, since it cannot use \( x' \) units of coal (such that the value of the coal’s marginal productivity equals its market price). However, pollution control and emission permits allow the producer to extend the use of coal beyond the shaded area without violating the sulfur regulation. Consider for example the line \( ux-a_c \) in figure 1. By choosing the control level \( a_c \), the constraint for coal use changes to input quantity \( x'^c \). This is a profitable choice if 1) the increase in profits by approaching \( x' \) exceeds the cost of pollution control and 2) the cost of pollution control is less than the cost of emission permits. If these requirements are met the firm will maximize profits.

Figure 1: The polluting technology
by employing pollution control so that the marginal increase in profits from extended access to technology $T1$ equals the marginal costs of pollution control.

This paper identifies marginal abatement costs by exploiting the common knowledge that environmental regulations force deviations from Business as Usual allocations (i.e. allocation $x'$ in figure 1). By using data to estimate directional distance functions, shadow prices of inputs and outputs can be computed and compared to their market prices. Assuming that each firm maximizes profits under environmental regulations, the deviations between shadow prices and market prices allow identifying marginal abatement costs.

Because of different treatments of pollution control in the literature I consider profit maximization for two different model specifications. The first specification considers desirable outputs to be jointly produced with pollution control (Coelli et al., 2007; Murty et al., 2012) while the second specification considers pollution control to be non-joint with the production of desirable outputs (Førsund, 2009). The implications of joint and non-joint pollution control are treated in section 4.3.

4.1. Joint pollution control

The first model specification assumes that the input vector $x$ can be employed both to desirable outputs and pollution control:

$$T1 = \{(x, y, a_c) : x \text{ can produce } (y, a_c)\}$$

(13)

Since the output vector (which includes the pollution control output) comprises outputs that are desirable for the firm by generating revenue and reducing controlled emissions, the standard axioms are assumed to hold. In particular, outputs are assumed to be freely disposable, which means that increases in one output take place at the expense of at least another output when the production is
technically efficient. Define the direction vector \( \mathbf{g} = (-g_x, g_y, g_c) \), where \( g_x \in \mathbb{R}_+^N \), \( g_y \in \mathbb{R}_+^M \), and \( g_c \in \mathbb{R}_+ \), and a suitable function representation for the technology with joint pollution control:

\[
\overline{D}(x, y, a_c; -g_x, g_y, g_c) = \sup \{ \beta \in \mathbb{R} : (x - \beta g_x, y + \beta g_y, a_c + \beta g_c) \in T_1 \}
\]

(14)

Under the standard axioms, the directional distance function completely characterizes technology \( T_1 \). Let \( p_c \in \mathbb{R}_+ \) be the market price for the pollution control output. I make the following assumption:

Assumption (A1): The market price for the pollution control output, \( p_c \), is zero.

Assumption (A1) simply states that the pollution control output has no market value. In my setting, pollution control’s sole purpose is to increase the producer’s access to technology \( T_1 \). That is, to increase his ability to employ polluting inputs while simultaneously comply with regulatory constraints. Emission permits serve the same purpose. Let \( p_z \in \mathbb{R}_+ \) denote the permit price, let \( a \in \mathbb{R}_+^2 \) be the vector that comprises pollution controls and emission permit purchases, and let \( p \in \mathbb{R}_+^2 \) be the corresponding price vector. The profit maximization problem for a producer that faces environmental regulations under assumption (A1) is then defined:

\[
\pi(r, w, p, b) = \sup_{x, y, a} \left\{ ry - wx - p_c a_c : \overline{D}(x, y, a_c; -g_x, g_y, g_c) \geq 0, ux - vy - ea \leq b \right\}
\]

(15)

