

## Chapter 8. Productivity measurement and the environment

Finn R Førsund

Department of Economics, University of Oslo

### 8.1. Introduction

The purpose of the chapter is to review and explore attempts to measure productivity when facing environmental problems caused by economic activities of production and consumption. The emphasis will be on how to include degradation of our natural environment by the discharge of pollutants from these activities. We are facing both regional and global negative effects, such as acid rain and global warming, the latter leading to sea level rise, forced human migrations, and adverse health effects such as diseases caused by the spread of mosquitos to new areas transmitting malaria and the zika virus.

Another aspect of environmental degradation not reflected in how conventional productivity is measured, is depletion of natural resources. The value of extracting or harvesting commercial resources is entered national accounts, but not the running down of stocks. A problem with the national accounts is that current market transactions only are included. The valuation of resource depletion and environmental degradation of public good aspects of Nature is absent. This chapter will focus on how to incorporate environmental degradation and not attempt to deal in depth with resource depletion.

Environmental problems had until the late 50ies been treated by economists as cases of externalities. The examples used had an innocent flair; Pigou (1913) used factory smoke dirtying washing hanging to dry outdoors as a negative externality, and Meade (1952) used the interaction of bees and apple blossoms as a positive externality. The publishing of the seminal paper Ayres and Kneese (1969) (see also Kneese et al., 1970; Kneese 1971), coining the phrase *materials balance*, heralded a new view within economics of the pervasiveness and seriousness of environmental pollution.<sup>1</sup> The conservation of mass (based on the first law of

---

<sup>1</sup> See Mishan (1971) for a review of the earlier externalities literature, and Fisher and Peterson (1976), and Cropper and Oates (1992) for reviews of the literature covering the 70-ies and 80-ies decades.

thermodynamics) tells us that matter cannot be created or disappear. If all the material inputs into an activity are not embedded in the products the activity is set up to deliver, then the difference must be contained in residuals discharged to the environment. In other words, if we weigh the inputs employed in an activity, and weigh the products that are the purpose of activities, the difference is the residuals (taking into consideration non-paid factors like oxygen from the air) that may turn out to be polluting the natural environment. Thus, the general feature of residuals is that they arise from use of material inputs in a wide sense. The concept of materials balance underlines the general inevitability of residuals generation and the pervasiveness of pollution when employing material resources. (The materials balance will be discussed in more detail in Section 8.3.2.) The materials balance principle leads naturally to a model of production activity (by firms or consumers) based on joint production of the intended goods of production and unintended products. The latter are in general termed residuals, but become pollutants if their discharge to the environment lead to degradations as evaluated by consumers.

The plan of the chapter is as follows. Section 8.2 introduces briefly the conventional way of measuring productivity as adopted when using data from the National Accounts. Macro issues such as dynamics of growth with pollutants, sustainability, green national product and the extension of input – output analysis to cover pollutants are touched upon. The introduction of physical satellite accounts for resource depletion and the environment is referred to. The concept of an environmental damage function is defined as the willingness to forsake man-made goods in order to sustain a certain quality of the environment. This cost term is a key in environmental economics. The effect on productivity of introducing environmental costs will then be shown in principle. In Section 8.3 the micro model used for studying the interaction of production and generation of pollution is introduced. The key assumption is that the use of material inputs in the production of intended (desirable) goods leads simultaneously to generation of residuals that may turn out to be polluting the natural environment. The model consists of two types of relations; a production function for the intended output, and a production function for the unintended product. This model satisfies the materials balance, and that there is no trade-off between desirable and undesirable outputs for given resources. Technical change is defined and end-of-pipe abatement is introduced as a separate activity. The social optimisation problem using a damage function representing evaluation of environmental degradation is set up and the impact of direct environmental regulation on productivity and profitability is shown. The Porter hypothesis is studied in Section 8.4. Basic regulatory

instruments such as the Pigou tax and cap and trade are studied. The key statements that environmental regulation represents a pressure on firms to reduce short-term inefficiency, and in the longer run leads to innovation reducing the generation of pollutants, are discussed. The impact on productivity change is commented upon. Productivity measures assuming persistent inefficiency of some firms compared with the contemporaneous best practice frontier function is discussed in Section 8.5. It is argued that the popular model in the inefficiency literature from the late 80ies is based on a model not satisfying the materials balance, and it is further argued that using this model should end. A summary of main results and conclusions is provided in Section 8.6 and indications of further research issues are offered.

## **8.2. Productivity measures**

### **8.2.1 Conventional productivity measures**

Productivity may be defined in several ways and used both in a macro and micro setting<sup>2</sup>. The most common definition is to use some measure of production per unit of labour used as an input in the activity in question; productivity can be measured simply as  $Y/L$  where  $Y$  is the output and  $L$  the input of labour. The output measures used based on the National Accounts are value added and gross production. Value added is the difference between revenue and variable cost excluding labour and capital, and gross production is measured as shipments, sales or revenue. Value added is the preferred measure at an aggregated level. In the national accounts it can be calculated from the supply side as defined above, or from the income side. Then value added is the sum of labour expenditures and remuneration to capital including depreciation and return on the assets (including owner income). At an industry level or lower levels of aggregation a multi factor measure of productivity based on gross output may be preferred. Then an index of inputs is used in the denominator. Gross output in the numerator also has to be represented by an index if there is more than one output.

We must distinguish between productivity and productivity change. The latter is the relative change in productivity from one period to another. The seminal definition of productivity

---

<sup>2</sup> See Chapters 2 and 3 in this volume for detailed accounts.

change in continuous time is the Solow residual (Solow 1956), where productivity change is the same as change in technology and defined as the growth in output that cannot be explained by the growth of inputs.

The labour productivity measure is a partial measure. Looking at the contribution of all factors of production a concept of multifactor productivity can be defined in the single output case:

$$TFP = \frac{Y}{F(x_K, x_L, x_E, x_M, x_S)} \quad (8.1)$$

where the  $F(.)$  function is an aggregator function, having positive partial derivatives and being linear homogenous, in the typical factors  $x_j$  ( $j=K,L,E,M,S$ ) with the sub-indices denoting production capital  $K$ , labour  $L$ , energy  $E$ , materials  $M$  and services  $S$ . In OECD and EU studies of production *KLEMS* is used as a term for these factors.<sup>3</sup> Multifactor productivity reflects the change in output that cannot be accounted for by the change in all the inputs. The best output measure corresponding to the use of all inputs is gross output. We can aggregate the individual outputs yielding revenue using an output aggregator function. A Törnqvist index of productivity on logarithmic form uses the revenue shares as weights in a linear aggregation of outputs and cost shares to aggregate inputs, using the behavioural assumption of revenue maximisation or cost minimisation.

### 8.2.2 Extending productivity measures to include the environment

The problem of pollution as a by-product of economic activity is a major topic in contemporary environmental economics, ranging from global warming due to generation of greenhouse gases, to local air-, water- and land quality deteriorations due to emission and discharge of a variety of polluting substances, as pointed out in Section 8.1. Other related issues both on the macro and the micro level should also be mentioned.

At the macro level the interest was focussed on the dynamics of growth including pollutants of the accumulating kind (Keeler et al. 1971; Plourde 1972; Smith 1972; d'Arge and Kogiku 1973;

---

<sup>3</sup> See Chapter 24 in this volume on the international *KLEMS* project.

Nijkamp and Paelinck 1973). The Pontryagin's maximum principle was the method adapted<sup>4</sup>. Resource depletion was also analysed at the aggregated level with this new mathematical tool following up on the early work of Hotelling in the 30ies. As stated in Section 8.1 this is an environmental problem of its own, but it will not be treated here.

The interest in sustainable growth led to some specific principle for utilising natural resources including environmental ones. Hartwick's rule (1977) for compensating resource depletion was to invest the resource rent. This was called genuine savings. Solow (1974) showed that one way to design a sustainable consumption program for an economy is to accumulate produced capital sufficiently rapidly. The pinch from the shrinking exhaustible resource stock can then be precisely countered by the services from the enlarged produced capital stock, given a sufficient degree of substitutability between produced capital and natural resources (see also Mäler 1991; Weitzman 1976 on the savings rule).

Since the construction of national accounts became common in most countries after the Second World War (through the work of the United Nation Statistical Division, UNSD) there has been a discussion about gross national product (GNP) (or gross domestic product; GDP) as a welfare measure (Stiglitz et al 2009). The blossoming of environmental economics also led to a wish to have the (mis)use of environmental resources included in a welfare measure as well as the running down exhaustible resources. Indeed, there is also a more recent interest in national accounts providing a measure of gross human happiness that would be a comprehensive welfare measure also including environmental concerns. (However, the only country that has attempted to make a human happiness account is Bhutan.)

The term Green national product (see e.g. Aaheim and Nyborg 1995; Asheim 2000) has been used for attempts to integrate national accounting with environmental considerations. A key concept is sustainable consumption that can be based on the Hicksian definition of sustainable income as consumption that can be had without degrading the resource base or environmental assets.

The attempt to integrate national accounts and emission of pollutants were also tried at a more detailed level of aggregation for input-output models (Leontief 1970; 1974; Leontief and Ford 1972). An abatement sector dealing with pollutants was introduced in the papers. The fixed

---

<sup>4</sup>The Maximum principle was published in 1956 in Russian and an English version in 1962, and the mathematical technique was applied to economic problems in the 60ies; a fast and widespread diffusion of an intellectual invention.

input-output coefficients were extended to include fixed emission coefficients for various pollutants calculated as emissions per unit of output. Recognising the role of material inputs, fixed coefficients related to outputs assume that there is fixed coefficients in production in general, as there are in the input-output model. Based on data for Norway, the costs of obtaining a “greener” mix of final deliveries for a given amount of primary inputs were shown in Førsund and Strøm (1976) and Førsund (1985).

However, it turned out to be too difficult to integrate environmental issues formally and statistically in the national accounts. Instead so-called Satellite accounts for natural resources, environmental resources and pollution emissions were established (UN 1993; Nordhaus and Kokkelenberg 1999). These accounts keep track of extraction and harvesting and development of stock of resources, and record emissions and change in various quality indicators of the natural environment.

We will limit our attention to the micro level of production, and for ease use a firm as the term for a production unit. The effect of environmental regulation on various aspects such as productivity and profitability in the short and long run will be studied.

The point of departure of this chapter is the way of modelling the interaction between economic activity and the natural environment within environmental economics (Førsund and Strøm 1987; Perman et al 2011). The standard productivity indexes used both on the macro level and on a micro level have to be extended to include environmental concerns.

How, then, may productivity measures be extended to include environmental assets? The general approach is to express environmental considerations in measuring units compatible with the way outputs are evaluated, i.e. money. Nature provides us with a number of amenities or services; clean air and water, wildlife, biodiversity, recreational possibilities, esthetic experiences, among many other forms of services. However, it should be realized that it is not meaningful to ask what is the value or price of such services. The point is that a valuation must be based on what we are willing to give up of man-made goods to enjoy such services at specific quality levels, and not based on any concept of intrinsic value.

We will narrow down our approach to investigate pollutants generated as by-product of economic activity. A key construct in environmental economics is to use a *damage function* monetising the degradation of the environment by discharging pollutants into Nature:

$$D(z) = D(z_1, \dots, z_R), \quad \frac{\partial D}{\partial z_i} \geq 0, i = 1, \dots, R \quad (8.2)$$

where  $z$  is a vector of  $R$  types of pollutants. The damage is measured as the value we are willing to give up (or demand in compensation) of produced goods, the production of which generates the pollutants as unintended products. Damage cost is used synonymously with environmental costs.

