DO ENVIRONMENTAL REGULATIONS LEAD FIRMS INTO TROUBLE? EVIDENCE FROM A TWO-STAGE DISTANCE FUNCTION MODEL WITH PANEL DATA

Lauri Hetemäki

The Finnish Forest Research Institute, Research Papers 509
(Metsäntutkimuslaitoksen tiedonantoja 509)

Department of Forest Resources
Helsinki 1994

DO ENVIRONMENTAL REGULATIONS LEAD FIRMS INTO TROUBLE? EVIDENCE FROM A TWO-STAGE DISTANCE FUNCTION MODEL WITH PANEL DATA

Lauri Hetemäki

The Finnish Forest Research Institute, Research Papers 509
(Metsäntutkimuslaitoksen tiedonantoja 509)

Department of Forest Resources
Helsinki 1994

Hetemäki, L. 1994. Do environmental regulations lead firms into trouble? Evidence from a two-stage distance function model with panel data. The Finnish Forest Research Institute, Research Papers 509. ISBN 951-40-1375-1, ISSN 0358-4283.*

Environmental regulations force firms to adjust their production processes. The purpose of the present study is to examine to what extent these adjustments affect firms' revenues and performance in general. The theoretical framework is based on the output distance function, which is used to derive the shadow prices of pollution and the impact of pollution reduction on firms' performance. The present study differs from the previous literature in that it uses a stochastic distance function rather than a deterministic one, by employing plant level panel data and imposing no a priori restrictions on the values of the shadow prices of pollution. Also, a novel two-stage approach which combines deterministic nonparametric linear programming with a stochastic econometric model is used to estimate the distance function. These extensions enable the estimation of a more general model with more robust parameter estimates, whose statistical significance can be explicitly assessed. The method is used to analyze the effects of water pollution regulations on eight pulp plants in Finland observed over a period of 19 years (1972-90).

Keywords: *distance function, shadow prices of undesirable outputs, panel data*

Publisher: Finnish Forest Research Institute; project KT 3013
Authors's address: The Finnsih Forest Research Institute, Unioninkatu 40 A, FIN-00170 Helsinki, Finland (internet: hetemaki@metla.fi)

*The paper is a revised version of a draft presented at the TIMS XXXII Conference, Anchorage, Alaska, June 12-15, 1994. I am grateful for helpful comments offered by the conference participants. I would also like to thank Robert Chambers, Mika Hakuni, Pekka Ilmakunnas, Jari Kuuluvainen and Markku Rahiala for their comments. The nonparametric distance scores were computed by Mika Hakuni. This research was supported by funds provided by the Academy of Finland, Maj and Tor Nessling Foundation and Nordic Environmental Research Program (NERP).

1. Introduction

In the last decade or so, there has been a surge of empirical literature which attempts to explicitly incorporate and assess the impacts of environmental regulations on producer behavior (e.g., Barbera & McConnell 1990, Conrad & Morrison 1989, Gollop & Roberts 1983, Jorgenson & Wilcoxen 1990, Myers & Nakamura 1980, Pittman 1981). This is a natural outcome of the increasing importance of environmental regulations and of the fact that pollution (undesirable output) can no longer be disposed of without cost. Firms have to adjust their production processes in order to reduce pollution to permitted levels or pay additional charges for their effluents. In the literature, interest has centered on the assessment of the effects of the regulations on productivity, efficiency, economic growth and factor demand, among other issues. The "standard approach" for including the environmental aspect in this context has been to add an additional input, such as abatement capital, to the producer's factor demand function or to introduce an environmental tax parameter for some input or output. Therefore, the inclusion of regulation or effluents has not essentially changed the conventional empirical neo-classical models of producer behavior.

However, Shephard (1970, 1974) noted earlier that the conventional assumptions and models of producer behavior should be modified in the case where the production process generates undesirable outputs which cannot be freely disposed.[1] Indeed, for many production processes which generate regulated undesirable outputs, the conventional single-product firm framework is not appropriate. Pollution (undesirable output) is often a side product created as a result of the production of the "good" (desirable) output. In other words, pollution does not necessarily enter the production process as an input but is instead created jointly during the production process. Consequently, there is generally a trade-off (marginal rate of transformation) in production between the desirable and undesirable outputs, the quality and quantity of which is of central importance when assessing the impact of pollution control. According to Shephard (1970), the polluting firm's production process is more accurately modeled as a production of multiple outputs, with strong disposable desirable outputs and weak disposable undesirable outputs. Färe et al. (1993) have extended Shephard's framework in order to examine the effects of waste water regulations on U.S. pulp mills. They use the output distance function as an analytical tool for representing the production technology and derive shadow prices for undesirable outputs.

Besides addressing the above issue concerning the way pollution is modeled in the production process, the empirical literature to date has often been based on somewhat unsatisfactory data. Due to a lack of micro level data, these studies have generally concentrated on measuring the effects of environmental regulations or pollution control using industry or country level aggregate data. However, if pollution regulations are set, for example, at the plant level, it is also important to be able to measure the effects of pollution control at the plant level. Firm or plant level regulations are particularly common for point source waste water effluents. For example, in the pulp and paper industry in Finland and Sweden, every plant is individually regulated with respect to the substances that it discharges. Similar practices are also common in many industries in the United States and Canada. Besides the "conventional" aggregation bias, using industry level data in such cases causes identification problems and measurement errors.