Equation 15 concerns the duality between the directional distance function and the profit function. When factors are fixed in the short run the duality between the short-run directional distance function and the short-run profit function (Blancard et al., 2006) can be considered.
where \( e \) is the vector with all its elements equal to 1. The necessary first order conditions for profit maximization are:

\[
\lambda_1 \nabla_y \bar{D} \geq r + \lambda_2 v, \quad \bar{y} \left[ r + \lambda_2 v - \lambda_1 \nabla_y \bar{D} \right] = 0 \quad (16)
\]

\[
\lambda_1 \nabla_x \bar{D} \geq - (w + \lambda_2 u), \quad \bar{x} \left[ - w - \lambda_2 u - \lambda_1 \nabla_x \bar{D} \right] = 0 \quad (17)
\]

\[
\lambda_1 \partial \bar{D} / \partial c \geq \lambda_2, \quad \bar{a}_c \left[ \lambda_2 - \lambda_1 \partial \bar{D} / \partial c \right] = 0 \quad (18a)
\]

\[
p_e \geq \lambda_2, \quad \bar{a}_c \left[ \lambda_2 - p_e \right] = 0 \quad (18b)
\]

\[
\bar{D} \geq 0, \quad \lambda_1 \bar{D} = 0 \quad (19)
\]

\[
ux - vy - ea \leq b, \quad \lambda_2 \left[ b - ux + vy + ea \right] = 0 \quad (20)
\]

where \( \lambda_1 \) and \( \lambda_2 \) are Lagrangian multipliers.

The first order condition 20 states that the shadow price on the emission constraint is zero if the constraint is not binding. The profit maximum is then equal to the Business as Usual profit maximum, which means that the producer is economically unaffected by the environmental regulation. A binding emission constraint, on the other hand, leads to forgone profits relative to the Business as Usual scenario. Intuitively, forgone profits are caused by reduced possibilities to employ polluting inputs.

The need for input and output substitution or downscaling of operations is counteracted by pollution control or emission permits. Consider the Lagrangian multiplier \( \lambda_2 \) which reflects the economic benefits from relaxing the emission constraint. The first order conditions from equation 18 state that \( \lambda_2 \leq \lambda_1 \partial \bar{D} / \partial c \) and \( \lambda_2 \leq p_e \), where \( \lambda_1 \partial \bar{D} / \partial c \) is the shadow price of pollution control.

The two first order conditions compare the marginal costs of pollution control and permit purchases to \( \lambda_2 \). Pollution control and permits are viable compliance strategies if their marginal costs are less
or equal to the marginal economic benefit from relaxing the emission constraint, i.e. their costs are
less or equal to the opportunity cost of input and output substitution or downscaling of operations.
Notice that the first order conditions from equation 18 perceive pollution control and permits as
perfect substitutes for relaxing the emission constraint, which means that the least costly approach
will be preferred. If for example 1) the increase in profits (from increased access to $T_1$) by pollution
control exceeds the cost of pollution control and 2) the cost of pollution control is less than the cost
of emission permits, then profits are maximized by employing pollution control until its marginal
costs equal its marginal benefits, $\lambda_2$.

Considering interior solutions, the first order conditions for desirable outputs $m$ and $m'$ from
equation 16 can be combined to define the optimum condition:

$$
\frac{r_{m'} + \lambda_2 y_{m'}}{r_m + \lambda_2 y_m} = \frac{\partial D(x, y, a_c; -g_x, y, g_c)/ \partial y_m'}{\partial D(x, y, a_c; -g_x, y, g_c)/ \partial y_m} \quad (21)
$$

Similarly, using equation 17 the corresponding optimum condition can be derived for inputs $n$ and $n'$:

$$
\frac{w_n' + \lambda_2 u_n'}{w_n + \lambda_2 u_n} = \frac{\partial D(x, y, a_c; -g_x, y, g_c)/ \partial x_n'}{\partial D(x, y, a_c; -g_x, y, g_c)/ \partial x_n} \quad (22)
$$

Equation 21 and 22 state that relative prices of desirable outputs and inputs equal the
corresponding ratio of distance functions derivatives in optimum. Notice that output prices, $r + \lambda_2 y$, and input prices, $w + \lambda_2 u$, deviate from their market prices by $\lambda_2 y$ and $\lambda_2 u$, respectively. The
vector $\lambda_2 y$ determines reductions in compliance costs due to increased recuperation of emissions
by a marginal increase in desirable outputs. The vector $\lambda_2 u$ represents increases in compliance costs due to increases in uncontrolled emissions following a marginal increase in inputs.