Nature has a spatial dimension. A damage function has to be detailed a lot more before we can talk about how to estimate the damage function. A distinction has to be made between local pollution from a single source and transport of pollutants from several sources to receptors in Nature where the damage is occurring. The latter form is called regional pollution. Damage depends on the location of the sources and the impacts of the emission from each source can in principle be identified. Acid rain is a typical example, and has been subject to large-scale modelling effort at IIASA initiated by the UN International Economic Commission for Europe (UNECE). Acid substances are transported via air from the sources, and where it is deposited on the ground or vegetation may be a long way from the source especially if high chimneys are used, or flue gases are ejected using pressurised air.

Another main type of damage function is emission to the air that mixes in the atmosphere and become global pollutants as climate gases causing climate change. The type of gas is then simply summed over all sources emitting this gas. Although the damages are local all sources contribute to these damages.

A technique developed within environmental economics is to measure the willingness to pay for preserving the present environmental qualities, or willingness to pay for an improvement. Necessary compensation demanded for accepting deterioration in environmental quality is also used. The damage function (8.2) is a theoretical construct commonly used in textbooks on environmental economics based on such willingness-to-pay concepts. Methods of estimation may be based on transport costs of visiting recreation sites, complementary equipment bought to enjoy recreation activities, or contingent valuation based on elaborated interview schemes (see e.g. Førsund and Strøm 1987; Perman et al 2011).

For a given level of aggregation (e.g. using the industrial classification system of national accounts) let  $p_i$  be the price and  $y_i$  the volume of desirable good  $i$ . Measuring the *social* gross output the damage cost of the environmental services caused by producing the goods must now be subtracted from revenue of desirable goods to be used in the numerator calculating the *social*

total factor productivity  $TFP_E$  :

$$TFP_E = \frac{\sum_i p_i y_i - D(z)}{F(x_K, x_L, x_E, x_M, x_S)} \quad (8.3)$$

In a static setting it is obvious that the social multi-factor productivity must go down because of the reduction in the numerator assuming no change in outputs and inputs; i.e.  $TFP_E < TFP$ . However, once environmental damage is included in the calculation of social gross output in a period predating period  $t$ , the productivity change may go either way:

$$\Delta TFP_E^{t,t+1} = \frac{\sum_i p_i y_i^{t+1} - D(z^{t+1})}{F(x_K^{t+1}, x_L^{t+1}, x_E^{t+1}, x_M^{t+1}, x_S^{t+1})} / \frac{\sum_i p_i y_i^t - D(z^t)}{F(x_K^t, x_L^t, x_E^t, x_M^t, x_S^t)} \begin{matrix} \geq 1 \\ < \end{matrix} \quad (8.4)$$

Prices are kept fixed in the numerator and damage cost must also be deflated by a price index to become a volume index. The direction of productivity change depends on the development of relative terms including environmental damage and may increase as well as decrease. If outputs and inputs used in production change over time, the scale properties of the underlying production function will play a role. Over time it is reasonable to assume change in the volume indices due to change in output- and input mixes. The damage function may also shift to show increased damage for the same level of pollution. In that case productivity change may be negative. However, damages may go down over time due to reduced discharge of pollutants, cf. the effort to decarbonise societies.

But we must not forget the macro effects due to resources re-allocation to reducing pollution, restoration of natural environment, and resources in general used on mitigation and abatement. There are also effects of lowering investment in polluting activities reducing their growth and may be reducing the pace of capital accumulation, and thus cause reduced future productivity growth. Research and development resources may be redirected to activities reducing the growth of economies.

### 8.3. Modelling production and the environment

#### 8.3.1. Introduction



When modelling environmental – economic interactions it is important to capture the main features of such interactions within a model small enough not to becoming unmanageable by trying to incorporate the numerous real life details. Adhering to the principle of Ockham’s razor as to the size of a model, we will develop such a model based on a multi output multi input production model termed “*Factorially determined multi-output system*” in Frisch (1965). This model will be termed the *Frisch pollution model* for short.<sup>5</sup>

### 8.3.2. The Frisch pollution model

#### *A multi-equation system*

Pollution is generically a problem with joint outputs in economic activities of production and consumption. As pointed out in Section 8.1 there is a materials balance that accounts for where the mass contained in material inputs end up; in the desirable output or in the natural environment. If all the material inputs into an activity are not embedded in the products the activity is set up to deliver, then the difference must be contained in residuals discharged to the environment. It seems important to capture these physical realities from use of material inputs in a modelling of the interaction economic activity and generation of pollutants. It will then be clarifying to distinguish between input factors  $x_M$  with material content (raw materials) being affected physically by the production process and factors unchanged (not used up) by the production process, the main types of the latter being labour, capital and external services, termed service inputs  $x_S$ . The *Factorially determined multioutput model* seems tailor-made for capturing the physical process of generation of residuals and the linkages to production of outputs (a unit index is suppressed for simplicity):

$$\begin{aligned} y &= f(x_M, x_S), f'_{x_M} > 0, f'_{x_S} > 0 \\ z &= g(x_M, x_S), g'_{x_M} > 0, g'_{x_S} < 0 \end{aligned} \tag{8.5}$$

Here  $y$  is the desirable output that it is the purpose to produce setting up the activity,  $x_M$  and  $x_S$  are material inputs and service inputs, respectively, and  $z$  is the unintended residual generated simultaneously utilising inputs to produce output. If we have  $D(z) > 0$  the residual is called undesirable or a pollutant. This will be assumed to be the case in the following. The multiple equation system of Frisch consists of a production function  $f(\cdot)$  for each desirable output as a

---

<sup>5</sup> The adaptation of the Frisch model to environmental pollution is a further development of the models presented in Førsund (2009); (2017).

function of the same bundle of inputs, and a production function  $g(.)$  for the generation of each residual as a function of the same set of inputs. The functions  $f(.)$  and  $g(.)$  are assumed to be efficient in the sense that  $y$  is maximised for a given bundle of inputs, while the residual  $z$  is minimised for the same given bundle of inputs, so equalities are used in (8.5)<sup>6</sup>. The system of equations (8.5) is a drastic simplification of engineering realities, but still captures the most essential feature of the type of joint production of goods and residuals; the crucial connection between these two types of outputs goes through the use of material inputs. The environmental problem of generation of residuals is that emitting them to the natural environments may create negative externalities due to degradation of environmental qualities or in general creating harmful effects. As defined above residuals are then called pollutants. To keep the model as simple as possible we consider a single desirable output  $y$  and an undesirable output  $z$  (generalising to multi output and multi pollutants can be done in the Frisch model just by adding more equations, one for each variable, keeping the same inputs as arguments in all relations, see Førsund 2009). The material inputs are fossil primary energy (in the form of e.g. coal, oil, gas and wood), and various types of raw materials. The service inputs are those that are not used up in the production process in a material sense, remaining physically intact and providing services such as labour and capital. Electricity used as input is in our context also a service input since it does not have any weight. The production of desirable outputs and undesirable residuals occur simultaneously and based on the technology of multiple output production. The separation into two types of equations does not mean that we have two technologies, but is done as a very helpful simplification of a possible complex technology without sacrificing key aspects of a technology such as identifying substitution between inputs, scale properties and the connection between desirable and undesirable outputs.

The partial derivatives (the productivities) of the desirable output function are assumed to have the usual positive signs. However, concerning the production function for the undesirable output the two types of inputs are assumed to have opposite signs: marginal increase in material inputs increases the undesirable output, while marginal increases in service inputs decreases the undesirable output. The positive partial productivity of service inputs in the desirable output production function and the negative sign in the residuals generation function can be explained

---

<sup>6</sup> In the case of inefficiency introduced in Section 8.5 the functions  $f(.)$  and  $g(.)$  will be the efficient border of the production possibility sets called frontier functions.

by the fact that more of a service input improve the utilisation of the given raw materials through better process control or increase internal recycling of waste materials.<sup>7</sup>

As a consequence of the typical way service inputs increase the output by utilising the material inputs more efficiently is that production of pollutants will then typically decrease. Therefore, the signing of the marginal productivity in the bads production function is negative.

The material inputs are *essential* in the sense that we will have no production neither of goods nor pollutants if  $x_M = 0$ ; <sup>8</sup>

$$y = f(0, x_S) = 0, z = g(0, x_S) = 0 \quad (8.6)$$

There will in general be substitution possibilities between material and service inputs. The rate of substitution evaluated at a point on an isoquant of the production function for the desirable output is  $(-f'_{x_M} / f'_{x_S}) < 0$ . This is the amount of material input that is reduced if the service input is increased with one unit, keeping output  $y$  constant. Considering several material inputs there may be substitution possibilities between them, e.g. between coal and natural gas, keeping the output constant, but decreasing the generation of bads if the marginal contribution of gas to creation of bads is smaller than the marginal contribution of coal;  $g'_{x_{coal}} > g'_{x_{gas}}$  for the same heat content..

The role of service inputs in the residuals production function is crucial as to the substitution effect of decreasing a material input and increasing a service input resulting in less generation of residuals for a constant production of the desirable output. The marginal rate of substitution is positive,  $(-g'_{x_M} / g'_{x_S}) > 0$  due to the marginal productivity of service inputs being negative.

This implies a special form of isoquants in the factor space and the direction of increasing residual level compared with a standard isoquant map for the output, as seen in Figure 8.1. The isoquants for the two outputs can be shown in the same diagram because the arguments in the functions are the same (see Frisch (1965, p. 272) or Førsund (2009, p.7) for the original illustration of isoquants in the factorially determined multioutput case). The level of the residual  $z$  is increasing moving South-East (see dotted black arrow) while the level of

---

<sup>7</sup> Cf. the famous chocolate production example in Frisch (1935), discussed in Førsund (1999), of short-run substitution between labour and cocoa fat due to more intensive recycling of chocolate with moulding defects the more labour that is employed. Moulding defects decrease with a higher proportion of cocoa fat. Stricter quality control using more labour reduces the share of defect products utilising the raw material more efficiently.

<sup>8</sup> One or more service inputs may also be essential, but the point is that residuals are in general an unavoidable feature using material inputs in production.

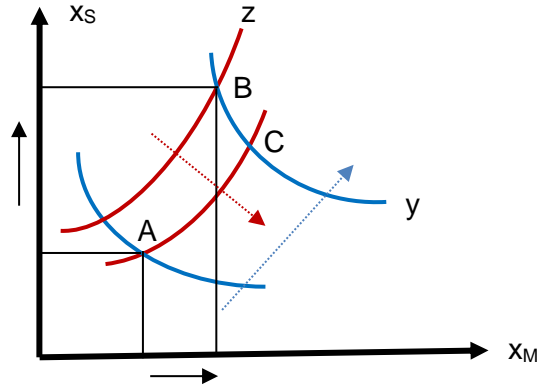


Figure 8.1. *Isoquants for the production of y and z*  
 Source: Førsund (2017)

the desirable good  $y$  is increasing moving North-East (see dotted blue arrow). Going from point A to point B in input space increasing both inputs, but non-proportionally, we see that the production of the residual  $z$  has decreased while the production of output  $y$  has increased.