Equation 21 and 22 allow estimating marginal abatement cost. The directional distance function from equation 14 can be estimated by applying quantities of inputs and desirable outputs. The ratio of distance function derivatives are further calculated as in Färe et al.'s (1993; 2005) procedure. The marginal abatement cost, $\lambda_2$, can thus be obtained from equation 21 or 22, when prices and emission factors are known.

In the efficiency measurement literature, equation 21 and 22 are considered expressions of allocative inefficiency since the ratio of shadow prices (the right hand sides of the equations) deviate from the ratio of market prices. Intuitively, equation 21 and 22 rationalize allocative inefficiency for producers that comply with an environmental regulation: the regulations raise the costs of polluting inputs relative to their market prices, simply because their related emissions require costly compliance strategies, e.g. permit purchases. As a consequence, the input mix that solves the emission constrained profit problem will employ a less amount of polluting inputs relative to the Business as Usual profit maximum. For outputs, the sales prices increase in the recuperation factors as the need for other compliance strategies becomes less when emissions are recuperated by intended outputs. The optimal output mix is thus composed of outputs with high recuperation factors, compared to the optimal output mix under Business as Usual.

Next, I evaluate the marginal abatement costs. Solving equation 22 with respect to $\lambda_2$ one obtains:

$$\lambda_2 = \left( MS \ast w_n \right) - w_n' \left/ \left( MS \ast u_n \right) - (MS \ast u_n)' \right., \text{ where } MS = \frac{\partial \tilde{D}(x, y, a_c; -g_x, g_y, g_c)}{\partial x_{n'}} / \frac{\partial \tilde{D}(x, y, a_c; -g_x, g_y, g_c)}{\partial x_n}. \quad (23)$$
The numerator in equation 23 defines the marginal costs of substituting factor \( n' \) with \( n \): The marginal rate of technical substitution, \( MS \), determines the increase in factor \( n \) that is sufficient to compensate a marginal reduction in factor \( n' \). The product \( MS \cdot w_n \) thus determines the increase in costs related to the increase in the use of input \( n \). Subtracting the reduced unit costs for \( n' \) provides the costs of substitution.

The denominator in equation 23 defines the marginal change in emissions caused by substituting factor \( n' \) with factor \( n \). This follows the same arguments as for the substitution costs. Marginal abatement costs are, in other words, determined by weighing the costs of input substitution against the change in environmental damage that must be offset by some costly compliance strategy in order not to violate the environmental regulation. A similar interpretation can be provided for equation 21 in terms of output transformation. By accommodating for flexible producer responses, abatement costs are now defined in terms of the least cost way of compliance. My model approach is therefore likely to produce more reliable estimates of marginal abatement costs compared to the model approach from section 2, since the latter only considers one producer response to environmental regulations.

4.2. Non-joint pollution control

Up until now I have allowed for full flexibility in choosing the output mix of pollution control and desirable outputs for a fixed input vector. However, inputs that are employed to pollution control may differ from those employed to produce desirable outputs. Assume that the input vector \( x \) can be cracked into inputs employed to desirable outputs, \( x_y \), and inputs employed to pollution control, \( x_c \). The technology then comprises two sub-technologies, one which converts “production inputs” into desirable outputs and one which converts “pollution control inputs” into the pollution control output. I assume that both technologies satisfy the standard axioms, which allow me to represent them by directional distance functions:
Let $w_y$ and $u_y$ be the vectors of input prices and emissions factors related to desirable outputs, while $w_c$ and $u_c$ are related to pollution control. I assume that pollution control does not contribute to increases in the regulated pollutant. This assumption, which is formalized by (A2), does not rule out that pollution control may increase other pollutants that are not under regulation:

**Assumption (A2):** The vector of emission factors for pollution control inputs, $u_c$, is the zero vector.

Under assumptions (A1) and (A2), the profit maximization problem for a producer that complies with environmental regulations is:

$$
\pi (r, w, p, b) = \sup_{x, y, a} \left\{ \frac{r - px}{w_x - p_x c_x - c} : \bar{D}^{-c} \leq 0, \frac{r - px}{w_x - p_x c_x - c} - \inf_{x} \left\{ w_x : \bar{D}^{-c} \leq 0 \right\} \right\}
$$

$$
= \sup_{x, y, a} \left\{ \frac{r - px}{w_x - p_x c_x - c} : \bar{D}^{-c} \leq 0, \frac{r - px}{w_x - p_x c_x - c} - \inf_{x} \left\{ w_x : \bar{D}^{-c} \leq 0 \right\} \right\}
$$

$$
= \sup_{x, y, a} \left\{ \frac{r - px}{w_x - p_x c_x - c} : \bar{D}^{-c} \leq 0, \frac{r - px}{w_x - p_x c_x - c} - \inf_{x} \left\{ w_x : \bar{D}^{-c} \leq 0 \right\} \right\}
$$

See Pethig (2006) for a discussion on cases where pollution control contributes to additional emissions.
where $c(w_c, a_c)$ is the cost function that defines minimal costs of pollution control given the prices for pollution control inputs.

The corresponding first order conditions for profit maximization are:

\begin{align}
\lambda_1 \nabla_y \tilde{D}^y &\geq r + \lambda_2 v, \\
\bar{y} \left[ r + \lambda_2 v - \lambda_1 \nabla_y \tilde{D}^y \right] &= 0 \tag{26} \\
\lambda_1 \nabla_x \tilde{D}^y &\geq -(w_y + \lambda_2 u_y), \\
\bar{x}_y \left[ -w_y - \lambda_2 u_y - \lambda_1 \nabla_x \tilde{D}^y \right] &= 0 \tag{27} \\
\partial c/\partial a_c &\geq \lambda_2, \\
\bar{a}_c \left[ \lambda_2 - \partial c/\partial a_c \right] &= 0 \tag{28a} \\
p_z &\geq \lambda_2, \\
\bar{a}_z \left[ \lambda_2 - p_z \right] &= 0 \tag{28b} \\
\bar{D}^y &\geq 0, \\
\lambda_1 \bar{D}^y &= 0 \tag{29} \\
u_y x_y - vy - ea &\leq b, \\
\lambda_2 \left[ b - u_y x_y + vy + ea \right] &= 0 \tag{30}
\end{align}

The major difference between first order conditions 26-30 and 16-20 relates to equations 18a and 28a, the first order conditions for pollution control. For the other first order conditions, differences relate only to the inclusion or non-inclusion of pollution control as an element in the directional distance function, and to the partition of the input vector. Expressions similar to equations 21 and 22 may therefore equally be calculated from equations 26 and 27, and estimates of marginal abatement costs – reflecting the least costly way of reducing pollution - may be identified by the procedure described in section 4.1.

Finally, note that pollution control may be unavailable or not economically viable in certain industries. Pollution control should not be included in the model in such cases, since it is not a feasible compliance strategy. The relevant profit maximization problem without pollution control corresponds to equation 25 when the cost function for pollution control and the pollution control output in the emission constraint are omitted. The first order conditions for profit maximization are
equivalent to equations 26-27, 28b-30, which means that optimum conditions similar to equations 21 and 22 may also be derived in this case.