The change in output mix and levels of  $y$  and  $z$  from A to B in Fig. 8.1 is due to changing the input mix and levels of  $x_M$  and  $x_S$  only. Going from B to C reducing the input  $x_S$  and increasing the input  $x_M$  the desirable output is constant, but the undesirable output has increased. The two types of outputs  $y$  and  $z$  are only implicitly related through being generated by the same inputs as shown by the model (8.5); there is no trade-off between  $y$  and  $z$  for given  $x$ . In the input – output space we only get one point each for  $y$  and  $z$  given a bundle of inputs  $x$  as implicitly shown by e.g. points A and B where the isoquants for the two outputs intersect.

### *The materials balance*

Assuming that one or more inputs to a production process of a production unit consists of physical mass this mass will not disappear during the production process, but, as stated previously in Section 8.1, must either be contained in the products being produced, or become residuals emitted to the external environment. Thus, for model (8.5) a materials balance exists for each production unit, and  $\alpha$ ,  $\beta$ ,  $\gamma$  are coefficients converting units of inputs  $x$  and desirable outputs  $y$  and residuals  $z$  to a common mass unit:

$$\alpha x_M \equiv \beta y + \gamma z \quad (8.7)$$

As above discussing (8.5) we simplify for convenience by operating formally with only a single input, desirable output and residual respectively, but generalisation to multiple inputs, outputs and residuals is a straightforward but cumbersome extension of the number of equations and a summation over the type of each variable yielding the same pollutant. Assuming that inputs and desirable outputs are homogenous across units, the coefficients  $\alpha$  and  $\beta$  must be equal for each type of input and output, and assuming that given homogenous inputs, respectively desirable outputs, will generate homogenous undesirable outputs, the coefficient  $\gamma$  is also constant across units. The parameters  $\alpha$ ,  $\beta$  and  $\gamma$  are technical unit-conversion coefficients and not parameters of a production relation.

It should therefore be realised that the materials balance is not a production function relation proper. It is without good meaning to e.g. differentiate (8.7) w.r.t.  $y$  and  $z$  holding  $x$  constant using (8.7) as a production relation. If (8.7) is regarded as a production technology  $y$  cannot increase without increasing  $x$ ; a given input  $x$  cannot be reallocated from producing  $z$  to produce  $y$ . All observations of a production activity, perfectly observed, must obey the materials balance as a physical law.

If a parametric or nonparametric efficient production function is estimated based on observations of  $x$ ,  $y$ ,  $z$  the materials balance must also hold for unobserved points obeying the production function. In this sense the materials balance puts a constraint on the efficient production function construction.<sup>9</sup> But the materials balance represents another type of link between the two types of outputs and the inputs than the relations representing the production relations. The materials balance does not tell us anything about the specific technology that has generated the observations in (8.7) (Førsund 2017).

Desirable outputs  $y$  cannot be produced without residuals  $z$  if the complete amount of the material content of the inputs is not contained in the outputs. This is the inevitability of generating pollutants together with the desirable product. Thermal electricity generation is an extreme case with no material content of the desirable output, implying  $\beta = 0$ , but where all material content of the material input ends up as waste products. The aging of a vintage wine may be an extreme example of an opposite case with zero residual. All the material content of the vintage wine is contained in the desirable output, the wine bottle; if we assume that no other material inputs are required such as heating/cooling of the storage facility (e.g. the wine is stored

---

<sup>9</sup> This theme is developed in the ecological economics literature; see e.g. Pethig (2006).

in a natural cave with the right temperature).

The materials balance is fulfilled for any combination of inputs in the model (8.5) because of the unique correspondence between input use and outputs generated by these inputs. The same inputs determine both outputs, so we only get two single values for the outputs for a given level of inputs. Then the materials balance (8.7) being a physical law is also obeyed by definition. Therefore, in our type of model, the materials balance can be regarded as more of an accounting identity since it holds at any point in input-output space satisfying the production functions.

This is illustrated in Fig. 8.1. At point A or B the desirable output cannot be increased by reducing the undesirable output keeping inputs constant; all points giving the values in output space are intersections of isoquants in the input space. The inputs supporting point A give simultaneous the amounts of desirable and undesirable outputs given by the intersection point of the respective isoquants, and similar at point B. Moving from A to B increase the desirable output and decrease the undesirable output, but this is happening due to a considerable reduction in material input and increase in service input.

Model (8.5) does not allow any transformation relationship between the desirable and undesirable outputs for a given level of inputs in the usual sense of a trade-off between the desirable and the undesirable outputs for a given bundle of inputs in a single equation system.

### **8.3.3. Technical change**

There are four main ways of influencing the generation of undesirable outputs when keeping the same type of desirable output. Two ways can be done within a static technology and two other ways require change in technology. Output of desirable goods may be reduced in such a way that use of material inputs is also reduced. This is the only option at a sectoral level within a standard input – output model with fixed coefficients in production (see Section 8.2.2). Another way is to utilise substitution possibilities between material inputs with different marginal productivities in the residuals production function (e.g. switch from coal to natural gas in thermal generation of electricity), and between material inputs and service inputs. These are the two options captured by our model (8.5).

A third way is to change the production technology so as to create less pollution for a constant

output of desirable goods. Technology improvements that may be small-scale and introduced in the short run (e.g. a period of a year or less) may be considered variable factors<sup>10</sup>, but technology improvement may also need large capital investments, changing the main production processes into technologies that use fewer amounts of raw materials, or process them in such a way that less waste of material inputs occur. Such changes will be more of a long-term character based on real capital with a long technical lifetime. This type of technical change is typical in industries using capital with embodied technology. In order to reduce generation of residuals investment in new technology developed to use fewer raw materials or reduce waste is needed.

Technology change, covering both short-time and long-time, means a simultaneous change of the functional forms  $f(\cdot)$  and  $g(\cdot)$  over time:

$$\begin{aligned} f^{t_2}(x_M, x_S) &> f^{t_1}(x_M, x_S) \\ g^{t_2}(x_M, x_S) &< g^{t_1}(x_M, x_S), \quad t_2 > t_1 \end{aligned} \tag{8.8}$$

Technical change in the two production functions in (8.5) may be illustrated in Fig. 8.1 by just changing the level labelling of the two sets of isoquants. A positive technical change of the desirable output production function means producing more for given inputs, while positive “green” technical change in the residuals production function means generating less residuals for the same input levels. Note that the unit conversion coefficients in the materials balance equation (8.7) remain the same as long as the inputs and the outputs stay the same. But technical change may also involve using new types of inputs or redesigning products so as to use fewer raw materials (“dematerialisation” of desirable outputs). These possibilities are not explored here.

A fourth possibility is to install a separate facility using the residuals from (8.5) as inputs and processing them in such a way that less harmful pollutants result, e.g. capturing particles using electrostatic filters on smoke stacks, converting an air pollution into a solid waste problem. Such facilities are called end-of-pipe treatment in environmental economics and will be addressed in the next subsection.

### 8.3.4. End-of-pipe abatement

---

<sup>10</sup> Retooling or re-engineering processes, and recycling internally more waste materials may be done within a reasonable short period, and small-scale investments like heat exchangers to recapture waste heat can reduce the amount of residuals for constant primary energy, and thus increase production (Martin, 1986).

We will add a specific abatement process to the multi-equation model (8.5) that may need inputs distinct from inputs employed in the production process described by the equations in (8.5). End-of-pipe abatement often consists of a facility separated from the production activity. Another abatement option in the short run is to retool the processes and do small-scale changes as mentioned above. This option is an alternative to integrated technological process solutions. However, it is often rather difficult to identify such activities and identify the inputs involved. It is easier to do this with a stand-alone abatement facility in terms of inputs used and outputs produced. Such a facility is called end-of-pipe. The residuals generated in the production process will then be channelled to a treatment facility and be regarded as an input in this activity. In addition other inputs - like labour, capital, chemicals, absorbing substances and energy, may have to be used in order to convert part of the original pollutants into the output being abated pollutants, creating less harm than the primary ones (Førsund, 2009). In the long run there may be a choice between end-of-pipe abatement and large-scale investment in new technology integrating production processes and abatement. The time horizon for environmental improvement, determined by a regulation authority, and certainty about what can be achieved may determine the choice between these two options.

Add-on abatement requires that we make a distinction between primary pollutants  $z^P$  from the production process of (8.5) and secondary pollutants  $z^S$  actually discharged to the environment. In Førsund (2009) a production function for end-of-pipe abatement is formulated following the single equation format of factorially determined multioutput production of having one production function for abatement of each type of residual. Here we will use a simpler formulation by focussing on a cost function in the abated amount  $a$  that is the difference between the original generation of residuals now termed primary residuals  $z^P$ , and the residuals emitted to the environment termed secondary residuals  $z^S$ . The cost function can be regarded as an index for abatement inputs and is written:

$$c(a) = c(z^P - z^S), \quad a = z^P - z^S, \quad c' > 0, c'' > 0 \quad (8.9)$$

The abatement cost function has standard textbook properties in the abated amount<sup>11</sup>. The secondary residuals are assumed expressed in the same units as the primary residuals. The abated amounts are in general of other forms than the primary residuals. As observed in Ayres and Kneese (1969, p. 283) abatement does not “destroy residuals but only alter their form”. We

---

<sup>11</sup> Factor prices of abatement inputs are not introduced because the employment of individual abatement inputs will not be studied (see Førsund 2009; 2017 for such studies).



assume that the cost of disposing of these amounts, e.g. using landfills, is included in the cost function, but that any environmental cost is zero.<sup>12</sup>

As to the materials balance principle the abatement activity will add to the total mass of residuals if material inputs are used, but the point is that abatement means less mass of the harmful residual emitted to the environment;  $z^S < z^P$ .

In the literature the resources of a firm are often regarded as given, and then increased abatement will imply fewer resources to produce the intended output and thereby decrease the generation of primary pollutants (Martin 1986; Murty et al 2012). To do this requires in principle that a restriction is imposed on the availability of inputs. However, this problem is created by the analyst and does not necessarily reflect decisions of a firm having access to markets for inputs to given prices. If it is assumed that abatement is a separate identifiable activity, as e.g. end-of-pipe, and inputs are sourced in competitive markets, there is no reason to assume that abatement resources are taken from the production inputs of a firm. Thus, abatement does not influence the output directly, but increases the cost of production and may then indirectly reduce output. It is closer to reality not to consider a common resource pool for the production unit, but to regard the activity behind (8.9) as a separate “profit centre”.

### 8.3.5. Environmental considerations

#### *The social optimisation problem*

In order to see how productivity can be influenced by environmental consideration we will use our model with joint production of desirable and undesirable outputs and add end-of-pipe abatement possibilities. The approach is as simple as possible and of a partial equilibrium nature. The social planning problem is to maximise consumer plus producer surplus, introducing a demand function on price form,  $p(y)$  for the desirable output, assuming a positive output price  $p$ . (The model is still kept simple with only a single output.) Fixed positive price  $w_j$  is used as social evaluation coefficient for input  $j$  for types  $M, S$ .

---

<sup>12</sup> This may not always be the case, there may be harmful run-offs from landfills and incineration of waste may create harmful gases, but these factors can in principle be entered into the analysis at some cost of extending the equation system (Førsund 2009).