4.3. The impact of joint and non-joint pollution control

Looking beyond the data issues treated in section 5.2, it is important to recognize that treating pollution control as a joint or non-joint activity have implications for the trade-off between the pollutant and desirable outputs. Starting with joint pollution control, I combine the first order conditions 16 and 18a to derive the trade-off between the pollutant and the desirable output $m$ in the case where pollution control is an economically viable compliance strategy (i.e. the first order condition 18a holds with equality):

$$\frac{\lambda_2}{r_m} = \frac{\partial \tilde{D}}{\partial a_c} \left( \frac{1}{1 - \nu_m \left( \frac{\partial D}{\partial a_c} / \frac{\partial D}{\partial y_m} \right)} \right) = \frac{\partial y_m}{\partial a_c} \frac{\partial a_c}{\partial b} = \frac{\partial y_m}{\partial b}$$

Equation 31 is a conventional optimum condition where the relative price equals the marginal rate of transformation. Since the marginal reduction in the pollutant here takes place by pollution control, the trade-off between the pollutant and the desirable output is the product of two related trade-offs; the first, $\left( \frac{\partial \tilde{D}}{\partial a_c} / \frac{\partial D}{\partial y_m} \right)$, defines the technical trade-off between the pollution control output and the desirable output in terms of the reduction in the desirable output required to increase pollution control marginally (for fixed inputs), while the second trade-off defines the increase in pollution control necessary to reduce the pollutant $b$ marginally. The latter depends on output $m$’s recuperation of the pollutant. With no recuperation, there is a one-to-one relationship between pollution control and emission reductions while with recuperation, the gains from increased pollution control, in terms of reduced emissions, are partly crowded out by forgone
recuperation due to the reduction in output $m$. The numerator of the second trade-off defines the gross reduction in emissions from increased pollution control, while the denominator defines the reduction in emissions net of recuperation losses.

Since $\frac{\lambda_2}{r_m}$ is greater or equal to zero by non-negative prices, equation 31 implies that the model with joint pollution control produces a non-negative trade-off between the pollutant and desirable outputs (under profit maximization). The non-negative trade-off is in line with the traditional approaches to polluting technologies from section 2, i.e. some desirable outputs must be forgone to reduce the pollutant, because pollution control takes place at the expense of producing desirable outputs. This result is also shown by Murty et al. (2012). Notice that in order for the result to be relevant, pollution control must be the least costly compliance strategy. Considerations about the relative costs of different compliance strategies are not taken into account by Murty et al. (2012).

In the case with non-joint pollution control, pollution control cannot take place at the expense of desirable output for a fixed input vector $x = \left( x_y, x_c \right)$ because pollution control inputs cannot be employed to producing desirable outputs and vice versa. Now, the trade-off between the pollutant and desirable outputs can be positive or negative. To see this, notice first that when inputs are fixed, the uncontrolled emissions related to input use, $u_y x_y$, must be fixed accordingly. Hence, emissions can only be reduced by producing more “high-recuperating” desirable outputs for given inputs (assuming that pollution control is at its optimal level). To study the relationship between the pollutant and a desirable output in this setting I rewrite equation 21:

$$\lambda_2 = \frac{(MS \cdot r_m) - r_{m'}}{v_{m'} - (MS \cdot v_m)} \text{ where } \frac{\partial D(x, y; -g_x, g_y)}{\partial y_{m'}}/\frac{\partial y_{m'}}{\partial y_m}$$

(32)
Similar to equation 23, equation 32 states that the marginal abatement costs, $\lambda_2$, equal the ratio of the change in revenue to the change in the pollutant by substituting output $m'$ with output $m$. Here, $v_{m'}$ defines the increase in emissions from a marginal reduction in output $m'$ (i.e. losses in recuperation) while $\left( MS^* v_m \right)$ defines the reduction in emissions due to increases in output $m$ (i.e. gains in recuperation). Clearly, if $v_{m'} \geq \left( MS^* v_m \right)$ there is a non-negative relationship between the pollutant and the desirable output $m$ (for fixed inputs). However, if $v_{m'} < \left( MS^* v_m \right)$, the relationship between the pollutant and the desirable output $m$ is negative. Notice that $\lambda_2 \geq 0$ by equation 30. This means that the numerator in equation 32 must be non-negative when $v_{m'} \geq \left( MS^* v_m \right)$ and non-positive when $v_{m'} < \left( MS^* v_m \right)$. This is an intuitive result which implies that additional reductions in the pollutant by output substitution lead to additional revenue losses, thus making it costly to comply with environmental regulations.

To sum up, the common assumption in the literature, that reductions in pollution takes place at the expense of desirable outputs (for given inputs), is in line with joint pollution control. The traditional models are thereby possibly intended for industries where pollution control is the least costly compliance strategy and where inputs are allocatable. Clearly, this limits the scope for empirical studies. When pollution control is non-joint or when it is not the least costly compliance strategy, the trade-off between the pollutant and desirable outputs can be both positive and negative. A negative trade-off is not consistent with modeling the pollutant as a freely disposable input, while Färe et al.’s model (1993; 2005) enables the trade-off to be both positive and negative (Kuosmanen and Kazemi Matin, 2011).

5. Discussion

The previous section established shadow pricing approaches that allow identifying marginal abatement costs. The method proposed has three main advantages over the shadow pricing
approach from section 2: First, there is no need for introducing non-standard axioms because the relevant outputs – pollution control and intended outputs – are not undesirable. The method is therefore applicable to all cases in which the usual axioms of production analysis are assumed to hold, including the before mentioned examples from electricity generation, agriculture, and aquaculture. Second, by applying the materials balance principle, the polluting technology explicitly accounts for underlying factors of pollution generation and separates uncontrolled emissions from controlled emissions. The approach is clearly in line with the materials balance condition and avoids the black-box treatment of pollutants in the literature. Third, the model allows evaluating the various ways in which a producer can reduce undesirable byproducts and to identify the least costly way of achieving these reductions.

The current paper derives marginal abatement costs within a static model framework that considers the impact of environmental regulations on firms’ potential to exploit their technical possibilities at a given point in time. Clearly, the impact of environmental regulations is likely to be dynamic, in the sense that new production- or pollution control technologies emerge over time. Investment costs related to regulatory compliance represent forgone profits that can be labeled abatement costs. The impact of new investments is captured by their implications for firms’ production possibilities and, hence, the technology set. Following the reasoning in section 4, a firm is likely to undertake a new investment when the economic benefits outweigh the investment costs. The benefits of new investments are reflected by intertemporal changes in the estimates of shadow prices for inputs and outputs following changes in the state of the technology, and their implications for marginal abatement costs.
5.1. Multiple pollutants

Production often causes multiple pollutants and producers face several regulations. Färe et al.’s (1993) model is a multi-output model that can easily be extended to settings with multiple pollutants. See Färe et al. (2012) for an example.

The natural way of dealing with multiple pollutants in the technologies from section 4 is to introduce materials balances for each pollutant. If there are K regulated pollutants, it means that K-1 additional emission constraints must be appended to the profit maximization problems in equations 15 and 25. For example, in the case with non-joint pollution control, the first order conditions for the regulated profit problem, corresponding to equation 22, are extended to:

\[
\frac{w_n + \sum_{k=1}^{K} \lambda_{2k} u_{nk}}{w_o + \sum_{k=1}^{K} \lambda_{2k} u_{nk}} = \frac{\partial D(x,y;g_x,g_y)}{\partial x_n} \]

(33)

Notice that, if the number of pollutants exceeds one, there is a problem with identifying each of the K abatement costs. However, in the case with two regulated pollutants, and where an allowance price exists for one of the two pollutants, the shadow price of the residual pollutant can be determined by assuming that the known allowance price equals the marginal abatement costs. In the case with more than two pollutants, each marginal abatement costs can no longer be directly determined from the derivatives of the distance function. Alternative procedures, such as estimating both the distance function and shadow prices directly, have to be employed.