Simplifying assuming single desirable and undesirable output as in model (8.5), a single abatement cost function in the abated amount as in (8.9), and a monetized damage function (8.2), now with a single pollutant emitted to the environment, the social planner's optimisation problem, maximising the consumer plus producer surplus, is:

$$\begin{aligned} \text{Max } & \int_{u=0}^y p(u)du - \sum_{j=M,S} w_j x_j - c(a) - D(z^S) \\ \text{s.t. } & \\ & y = f(x_M, x_S), f'_{x_M}, f'_{x_S} > 0 \\ & z^P = g(x_M, x_S), g'_{x_M} > 0, g'_{x_S} < 0 \\ & a = z^P - z^S \end{aligned} \quad (8.10)$$

The first term in the objective function is the area under the demand curve from 0 to  $y$ , i.e., the sum of consumers' willingness to pay for  $y$  and the revenue generated by production. The next two terms are the input- and abatement costs of production. The last term is the social cost of emission of secondary pollutants. The model can straightforwardly be extended to multiple desirable and undesirable outputs as suggested previously. When considering several units demand functions have to be adjusted according to type of demand interactions, and it must be specified whether the damage functions are unique to each unit, or the nature of interactions between damage functions must be specified if there are any (see Section 8.2). This simple model makes the fundamental trade-off between desirable output  $y$  and undesirable pollutant  $z^S$  tractable, allowing for Pareto-optimal allocation rules both for the two types of inputs and for abatement effort. From the objective function we see that the abatement activity does not directly influence the generation of pollutants, but indirectly by influencing the optimal solution both for primary and secondary pollution.

Inserting the production functions for the good  $y$  and the pollutant  $z^P$  that is the primary pollutant into the objective function, and substitution for abatement  $a$  in the cost function yields:

$$\text{Max } \int_{u=0}^{f(x_M, x_S)} p(u)du - \sum_{j=M,S} w_j x_j - c(g(x_M, x_S) - z^S) - D(z^S) \quad (8.11)$$

There are three endogenous variables remaining in the problem;  $x_M$ ,  $x_S$ , and  $z^S$ . The necessary first-order conditions for interior solutions are:

$$\begin{aligned} pf'_{x_j} - w_j - c'g'_{x_j} &= 0, j = M, S \\ c' - D' &= 0 \end{aligned} \quad (8.12)$$

This is three equations in the three endogenous variables. When these variables are determined the optimal solutions for  $y$ ,  $z^P$  and  $a$  follow directly. The two first necessary conditions tell us that for each type of factor  $j = M, S$ , the revenue of increasing factor  $j$  marginally and consequently increasing the desirable output, is equal to the unit cost of the factor plus the abatement cost of increasing factor  $j$ , generating primary pollutant that is abated. For services negative productivities in the generation of primary pollutants imply that the last term in the first condition is positive and show the abatement cost *savings* of a marginal increase in a service input. The third condition tells us that at the optimal level of abatement, i.e. both primary and secondary pollutants are at their optimal levels, the marginal abatement cost should be equal to the marginal damage of the secondary pollutant.

The two first conditions in (8.12) can be written

$$f'_{x_j} p = w_j + D'g'_{x_j}, \quad j = M, S \quad (8.13)$$

Marginal revenue
Unit factor cost
Marginal Damage cost = marginal abatement cost

At the margin a material input generates environmental damage cost, in addition to the input price, while a service input generates a saving of damage cost by subtracting the marginal damage from the input price. The equation shows the fundamental trade-off between the value of the man-made good and the total social cost including the environmental damage at the margin.

To see the impact of the use of inputs due to different productivities in the generation of primary pollutant, the rate of substitution between a material input and a service input is:

$$\frac{f'_{x_M}}{f'_{x_S}} = \frac{w_M + c'g'_{x_M}}{w_S + c'g'_{x_S}} \quad (8.14)$$

The optimal unit price on the material input faced by the decision-maker is higher than the given market price, but the opposite is the case for the service factor because the impact  $g'_{x_S}$  on cost is negative. The optimal solution implies a relative reduced use of the material input compared with a solution without a damage function and abatement cost function. As to the substitution between material inputs the increase in the social price follows the marginal productivity in pollutant generation, thus stimulating a shift to “greener” material inputs. It is easy to see that if a material input has a social factor price that is greater than the value of the marginal product in the desirable goods production for any value of the input, then this input should not be used.

In a realistic case of a technologically forced choice between e.g. coal and natural gas as primary energy source the input with the lowest social factor price given the same level of marginal revenue will be chosen.

The social productivity measure  $TFP_E$  introduced in (8.3) for the solution to problem (8.10) and productivity change are

$$TFP_E = \frac{y - D(z^S)}{F(x_M, x_S) + c(a)}, \quad (8.15)$$

$$\Delta TFP_E^{t,t+1} = \frac{y^{t+1} - D(z^{S,t+1})}{F(x_M^{t+1}, x_S^{t+1}) + c(a^{t+1})} / \frac{y^t - D(z^{S,t})}{F(x_M^t, x_S^t) + c(a^t)} \geq 1$$

The cost term in the denominator represents an index for abatement resources used and is deflated over time. This measure of productivity may increase as well as decrease over time depending on the optimal solution for each period. Technical change and change in input mix may influence the input variables, and the damage- and abatement functions may change over time, the former due to preference changes and the latter due to technical change. The optimal level of both the secondary residual in the damage function and amount of abatement may then change over time also.

The macro factors mentioned in Section 8.2 - reallocation of resources to less polluting activities and mitigation and abatement effort – that may drive prices on the resources up slowing capital accumulation and growth, are also relevant for the development of productivity at the micro level in (8.15).

### *Imposing a constraint on emission*

The social planner is a hypothetical construct to give us a frame of reference for evaluating a real world situation with private firms as decision-makers and a regulator imposing constraints on private activities. To see how productivity is influenced we look at a firm regulated in such a way as to implement the social optimal solution to (8.10). We will investigate the effects on a private firm operating in a competitive environment of the cost burden of environmental regulation. Two typical instruments of direct control is, firstly, that the environment agency imposes a change in technology, most often concerning end-of-pipe abatement technology, and, secondly, that an upper limit on the amount of the secondary pollutant emitted during a specific

time period  $z_R^S$ , is imposed;  $z^S \leq z_R^S$ . In the latter case the firm's optimisation problem of private profit under an environmental constraint becomes

$$\text{Max } py - \sum_{j=M,S} w_j x_j - c(z^P - z^S)$$

subject to

$$y = f(x_M, x_S)$$

$$z^P = g(x_M, x_S)$$

$$z^S \leq z_R^S$$

(8.16)

where the last inequality is the restriction on emissions of the secondary pollutant. A regulation is often expressed as restricting emission per unit of output;  $z^S / y \leq z_R^S / y$ . However, this is the same condition as above.

We assume perfect compliance and costless enforcement of the regulation. Substituting for the output and the primary pollutant in the cost function the Lagrangian function can be written:

$$L = pf(x_M, x_S) - \sum_{j=M,S} w_j x_j - c(g(x_M, x_S) - z^S) - \lambda(z^S - z_R^S)$$

(8.17)

The necessary first-order conditions are:

$$pf'_{x_j} - w_j - c'g'_{x_j} = 0, j = M, S$$

$$c' - \lambda \leq 0$$

$$\lambda \geq 0 (= 0 \text{ for } z^S < z_R^S)$$

(8.18)

Here  $\lambda$  is the shadow price on the emission constraint. If the constraint is not binding we see from the complementarity slackness condition that the firm is not influenced by the environmental regulation and will not use any abatement. If the value of the desirable output cannot meet the social cost at any level of the input  $j$  (assuming several inputs in group  $M$ ) this input should not be used. Assuming that our input  $j$  is used in a positive amount and that the emission constraint is binding, the shadow price shows the gain in profit of marginally relaxing the constraint. We see comparing (8.12) and (8.18) that the shadow price takes the place of the marginal damage of the secondary pollutant in the optimal solution. When the emission constraint is binding the cost in the optimal solution becomes greater than the private optimal

solution without emission regulation because abatement has to be used. The substitution effects between inputs triggered by the use of abatement resources shift the use of production inputs away from material inputs. The shift is stronger the higher the marginal productivity of the material input in question is in residuals production. The use of service inputs is increased. These substitution effects counter some of the reduction in output and profit, but these latter variables will inevitably decrease. Levels of service inputs may increase to higher levels than in the pre regulation situation, but not enough to compensate the reduction in output and profit, if so the firm would itself have imposed a subsidy on its use of service inputs in the pre-regulation case.

The private productivity development is equal to the social productivity development in (8.15) with the important exception that the environmental damage term is absent from the numerators. We have that due to abatement costs in the denominator the private productivity level goes down if output and inputs remain constant. Once regulation is introduced the development of productivity can both be positive and negative depending on relevant factors discussed in connection with (8.15). A new element is that the regulated amount is subject to political changes over time. Stricter regulation is to be expected, strengthening, *ceteris paribus*, the case of reduced private productivity.

As stated above the shadow price on the emission constraint in (8.17) shows the impact on the profit of a change of the allowed emission. Marginal abatement cost has been suggested as measure of the price on pollution. In the optimal solution with a binding emission constraint we see from (8.18) that marginal cost of abatement equals the shadow price. However, this shadow price is not in general equal to the social price on pollution. In the optimal solution of the social model (8.10), marginal damage is equal to marginal cost, so setting a price on pollutants equal to the shadow price  $\lambda$  in (8.18) may either undervalue or overvalue the social marginal cost, depending on the difference between the regulated amount  $z_R^S$  and the optimal solution of  $z^S$  to (8.11). It is only if the allowed amount of secondary pollutants is equal to the optimal solution of the social planner's problem (8.11) that marginal abatement cost should equal marginal damage.

#### **8.4. The Porter hypothesis**

On the backdrop of the huge cost climate change may inflict on us all, and other pressing environmental problems of more regional nature, stricter and stricter environmental regulation is introduced and planned to be introduced in the near future. The Stern Review (2006) gives estimates of mitigating and abatement costs necessary to reach global environmental targets. A worry is that environmental regulations will set back the growth of standard goods (exclusive of environmental goods) and prevent developing countries to achieve the standard of living enjoyed in developed countries. At the micro level a strict environmental policy imposes costs on firms and leads to lower profit and activity, according to the traditional view by most economists. This is also the case for private productivity as pointed out in the previous section, but not necessarily the case for productivity growth.

In Porter (1991) and Porter and van der Linde (1995) a more optimistic Panglossian hypothesis, the *Porter hypothesis*, is put forward: strict environmental regulation may induce firms to innovate to such a degree that profit increases, and thus strict regulation represents a win-win situation. It is stated that the pessimistic view stems from considering a static situation only, but that the pressure of environmental regulation induces a dynamic process of change representing retooling, process improvement and technical change, which more than offsets the abatement costs. However, Porter and van der Linde do not present any formal mechanism supporting the cost offset hypothesis, but refer to a few examples of successful adaptation and technical change. The Porter hypothesis, and attempts to model the positive dynamics (Ambec and Barla 2002; Xepapadeas and deZeeuw 1999), and empirical studies and critique of the hypothesis (e.g. Palmer et al 1995) is extensively reviewed in Brännlund and Lundgren (2009), Lanoie et al (2011) and Ambec et al (2013). The latter three references provide long lists of references to the literature on the Porter hypothesis.

Notice that in the discussion of the Porter hypothesis environmental costs, the damage costs, used in Sections 8.2 and 8.3 are not considered. The focus is on the possibilities of reducing emissions and the private cost of this.

#### *Dynamic effects of environmental regulation*

Porter and van der Linde suggest two different dynamic effects. First, assuming that there is inefficiency in utilisation of resources before the introduction of environmental regulation this inefficiency is reduced or even removed after regulation has been introduced. Second, the

regulation induces new technology to be developed shifting the production function outwards. This is set out in Fig.2. Panel (b) illustrates these two effects. In the space of the desirable

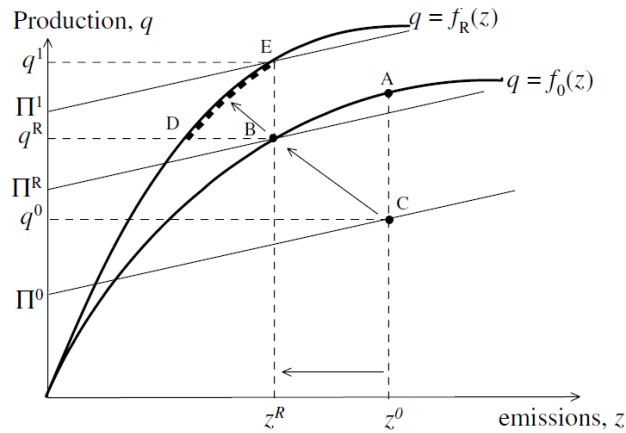


Figure 8.2. The Porter hypothesis  
Source: Brännlund and Lundgren (2009), p.83

output ( $q$  in the original figure) and emissions  $z$  the pre-regulation position of the firm is at the inefficient point C below the initial frontier production function  $f_0(z)$ . The efficient point A on the frontier shows the production the firm could have had corresponding to emission  $z^0$ . After introducing regulation the firm improves its efficiency and reduces the emissions down to  $z^R$  and increases output from  $q^0$  to  $q^R$  at point B on the initial frontier. Then there is a shift of the frontier due to innovation after introducing regulation to  $f_R(z)$  where the point E is the efficient point for the level  $z^R$  of the reduced emission. The firm continues to reduce emissions and increase output  $q$  and profit  $\Pi$  moving towards the new frontier.<sup>13</sup> This is a neat possible illustration of the story told in Porter and van den Linde (1995) of the increased efficiency effect and the shift in technology effect, but does not explain the mechanisms behind the moves.

Establishing credible dynamic mechanisms has proven to be rather difficult. Regarding inefficiency this difficulty is underlined by the fact that there are hardly any dynamic mechanisms in the huge body of specialised literature on inefficiency. Estimating efficiency scores based on parametric and non-parametric production functions, and indices for productivity change of the Malmquist type using technical efficiency scores, may be said to be no more than more or less advanced description of implications of observed data on outputs and inputs and do not represent explanations of inefficiency. There is a general methodological

<sup>13</sup> As opposed to the model (8.5) the pollutant is functioning as an input. This can be the case if there is a fixed relationship between the real input and the pollutant in a single desirable output and a single input and pollutant world. The pollutant is then a *shadow factor* to the material input, using the terminology of Frisch (1965).



problem here, namely the difficulty in explaining inefficiency based on rational decisions. One type of explanation mentioned in the early efficiency literature (see e.g. Førsund (2010) for a review) is that the inputs may be heterogeneous (Farrell (1957) mentions capital) implying that the estimated inefficiencies are biased. Stigler (1976) goes so far, in reviewing the Leibenstein (1966) concept of X-efficiency, to state that there is no inefficiency in competitive industries, just inhomogeneous inputs including not only capital and labour, but also management (Charnes et al (1978) point to the latter factor as a cause of inefficiency). Inefficiency can be rationalised if the technology in production capital and equipment is embodied (Førsund and Hjalmarsson 1974). Then if a firm having old technology is compared with one having a new technology may appear as inefficient. However, with long-lived equipment this type of inefficiency may well serve an objective of maximising the present value of profit (Førsund 2009).

The standard critique of the Porter hypothesis builds on the questions: why have not good ideas already been utilised, and why should the firms wait until environmental regulation comes along? If profitable opportunities to reduce emissions of pollutants exist, profit-maximising firms would already be taking advantage of them (Ambec et al 2013). Apart from pointing to limited information about possibilities and limited resources, support for the dynamics of the Porter hypothesis can be found in the theory of induced innovations of Hicks (1932). Porter and van der Linde (1995) talk about the pressure put on firms by environmental regulation to comply with regulations. The pressure in the theory of Hicks is represented by costs of inputs, i.e. resources are devoted to research and development in order to reduce costs. However, the development of technology may have deeper roots than mere cost savings. There is a path from basic research done without any cost considerations firms may have about usefulness, to applied research and to innovations within firms themselves or equipment-producing industries. But it is not easy to establish a dynamic mechanism for events that are stochastic by nature.<sup>14</sup>

We can distinguish between three types of reactions to new environmental regulation by a firm: short-run measures like better process control, small-scale reengineering, introducing more internal recycling of waste, leading to improved efficiency of utilizing material inputs and thereby reducing pollutants, medium-term measures like investing in end-of-pipe technologies, and long-term measures involving developing new technologies for the main processes saving material inputs and reducing waste products, and even developing waste products into saleable

---

<sup>14</sup> Innovations may also be directed to develop new products. As mentioned before, this is left out of the story in this chapter.

products. Technologies used in material-processing industries often have embodied specific technologies so technical change can only come by investing in new technologies and scrapping old machinery and equipment. It may be the case that a choice has to be made between investing in short-term measures or long-term measures. But going for short-term measures may lead to less emphasis on developing new process technologies and thus may not be a socially optimal decision. Uncertainty about future regulation may imply favouring short-term measures.

The requirement for technical change to lead to increased private profit can be expressed using the relations (8.8). In addition there may be technical change in the abatement function that we can express using the cost function (8.9):

$$c^{t_2}(z^P - z^S) < c^{t_1}(z^P - z^S), t_2 > t_1 \quad (8.19)$$

The periods  $t_1$  and  $t_2$  may not necessarily be consecutive years, but spanning a period more reasonable for technical change to take place. We will therefore consider capital as one of the factors, but without formulating any investment problem that should have to be done in a realistic setting. Let us assume that the environmental regulator set an upper limit on allowable emission as in model (8.16), but let the firm itself determine how to obtain this level. We assume that the regulation starts in period  $t_1$  and that the regulated amount is the same and binding in both periods. It is reasonable that it takes some time before the regulatory pressure leads to new technology being available, so the difference between  $t_1$  and  $t_2$  may be several years (Lanoie et al 2008). The product price and the factor prices, respectively, are the same in both periods.

The condition for the pressure to lead to increased profit is then

$$\begin{aligned} & py^{t_2} - \sum_{j=M,S} w_j x_j^{t_2} - c^{t_2}(z_{t_2}^P - z_R^S) > py^{t_1} - \sum_{j=M,S} w_j x_j^{t_1} - c^{t_1}(z_{t_1}^P - z_R^S) \Rightarrow \\ & \underbrace{p(f^{t_2}(x_M^{t_2}, x_S^{t_2}) - f^{t_1}(x_M^{t_1}, x_S^{t_1}))}_{\text{Output effect}} + \underbrace{\sum_{j=M,S} w_j (x_j^{t_1} - x_j^{t_2})}_{\text{Input effect}} \\ & + \underbrace{(c^{t_1}(g^{t_1}(x_M^{t_1}, x_S^{t_1}) - z_R^S) - c^{t_2}(g^{t_2}(x_M^{t_2}, x_S^{t_2}) - z_R^S))}_{\text{Abatement cost effect}} > 0 \end{aligned}$$

(8.20)

In period  $t_1$  the initial output effect is negative but input costs probably goes down- but then abatement costs contribute to reduce the profit. There are three effects of technical change in period  $t_2$ ; the change in the amount of output, the change in the amounts of inputs and the change in the abatement costs. As to the latter effect we also have a negative cost consequence in the reduction in primary pollutant due to technical change. Technical change satisfying the

condition (8.20) will have a positive output effect, an uncertain input effect and a reduction in the abatement costs because of the two technology shifts in the second period.

If the effect on profit is positive, then private productivity will also increase over time. To keep up a pressure on firms that have managed to comply to the regulation it may be necessary to have more stringent regulation over time to keep the dynamic process going.

### *Regulation using economic incentives*

The type of regulation introduced can play a decisive role for the reactions of firms. Porter and van der Linde (1995) base their dynamic innovation process on well-designed environmental regulations based on economic instruments. They go strongly out against using command and control that demands a specific technology to be used. This will stifle any innovation activity of the firms exploring other solutions and may not lead to any further development of technology if the command of a specific technology to be used satisfies the required reduction in emissions. The technology imposed by the regulator is usually of the type best available technology (BAT) or “best available technology not entailing excessive cost” (BATNEEC). Installing end-of-pipe abatement may be cheaper and less complicated than demanding new basic process technologies, so the regulator may be biased toward imposing end-of-pipe solutions without considering the socially optimal solution in a longer run. However, regulators may also give firms a certain number of years before complying if changing the process technology is the best solution. Thus, existing firms may enjoy a positive quasi-rent from “dirty” processes, biasing against new firms entering the market.

The seminal economic instrument is to introduce a Pigou tax on the emitted pollutant (Pigou 1920). Using our model (8.5) and (8.9) the environmental regulator uses a tax on secondary pollution as the instrument giving incentive to react both in the short- and the long run. Regarding the unit as a firm that maximises profit, facing competitive markets both for output and inputs, introducing a Pigou tax  $t$  on secondary pollutants yields the following optimisation problem:

$$\text{Max } pf(x_M, x_S) - \sum_{j=M,S} w_j x_j - c(g(x_M, x_S) - z^s) - tz^s \quad (8.21)$$

The production functions for output and pollutant from (8.5) are inserted for output and primary pollutant, respectively, and a flat unit tax  $t$  is imposed on the secondary pollutant. We assume

that the firm has access to an end-of-pipe abatement technology, but any investment cost is disregarded, we look at the management problem of using capacities and do not perform investment calculations.

The necessary first-order conditions, assuming that abatement will be used, are

$$\begin{aligned} pf'_{x_j} - w_j - c'g'_{x_j} &= 0, j = M, S \\ c' - t &= 0 \end{aligned} \tag{8.22}$$

The optimal firm solution will be a function of the exogenous tax rate  $t$ . Going back to (8.12) we see that the firm solution will conform with the social solution if the tax rate is set equal to the marginal environmental damage in the latter optimal solution.

The choice between using and not using abatement is decided by the last condition in (8.22): if the marginal cost is higher than the unit tax for all levels of abatement, then it is optimal to set abatement equal to zero and emitting the primary pollutant with the tax levied on it, yielding the condition  $pf'_{x_j} - w_j - tg'_{x_j} = 0$ . With positive abatement the optimal solution to problem (8.21) implies that the marginal abatement cost is set equal to the tax. We then see that we get the same solution for the inputs and the desirable output in the two cases, but without the abatement being optimal to use, the emission to the environment is higher than in the case of the abatement facility being used. The relative use of factors will be changed in the same way as illustrated by (8.14) and output will be scaled down compared with the profit maximising solution without a charge levied on the pollutant.

The results are based on static technologies; substitution takes place within the pre-regulation technology. According to the Porter hypothesis the tax introduces a pressure on the firms to try to reduce the tax “punishment” in the short run by diminishing any existing polluting waste, and in a longer run adopt end-of-pipe solutions that reduces the tax burden, and develop new technologies using less raw materials that generates pollutants per unit of output.

The impact on productivity and productivity change is similar to what is described in the previous subsection. But a new element has entered the numerator; the tax paid on the secondary pollution is deducted the private profit. Again, the static productivity is reduced, but the productivity change may be both negative and positive according to the strength of the different factors as discussed above. But now the tax amount is a new influence on the productivity change if it increases over time.