---

\(^6\) Equation 33 resembles equation 38 in Førsund (2009).
### 5.2. Computational approaches

First, I consider data accessibility and its implications for modeling joint and non-joint pollution control. In terms of data requirement, it is clear that estimating \( \overline{D}(x, y; y^- g_x, g_y) \) is more convenient than estimating \( D(x, y; a_c; g_x, g_y, g_c) \) since that latter is more data intensive. Data on the pollution control output is usually not readily available, but must be estimated separately. This makes \( D(x, y; a_c; g_x, g_y, g_c) \) more vulnerable to data error since the estimates of the pollution control output may differ from the true output. For inputs, it is likely that data on “production inputs” are more easily accessible than on “pollution control inputs”. However, in some cases it may be difficult to distinguish inputs used for production purposes from inputs used for pollution control, or data sources may simply merge the inputs used for production and pollution control purposes into one category. Using data on inputs related to both desirable outputs and pollution control to estimate \( \overline{D}(x, y; y^- g_x, g_y) \) leads to overestimation of the distance function (or inefficiency), since inputs related to pollution control are assumed to be unproductive (Färe et al., 2001); see for example Shadbegian and Gray (2005) who distinguish production inputs from pollution control inputs in order to show that pollution control inputs contribute little or nothing to intended production in their empirical study. Having identified possible pitfalls related to data accessibility, I conclude that empirical testing should guide the decision on the empirical model specification.

Emission factors are essential for identifying marginal abatement costs and the pollution control output. They are available from a wide range of sources, including the intergovernmental panel on climate change’s (IPCC) emission factor database, the U.K. National Atmospheric Emissions Inventory, and AP 42 Compilation of Air Pollutant Emission Factors. The emission factors relate the quantity of an undesirable byproduct released to the environment with a certain activity. Examples include air pollutant and greenhouse gas emissions from the combustion of fossil fuels. The factors...
are usually defined as the weight of pollutants divided by a unit weight, e.g. tons of sulfur dioxide emitted per tons of coal burned.

Uncontrolled emissions are calculated according to equation 9, by multiplying emission- and recuperation factors with the input and output vectors. Pollution control efforts can further be quantified if data on controlled emissions are available\(^7\). The pollution control output is calculated according to equation 10, by subtracting the controlled emissions from the estimated uncontrolled emissions. Data on pollution control is required for estimating the directional distance function from equation 14.

The estimation of a parametric distance function requires the selection of a functional form. Chambers (1998) suggests the quadratic functional form for the directional distance function, which has been followed up in the literature; see for example Färe et al. (2005; 2006). A programming method developed by Aigner and Chu (1968) is often applied to estimate parametric directional distance functions. Its drawback is that it does not account for random shocks that may affect the performance of firms. Stochastic frontier analysis (Aigner et al., 1977; Meeusen and van den Broeck, 1977) can be applied to address this problem.

6. Conclusion

This paper proposes a new approach to estimating marginal abatement costs. A model framework that utilizes the materials balance condition as the pollution generating mechanism is applied to address how undesirable byproducts come into existence, as well as the relative costs of different compliance strategies. Such dynamics are generally neglected in comparable studies, which suggest that they are in danger of misinterpreting the restrictions environmental regulations impose on firms. I evaluate the trade-off between pollutants and desirable outputs in the cases where the

\(^7\) One example where the required information on controlled and uncontrolled emissions is readily available is the case of U.S. electricity generation, where Continuous Emission Monitoring (CEMS) provides detailed data on regulated plants’ emissions of air pollutants.
production of desirable outputs is joint and non-joint with pollution control, and find that a non-negative trade-off is secured by joint pollution control. This finding indicates that other approaches to marginal abatement cost estimation consider pollution control as the most favorable compliance strategy. In my approach, the relative costs of different compliance strategies are the determinants of the most favorable strategy.

7. Acknowledgement

I am grateful to Eirik Romstad, Yohei Mitani, Yves Surry, Kjetil Gramstad, Finn R. Førsund, and three anonymous referees for their comments. The usual disclaimer applies.

8. References