Direct regulation in the form of an upper limit on pollutant emitted to the environment should also work as a pressure to adapt and adjust the technology and to innovate also basic process technologies. We see from the necessary first-order conditions (8.12) for optimal solution of model (8.11) that these conditions are of the same form as in (8.22). However, there is a difference as to the strength of the pressure because the regulated amount  $z_R^S$  is not necessarily the optimal amount  $z^{S*}$ , and in the case of using a tax this has to be paid in addition to the abatement costs thus reducing the private profit below the social optimal one<sup>15</sup>. With the direct regulation of this type the firm has the same freedom as under a tax regime to develop ways of reducing the use of pollution-generating raw materials and to develop technologies either of the end-of-pipe type or of process type doing the same. For the regulator the difference is that a known maximum amount  $z_R^S$  is emitted to the environment imposing such a limit, while with the tax solution the regulator may be uncertain about what will actually be realised as the emission (assuming that the regulator does not have perfect information about the firm's technology).

### *Cap and trade*

It is well known from the environmental economics literature that “cap and trade” can combine both types of instruments for the situation that a number of sources emit the same type of pollutant or pollutants that have the same effect on the environment independent of the location of the source (e.g. greenhouse gases and their effects on global temperature). By imposing an upper limit on the total amount of emission from all firms and then introducing quotas for each firm made tradeable, the equilibrium price of quotas will act as if a tax on pollutants is introduced as in problem (8.21).

The optimisation problem for firm  $i$  ( $i=1, \dots, N$ ) is now:

$$\text{Max } pf_i(x_{iM}, x_{iS}) - \sum_{j=M,S} w_j x_{ij} - c_i(g_i(x_{iM}, x_{iS}) - z_{iC}^S) + p_Q(z_{iQ}^S - z_{iC}^S) \quad (8.23)$$

We have three types of endogenous variables;  $x_M$ ,  $x_S$ , and  $z_C^S$ . The quota for firm  $i$  is  $z_{iQ}^S$  and we assume that  $z_{iQ}^S < z_i^S$  for all firms, so the total emission goes down due to the caps. If it is not optimal to use abatement then the emissions  $z_i^P$  is the optimal emission emitted to the

---

<sup>15</sup>This is a well-known result in textbooks on environmental economics (Førsund and Strøm 1988).

environment. We assume that the quota price  $p_Q$  is clearing the quota market. The emission  $z_{iC}^S$  is the optimal emission under the cap and trade regulation.

The necessary first-order conditions are

$$\begin{aligned} pf'_{ix_j} - w_j - c'_i g'_{ix_j} &= 0, j = M, S \\ c'_i - p_Q &= 0, i = 1, \dots, N \end{aligned} \quad (8.24)$$

Having obtained the solution for the endogenous variables  $x_M, x_S$ , and  $z_C^S$  the solutions for  $y$  and  $z^P$  can be obtained. The first condition shows that use material inputs will be reduced; the unit cost will go up, but the use of service inputs will go up; their unit cost is reduced. It is reasonable to assume that the firm will react with a cutback in production thus reducing revenue and also profit. As to buying or selling quotas the crucial condition is the second one saying that marginal abatement cost should be equal to the quota price. Since the quota price is the same for all firms the marginal abatement costs will also be the same across firms. The firms that abate, but have optimal emissions that exceeds the quota,  $z_{iC}^S > z_{iQ}^S$ , have to buy additional quotas. It is possible that marginal abatement cost is higher than the quota price for zero abatement, then  $z_i^P - z_{iQ}^S$  has to be bought on the quota market. The sellers will be firms for which it is profitable to abate so much that  $z_{iC}^S < z_{iQ}^S$ . It is formally possible that marginal abatement cost is lower than the quota price for maximal abatement, i.e.  $a = z_i^P$  and  $z_{iC}^S = 0$ , although observations typically show that this is not the case.

We see that all firms that abate will face the same marginal abatement cost equal to the clearing price  $p_Q$  in the market. Comparing this solution to the solution (8.24) to problem (8.23) with a tax on the emission we have that, provided that the total restriction on emissions is the same as the sum of optimal solutions for emissions when using a tax, then the market clearing price is equal to the tax.

The advantage with cap and trade is that it combines the certainty of regulated emissions, provides perfect compliance, and gives economic incentives or pressures to improve efficiency and develop new technologies reducing emissions. An additional advantage is that the regulator can tighten the total amount of quotas over time as a single decision.

However, a tax on emissions and cap-and-trade may have different incentives in a dynamic setting. A tax represents a predictable form of regulation for a firm, assuming that it is not

changed so frequently. But the price on emission permits may fluctuate quite a lot and represent an unpredictable factor for investment decisions related to environmental regulations.

Behind fluctuations are the business cycle, technology change, competition from firms in countries not regulating carbon emissions, and for thermal electricity generators competition from renewable energy and its fluctuations of production. The incentives of the two different instruments may give incentives to investment in different technologies, presumably more short-term considerations in the form of end-of-pipe technologies than long-term more costly and fundamental process innovations due to uncertainties about the price. The fluctuations are illustrated in Fig. 8.3 showing the price development for carbon certificates<sup>16</sup> within the EU Emission Trading System (ETS) for 10 years 2005-2014.

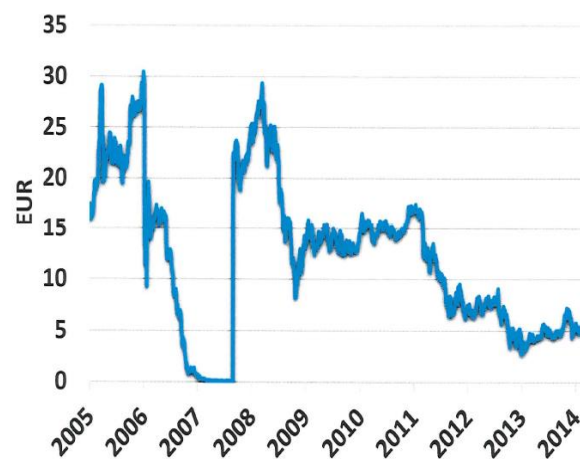


Figure 8.3 Volatility of the EU ETS carbon certificate market

Source: Alexandru P. Luta. *The current state of the EU ETS.*

The Sanbag Climate Campaign July 22, 2014

[https://www.google.no/webhp?gws\\_rd=ssl#q=The+current+state+of+the+EU+ETS](https://www.google.no/webhp?gws_rd=ssl#q=The+current+state+of+the+EU+ETS)

The theoretical model above of the quota market links marginal abatement costs to the quota price. However, estimates of abatement costs of participating firms fluctuate much less. The EU ETS has had three phases, the first phase covering the period 2005-2007, the second phase 2008-2012, and the third phase 2013-2020. The scheme started out with free quotas for about the total emissions. A positive price is due to mismatch between firm specific quotas and actual emissions. The fall in price during 2007 is due to earlier trades leading to a better match quotas

<sup>16</sup> Note that it is carbon that is the physical basis in the market, not e.g. CO<sub>2</sub>. This means that  $\gamma z$  in (8.7) is the basis.

and emissions, and the fact that quotas bought earlier but saved to be used later could not be transferred to Phase II.

In Phase II the total quota was reduced somewhat below the emissions and more countries joined the system. The price started at levels similar to the start of Phase I, but then the economic recession caused by the financial crisis started to bite and reduce output of firms within the system and the quota price plummeted. The continuing slide and low level, also kept up to now, is due to structural adjustment of firms in the participating sector like fuel substitution of thermal electricity generators and the increased share of renewables in electricity production.

### *Empirical studies of the Porter hypothesis*

In the substantial literature on the Porter hypothesis it is difficult to find general theoretical arguments on which to build mechanisms that give the results of the hypothesis. Therefore the validity of the hypothesis may be regarded as an empirical question. There is also a rich literature trying to do this, but there is a high ratio of speculation and anecdotes to evidence (Jaffe et al 1995; Lanoie et al 2008), so it is difficult to generalise the obtained results. We will not give a detailed account of empirical studies here, but focus on main findings.<sup>17</sup>

A common approach is to study separately three aspects of the Porter hypothesis; the weak version, the narrow version and the strong version (Lanoie et al 2011). The weak version posits that environmental regulation will stimulate innovations and technology choice to make production “greener”, as measured by R&D, investment in capital and patents, the narrow version states that flexible environmental regulation give firms stronger incentive to innovate than command and control regulation prescribing specific technologies, and the strong hypothesis states that properly designed regulation may induce innovations that more than compensates for the cost of compliance measured by firms’ performance such as productivity and costs.

---

<sup>17</sup> For detailed accounts see Brännlund and Lundgren (2009) that give extensive tables of references with information on purpose and method, data, and results, Lanoie et al (2011) that used a data set for approximately 4200 facilities collected from seven OECD countries by postal survey in 2003, and Ambec et al (2013) that also give an account of many empirical studies.



The main findings are that the weak hypothesis have some empirical support, as also is the case for the narrow hypothesis, but that there is not much evidence for the strong hypothesis of increased profit due to regulatory-induced innovations.

In the studies of innovations attention is paid to time lags between introduction of environmental regulation and innovations. There is a transition period in most cases. Contemporaneous impact on economic performance may be negative, but can turn out positive when using lags between regulatory variables and effects (Lanoie et al 2008).

There is a data problem identifying the environmental cost part of process changing investments that may have been done without regulation. The lack of accurate accounting of regulation cost that allocates the parts attributable to environmental regulation and to process changes undertaken without regulation is becoming a problem because there is an increase in process changes and product reformulations rather than installation of end-of-pipe equipment.

## **8.5. Productivity measures based on inefficiency**

An interpretation of the Porter hypothesis is that existing efficiency before introduction of environmental regulation is reduced or even eliminated as low hanging fruit in the short run, and in the long run inefficiency seems to disappear completely, at least it is not mentioned in connection with long run. However, there is a body of literature on productivity that assume that inefficiency may be persistent over time, the main reason being just innovation and new technology shifting the frontier production function upwards and consequently making it harder for firms to become efficient. This strand of research is commonly based on the assumption that prices are not available and estimation of the efficiency scores are mainly based on a non-parametric frontier. (See Førsund (2009); (2017); Murty et al (2012); Dakpo et al (2016) for critical reviews of this literature.)

### *Single equation models*

The inspiration of Färe et al (1989) to explicitly address the question of how to incorporate generation of pollution when estimating the productivity of both desirable and undesirable

outputs, was Pittman (1983). In the latter paper the approach of Caves et al (1982a), introducing translog multilateral productivity indexes, is extended to include undesirable outputs or pollutants. Shadow prices, mainly based on abatement costs, on undesirable outputs are used in calculating the output index based on revenue shares and an input index based on cost shares.

Färe et al (1989. p. 90) introduced an enhanced hyperbolic efficiency measure based on estimating Farrell (1957) efficiency measures, solving an LP problem for a non-parametric frontier function. This was done in order to credit producers for “their provision of desirable outputs and penalize them for their provision of undesirable outputs.” A single equation model in the form of a distance function was used. The starting point is to introduce a technology using the output possibility set that is conditional on the inputs

$$P(x) = \{(y, z) : x \text{ can produce } (y, z)\} \quad (8.25)$$

Some general assumptions are that the set is convex and is monotonic in  $y$  and  $x$ , that is both  $y$  and  $x$  are strongly disposable. The trade-off between desirable and undesirable output  $y$  and  $z$  must be restricted in some way for the technology to lead to sensible results (Førsund 2009). The way it is done in Färe et al (1989) and the literature that followed, is to impose weak disposability (introduced in Shephard 1970; 1974) for  $y$  and  $z$  together

$$(y, z) \in P(x) \text{ and } 0 \leq \theta \leq 1 \text{ imply } (\theta y, \theta z) \in P(x) \quad (8.26)$$

Furthermore it is assumed that  $y$  cannot be produced without  $z$  also being produced.

However, although the intention of measuring efficiency taking into account generation of pollutants is a good one, the solution using the assumption of weak disposability is not. As pointed out in Section 8.3.2 the generation of pollutants stems from the use of material inputs. Then the materials balance accounts for where the mass content of the raw materials ends up; incorporated in the desirable output or as pollutants. There is physically no possibility of a technical trade-off between a desirable output and a pollutant for given amounts of inputs (see Section 8.3.2.)<sup>18</sup>. But such a possibility is not only allowed in the formulation (8.26), but is crucial for the modelling, using a radial or directional distance function as the single equation model (cf. the many illustrations in the weak-disposability literature of such a trade-off). One should not feel comfortable neglecting a basic physical law when there are alternatives available such as the model presented in Section 8.3.2.

---

<sup>18</sup> This is extensively discussed in Førsund (2009) and (2017).

The problem with dealing both with desirable and undesirable goods when prices on the latter do not exist, have been met in the last decade by estimating efficiency scores using directional distance functions (Chung et al 1997). The directional distance function is based on choosing a direction from an inefficient observation (adding or subtracting from the observed values), to the frontier that is not necessarily radial. For an output-oriented efficiency score the value of the output point on the frontier is measured adding to the observed output the product of a common expansion factor and the desirable output, and in the case of an undesirable product to subtract the product of the expansion factor and the value of the observation of an undesirable output.

The directional output distance function is defined by

$$\bar{D}_o(x, y, z; g_y, -g_z) = \sup[\beta : y + \beta g_y, z - \beta g_z] \in P(x) \quad (8.27)$$

The directions  $g_y$  and  $g_z$  are usually chosen as  $(y, -z)$ . The maximal value of the directional distance function is zero when the unit in question is on the frontier, and it is restricted to be less than or equal to 1. The expansion factors for desirable outputs are  $(1+\beta)$  and the contraction factor for undesirable outputs is  $(1 - \beta)$ .

A Malmquist type of productivity change index, termed Malmquist-Luenberger index (*ML*) in Chung et al (1997), using directional output distance functions for period  $t$  and  $t + 1$ , can be written<sup>19</sup>

$$ML_t^{t+1} = \frac{1 + \bar{D}_o^t(x^t, y^t, z^t; y^t, -z^t)}{1 + \bar{D}_o^t(x^{t+1}, y^{t+1}, z^{t+1}; y^{t+1}, -z^{t+1})} \quad (8.28)$$

$ML_t^{t+1} > 1$  implies  $\bar{D}_o^t > \bar{D}_o^{t+1}$ , i.e. that period  $t$  efficiency for the unit in question is less than the efficiency in period  $t + 1$  resulting in an increase in productivity. Because the *ML* productivity index is based on the assumption of weak disposability it is open to the problems stated above.

Another type of problem, if accepting the directional distance model, is that measuring productivity change by using functions based on a common proportionality growth factor with opposite signs for the two types of outputs, are rather arbitrary as the basis for a productivity

---

<sup>19</sup> In Chung et al (1997) the *ML* index is computed as the geometric mean using the technology from period  $t$  and period  $t + 1$ , respectively, as the base technology following the practice in the literature set by Färe et al (1992); (1994). However, we follow here the original Caves et al (1982b) using the technology from period  $t$  only as the base technology. Notice that there is a problem of comparability over time if the directions for each period are taken as the observations.

measure. “Punishing” generation of pollutants with the same factor as “rewarding” desirable goods is just one of many ways of implicitly “weighting” the two types of outputs. The fundamental definition of environmental damage we are seeking, may be far from the interpretation of efficiency measures based on directional distance functions. One cannot really know if such measures can be of any use in decision-making concerning environmental actions.

Most, may be all, of polluting industrial activities have available prices on desirable outputs and inputs, but not on pollutants. This has been utilised to estimate shadow prices on pollutants. However, this can at best only be used to estimate loss of desirable output when reducing output is the only abatement option.

Another weakness with the way single equation models based on distance function model that has been used is the lack of explicit modelling of abatement activities as is done in the model in Section 8.3.4. Measuring abatement in the form of lost output only is just taking into account one of the ways to reduce generation of pollutants mentioned in Section 8.3.3. It is standard in the environmental economics literature to model abatement as end-of-pipe as used in Section 8.3.4 (see Førsund (2009); 2017).

The crucial point for environmental policy is a trade-off between values of goods and pollutants. But this must be based on a trade-off between man-made desirable goods and jointly generated pollutants measured in the same unit. It is difficult to see that this can be captured by defining productivity as changes in directional distance functions with a fixed goods expansion factor and the same factor as a contraction factor of pollutants. Productivity measures consistent with real trade-offs are inevitably required by regulatory activities, carefully balancing benefits and costs (Jaffe et al 1995).

#### *The multi-equation model and inefficiency*

The model factorially determined multioutput used in Section 8.3 highlights insights from environmental economics. This model can easily be generalised to include inefficient operations. The simplest way is to use the three relations (8.5) and (8.9) depicting frontier functions, obeying the materials balance (8.7), by expanding the functions involved into sets:

$$\begin{aligned}
 y &\leq f(x_M, x_S) \\
 z^P &\geq g(x_M, x_S) \\
 c^A &\geq c(a) = c(z^P - z^S)
 \end{aligned}
 \tag{8.29}$$

$c^A$  is the observed cost. The production and cost frontiers are on the right-hand side of the inequalities. Strict inequalities open for inefficient operations while equations holding with equalities means efficient operations. Notice that introducing abatement activity makes it necessary to operate with two types of emission; primary pollutants and secondary pollutants. We can distinguish between three types of efficiency:

- a) Efficiency in producing desirable outputs
- b) Efficiency in generating pollutants
- c) Efficiency in abatement cost

Murty et al (2012) work with two production possibility sets corresponding to the two first relations in (8.5) or (8.29), but introduce abatement in a general way that hides the abatement activity as such. The intersection of these two sets represents the overall production possibility set.

Productivity development can be calculated using e.g. a Malmquist productivity change index<sup>20</sup> for each of the three equations (8.29). A problem is that the observations are generated simultaneously from the system (8.29), so there are cross-links between variables. This means that when calculating separate productivity changes each of the productivity developments are conditional on the other developments. Causal interpretations of the productivity developments are then difficult, but the separate developments will still give us useful information about productivity change for each output and abatement costs, and the development of frontiers. Keeping the developments separate also underlines the point that to see the trade-off between desirable and undesirable goods is only something that can be achieved if the effects of pollution in the environment are made comparable with man-made goods by measuring each type of output in the same unit. It does not make much sense to try to capture a trade-off by forcing a numerical trade-off not based on real evaluation of values.

---

<sup>20</sup> The original popular Caves et al (1982b) Malmquist productivity change index is not a true total factor productivity index. However, the total factor productivity index developed in Bjurek (1996) is a productivity change index defined as a Malmquist volume index for outputs in the numerator and a Malmquist input index in the denominator and is therefore a true total factor productivity index. This index is also called the Hicks-Moorsteen index following a suggestion in Diewert (1992).

## 8.6. Conclusions

The main question has been how to include considerations of the natural environment when measuring productivity of production and its change over time. The natural environment provides us with services of various kinds influencing our health and wellbeing in a profound way. However, many of these services are not traded in markets and also have a public good character. It may therefore be futile to search for prices of these services. What can be done for monetised evaluation is to estimate the damages inflicted on the environment as evaluated by consumers of environmental services. Concentrating of how to deal with pollution from production using material resources, estimates of damages constituting the damage function are based on the value of man-made desirable goods that we are willing to forsake in order to reduce the pollution from residuals that are generated jointly with the desirable outputs.

The guiding principle for formulating a model describing joint generation of intended and unintended outputs has been the materials balance principle. Matter cannot be created or disappear, so matter of raw materials must either be part of the desirable outputs or end up as residuals discharged to the environment. A multi equation model is then most suitable to be used. The choice was the Factorially determined multioutput model; the Frisch pollution model for short, having separate production functions for desirable outputs and residuals. This model does not allow any relationship between the desirable output and the undesirable one for given amounts of inputs; just a single point in output space is determined, thus obeying the materials balance.

Productivity is measured as the ratio between outputs and inputs. Outputs are either value added or gross output, and volume indexes have to be constructed in the multioutput case. Inputs can be just labour, or comprising all of them in the case of multifactor productivity pursued here. Environmental concerns are then entered forming the social gross output subtracting damage costs from gross “economic” outputs. For constant outputs and inputs static productivity has then to go down. However, once the adjustment is done productivity change may be both positive and negative depending on the time path of the different components. Within a general equilibrium framework resources are channeled to pollution-reducing activities, thus leading to a general reduction of man-made goods, but it is still possible for productivity change incorporating damage cost to show an increase. However, capital accumulation may go down

in general and being a brake on productivity growth if not compensated sufficiently by increase in environmental qualities (i.e. a reduction in damage cost).

Generation of pollutants can be reduced in this model by substitution between inputs (fuel substitution, and using non-material inputs instead of material ones) keeping desirable output constant, and also reducing the amount of desirable outputs. Another way of abatement is changing technology, but usually possible in the long run only, but there is also an option of installing end-of-pipe facilities that can be done within a shorter time horizon. Such a facility is added to the Frisch pollution model. As to productivity abatement costs represents additional inputs reducing static productivity, but again productivity change may go either way; this is an empirical question also influenced by technical change in the long run.

The Porter hypothesis postulates environmental regulation based on the “right” incentives to be a win-win situation; regulation put pressure on firms to reduce current inefficiency and innovate to such an extent that profit and productivity increases. However, there is not much systematic evidence for such a Panglossian view. No proper formulations of formal dynamic mechanisms have been put forward.

Environmental considerations have also been developed within literature on inefficiency. The bulk of inefficiency modelling has used a single equation model based on a special assumption called weak disposability linking desirable and undesirable outputs to move proportionally at the efficient border of the production possibility set allowing inefficient operation to take place. Weak disposability (together with null jointness) blocks the reduction of pollution to zero within the adapted assorted production framework, but no economic or engineering explanation has been put forward backing up the assumptions. The physical background for generation of residuals when using raw materials is not commented upon at all, and the materials balance is neglected. Indeed, the weak disposability goes counter to the materials balance principle because it is based on a transformation possibility (assorted production) between the desirable and undesirable outputs for given amount of resources. However, it is impossible to decrease both desirable and undesirable outputs simultaneously for given resources. Saying that resources are just made idle goes against the definition of an efficient border of the set, and telling a story that resources in production are reduced because they are reallocated to abatement without modelling such activity is just wishful thinking.

Calculation of productivity and its change are done applying directional distance functions within the weak disposability framework and using a Malmquist type of productivity measure,

To make productivity of desirable goods comparable with pollutants the expansion factor for undesirable outputs are restricted to be the same as for desirable outputs, but with a negative sign. But this is just an arbitrary way of comparing growth having nothing to do with a proper evaluation of pollutants and desirable goods; the task is to make desirable man-made goods comparable with environmental services in value terms.

The strong conclusion of this chapter is that the use of single equation models based on the assumption of weak disposability models is at the end of its road. Multi-equation models capturing essential mechanisms behind simultaneous generation of pollutants and desirable goods should be developed.

The solution to the problem of incorporating pollution into productivity measures has been based on introducing an environmental damage function. However, although a neat theoretical solution, it may not be so helpful for practical calculation. Further building up of satellite accounts tracking physical changes in various types of natural environments and stocks of resources, both biological renewable ones and non-renewable resources, is needed. An alternative to estimating damage costs is to impose physical limits on discharge of pollutants if the physical damages are known well enough. As demonstrated with cap and trade schemes setting physical limits imply that mitigation costs are revealed, thus making a trade-off between man-made goods and pollutants possible.

## References

Aaheim A and Nyborg K (1995) On the interpretation and applicability of a “green national product”. *Review of Income and Wealth* 41(1), 57-71

Asheim GB (2000) Green accounting: why and how? *Environment and Development Economics* 5(1), 25-48

Ambec S and Barla P (2002) A theoretical foundation of the Porter hypothesis. *Economics Letters* 75(3), 355-360

Ambec S, Coheny MA, Elgiez S, and Lanoie P (2013) The Porter hypothesis at 20: can environmental regulation enhance innovation and competitiveness? *Review of Environmental Economics and Policy* 7(1), 2-22

Ayres RU and Kneese AV (1969) Production, consumption and externalities. *American Economic Review* 59(7), 282–297



- Balk BM (2016) Empirical productivity indices and indicators. In *The Oxford handbook of productivity analysis*. Grifell-Tatjé E, Lovell CAK and Sickles RC (eds.). Chapter 2, x-y. Oxford: Oxford University Press
- Bjurek H (1996) The Malmquist Total Factor Productivity Index. *Scandinavian Journal of Economics* 98(2), 303–313
- Brännlund R and Lundgren T (2009) Environmental policy without cost? A review of the Porter hypothesis. *International Review of Environmental and Resource Economics* 3(1), 75-117
- Caves DW, Christensen LR and Diewert WE (1982a) Multilateral comparisons of output, input, and productivity using superlative index numbers. *Economic Journal* 92, 73-86
- Caves DW, Christensen LR and Diewert WE (1982b) The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica* 50(6), 1393-1414
- Charnes A, Cooper WW, and Rhodes E (1978) Measuring the efficiency of decision making units. *European Journal of Operational Research* 2(6), 429-444
- Cropper ML and Oates WE (1992). Environmental economics: a survey. *Journal of Economic Literature* 30(2), 675-740
- Chung YH, Färe R and Grosskopf S (1997) Productivity and undesirable outputs: a directional distance approach. *Journal of Environmental Management* 51(3), 229-240
- Dakpo KH, Jeanneaux P and Latruffe L (2016) Modelling pollution-generating technologies in performance benchmarking: recent developments, limits and future prospects in the nonparametric framework. *European Journal of Operational Research* 250(2), 347-359
- Diewert WE (1992) Fisher ideal output, input and productivity indexes revisited. *Journal of Productivity Analysis* 3(3), 211–248
- Färe R, Grosskopf S, Lovell CAK and Pasurka C (1989). Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach. *Review of Economics and Statistics* 71(1), 90–98
- Farrell MJ (1957). The measurement of productive efficiency of production.” *Journal of the Royal Statistical Society, Series A*, 120(III), 253–281
- Fisher AC and Peterson FM (1976). The environment in economics: a survey. *Journal of Economic Literature* 14(1), 1-33
- Frisch, R (1935) The principle of substitution. An example of its application in the chocolate industry. *Nordisk Tidskrift for Teknisk Økonomi* 1(1-2), 12 - 27
- Frisch, R (1965). *Theory of Production*, Dordrecht: D. Reidel
- Førsund FR (1985) Input-output models, national economic models, and the environment. In *Handbook of natural resource and energy economics, vol. I*, edited by AV Kneese and JL Sweeney. Amsterdam: Elsevier Science Publishers BV
- Førsund FR (1999) On the contribution of Ragnar Frisch to production theory. *Rivista Internazionale di Scienze Economiche e Commerciali (International Review of Economics and Business)* 46(1), 1-34

- Førsund FR (2009) Good modelling of bad outputs: pollution and multiple-output production. *International Review of Environmental and Resource Economics* 3(1), 1-38
- Førsund FR (2010) Dynamic efficiency measurement. *Indian Economic Review* 45(2), 123-157, also published as Chapter 4 (p. 187-219) in Ray SC, Kumbakahr SC, Dua P (Eds): *Benchmarking for performance evaluation. A frontier production approach*, Springer India (2015).
- Førsund FR (2017) Multi-equation modelling of desirable and undesirable outputs satisfying the material balance. *Empirical Economics*, published online 12 January 2017, DOI 10.1007/s00181-016-1219-016)
- Førsund FR and Hjalmarsson L (1974) On the measurement of productive efficiency. *Swedish Journal of Economics* 76 (2), 141-154
- Førsund FR and Strøm S (1976) The generation of residual flows in Norway: an input-output approach. *Journal of Environmental Economics and Management* 3(2), 129-141
- Førsund FR and Strøm S (1988) *Environmental economics and management: pollution and natural resources*. London: Croom Helm. Also published in the series Routledge Revivals, 2013. Abingdon-New York: Routledge
- Hartwick JM (1977) Intergenerational equity and the investment of rents from exhaustible resources. *American Economic Review* 67 (5), 972-974
- Hartwick JM (1990) Natural resources, national accounting and economic depreciation. *Journal of Public Economics* 43(3), 291-304
- Hicks JR (1932) *The theory of wages*. (Second edition 1963). London: MacMillan
- Jaffe AB and Stavins RN (1995) Dynamic incentives of environmental regulations: the effects of alternative policy instruments on technology diffusion. *Journal of Environmental Economics and Management* 29(3), S-43 - S-63
- Jaffe AB, Peterson SR and Portney PR (1995) Environmental regulation and the competitiveness of U.S. manufacturing: what does the evidence tell us? *Journal of Economic Literature* 33(1), 132-163
- Jorgenson D (2016) Measuring productivity. In *The Oxford handbook of productivity analysis*. Grifell-Tatjé E, Lovell CAK and Sickles RC (eds.). Chapter 24, x-y. Oxford: Oxford University Press
- Kneese AV (1971) Background for the economic analysis of environmental pollution. *Swedish Journal of Economics*. Special issue on environmental economics 73(1), 1-24
- Kneese AV, Ayres RU and d'Arge RC (1970) *Economics and the environment. A materials balance approach*. Baltimore: Johns Hopkins Press
- Lanoie P, Patry M, and Lajeunesse R (2008) Environmental regulation and productivity: testing the Porter hypothesis. *Journal of Productivity Analysis* 30(2), 121-128

- Lanoie P, Laurent-Lucchetti J, Johnstone N, and Ambec S (2011) Environmental policy, innovation and performance: new insights on the Porter hypothesis. *Journal of Economics and Management Strategy* 20(3), 803–842
- Leibenstein H (1966) Allocative efficiency vs. 'X-efficiency'. *American Economic Review* 56(3), 392-415
- Leontief W (1970) Environmental repercussions and the economic structure: an input-output approach. *The Review of Economics and Statistics* 52(3), 262-271
- Leontief W (1974) Environmental repercussions and the economic structure: an input-output approach: a reply. *The Review of Economics and Statistics* 56(1), 109-110
- Leontief W and Ford D (1972) Air pollution and the economic structure: empirical results of input – output computations. In Brody A and Carter A (eds). *Input – output techniques*. Amsterdam-London: North-Holland, 9-30
- Mäler K-G (1991) National accounts and environmental resources. *Environmental and Resource Economics* 1(1), 1-15
- Martin RE (1986) Externality regulation and the monopoly firm. *Journal of Public Economics* 29(3), 347–362
- Meade JE (1952) External economies and diseconomies in a competitive situation. *Economic Journal* 62 (245), 54-67
- Mishan E J (1971) The postwar literature on externalities: an interpretative essay. *Journal of Economic Literature* 9(1), 1-28
- Murty S, Russell RR and Levkoff SB (2012) On modelling pollution-generating technologies. *Journal of Environmental Economics and Management* 64(1), 117-135
- Nordhaus WD and Kokkelenberg EC (Eds) (1999) *Nature's Numbers: Expanding the National Economic Accounts to Include the Environment*. National Research Council. Washington, DC: The National Academies Press, doi:10.17226/6374
- O'Donnell C (2016) Linear programming, least squares and maximum likelihood estimation of total factor productivity indexes. In *The Oxford handbook of productivity analysis*. Grifell-Tatjé E, Lovell CAK and Sickles RC (eds.). Chapter 5, x-y. Oxford: Oxford University Press
- Palmer K, Oates WE, and Portney PR (1995) Tightening environmental standards: the benefit-cost or the no-cost paradigm? *Journal of Economic Perspectives* 9(4), 119-132
- Pigou AC (1913) *The economics of welfare*. London: Palgrave Macmillan
- Perman R, Ma Y, Common M, Maddison D and McGilvray J (2011) *Natural resource and environmental economics*, 4<sup>th</sup> edition. Harlow: Pearson education Limited
- Pethig, R (2006) Non-linear production, abatement, pollution and materials balance reconsidered. *Journal of Environmental Economics and Management* 51, 185-204

- Pittman RW (1983) Multilateral productivity comparisons with undesirable outputs. *The Economic Journal* 93 (372), 883-891
- Porter ME (1991) America's green strategy. *Scientific American* April, 168
- Porter ME and van der Linde C (1995) Toward a new conception of the environment-competitiveness relationship. *Journal of Economic Perspectives* 9(4), 97-118
- Ruser J and Stewart J (2016) Measuring productivity in the United States. In *The Oxford handbook of productivity analysis*. Grifell-Tatjé E, Lovell CAK and Sickles RC (eds.). Chapter 3, x-y. Oxford: Oxford University Press
- Russell RR (2016) Theoretical productivity indexes. In *The Oxford handbook of productivity analysis*. Grifell-Tatjé E, Lovell CAK and Sickles RC (eds.). Chapter 4, x-y. Oxford: Oxford University Press
- Shephard RW (1970) *Theory of Cost and Production Functions*, Princeton NJ: Princeton University Press.
- Shephard RW (1974) *Indirect production functions*. Meisenheim Am Glan: Verlag Anton Hain
- Solow RM (1956) A contribution to the theory of economic growth. *Quarterly Journal of Economics* 70(1), 65-94
- Solow RM (1974) Intergenerational equity and exhaustible resources. *Review of Economic studies. Symposium on the Economics of Exhaustible Resources*, 41(5), 29-46
- Stern NH (2006) *The economics of climate change*. HM Treasury- 2006, hm-treasury.gov.uk
- Stigler GJ (1976) The Existence of X-efficiency. *American Economic Review* 66(1), 213-216
- Stiglitz JE, Sen A, and Fitoussi J-P (2009) *Report by the commission on the measurement of economic performance and social progress*.  
<http://graphics8.nytimes.com/packages/pdf/business/Stiglitzreport.pdf>
- UN (1993) *System of accounts 1993*. United Nations Department for Economic and Social Information and Policy Analysis – Statistics Division. Brussels/Luxembourg, New York, Paris, Washington D. C.
- Weitzman M (1976) On the welfare significance of national product in a dynamic economy. *Quarterly Journal of Economics* 90(1), 156-162
- Xepapadeas A and deZeeuw A (1999) Environmental policy and competitiveness: the Porter hypothesis and the composition of capital. *Journal of Environmental Economics and Management* 37(2), 165-182