The purpose of the present study is to examine the impact of water pollution control on the production technology of the Finnish pulp industry and to derive a measure for the cost of reducing different water pollution effluents using the theoretical approach of Shephard (1970) and Färe et al. (1993). However, while Färe et.al. use deterministic linear programming analysis and plant level cross section data, the present study is based on the *stochastic* distance function and plant level *panel data*. Moreover, Färe et al. restrict the shadow prices of undesirable outputs to be nonpositive, whereas no such a priori restriction is set in the present study. Finally, the estimation procedure used in the present study differs from the existing stochastic distance function studies in that it combines nonparametric linear programming with a stochastic econometric model. The present approach allows one to relax the assumption that plants operate on the technology frontier.

2. Methodological background

Although the assessment of pollution control costs is straightforward in principle, in practice the problem is that the regulatory agency rarely knows the marginal abatement costs of individual plants. In general, reliable data on abatement costs and the resulting reduction in effluents are not available. Thus, one usually has to try to infer the marginal treatment costs indirectly. As Färe et al. (1993) show, one possible way to derive the marginal treatment costs indirectly is to use the duality of the *output distance function* and the revenue function. This approach also provides information on the production technology of the plants. Furthermore, they assume *weak disposability* of undesirable (bad) outputs, which allows them to model the fact that regulations restrict the firms' ability to costlessly dispose of effluents. The output distance function combined with the weak disposability assumption allows the computation of shadow (virtual) prices of pollutants without requiring detailed information about the actual regulations or abatement costs. Furthermore, the method allows one to identify the shadow prices of pollution at the level of the individual plant. These shadow prices reflect the impact of pollution control (or environmental regulations) on a plant and indicate to what extent the revenues and performance of the plants are affected by pollution control measures.

The output distance function has important advantages over more traditional means of representing production technology. In comparison to a production function, a distance function allows one to model multiple output and joint production technologies. On the other hand, the advantage of the distance function over cost, profit and revenue functions (which can also be readily used to model multiple output technology) is that no maintained behavioral hypothesis (cost min. or profit/revenue max.) is required. A distance function only identifies the technology frontier and gives the distance to the frontier for each observation. As a result, the different measures of economic effects (e.g., substitution) are not conditional on the behavioral hypothesis. Also, of great practical importance is the property that the distance function can be computed with data on quantities of inputs and outputs alone; prices are not needed.

In empirical applications of distance functions, it has been common practice to use deterministic linear programming, and so the econometric approach has rarely been used. Indeed, Lovell et al. (1990, footnote 6) state that "Although empirical computation of distance functions using econometric techniques is certainly in its infancy, mathematical programming techniques have been used to calculate distance functions for many years now."

As far as we know, the only empirical applications of the parametric output distance function in the context of the environmental regulations literature are Färe et al. (1993) and Hetemäki (1994). As was stated above, the Färe et.al. study is deterministic. The problem with the deterministic model is that the computed parameters are affected by random factors not controlled by the plants, which may be numerous and complex and not observable, and hence not measurable. On the other hand, the stochastic model allows for random error and so also permits direct the testing for statistical significance and consistency of the estimates. Furthermore, there may be production processes for which it may be inappropriate to restrict the shadow prices of undesirable outputs to be nonpositive, as Färe et al. (1993) do.[2] In any case, because the axioms behind the theoretical model do not require such a restriction, it is unecessary to impose it. Consequently, in the present study the shadow prices of "bad outputs" are not restricted to be nonpositive. Finally, using panel data rather than cross section data, provides more precise and consistent estimation even in the presence of correlated plant-specific effects (Hsiao 1986). Naturally, panel data also allow one to examine the impacts of pollution control over time.

The approach used here differs from the existing stochastic distance function studies in that the empirical analysis is based on a two-stage approach. In the first stage, a nonparametric linear programming model is used to compute the distance measures (efficiency scores) for each plant. In the second stage, these distance measures are used as a dependent variable in the stochastic output distance function model. The main advantage of this approach is that the assumption that the producers are operating on the production frontier (i.e., the value of the distance function is set equal to 1) can be relaxed. This provides a more realistic approximation of the production technology for most applications.

2. Theoretical model

The initial incidence of much of pulp industry pollution control falls on the firms. The ultimate incidence depends on the ability of firms to shift the cost burden to consumers by raising prices or to workers or other factors of production. The theoretical model in this study focuses on the initial incidence of pollution control and therefore is restricted to a partial equilibrium analysis of how a firm (or plant) reacts to a change in its cost structure when faced with a stable demand curve. Within this framework, the output distance function can describe the effects of pollution control on the production technology of a firm. The conventional production function gives the maximum output obtainable from a given input vector. The distance function is a generalization of this notion, serving as a functional representation of the output set in the context of multi-outputs (Shephard 1970, Färe 1988).

A production technology transforming factors of production $x = (x_1, x_2, \ldots, x_n) \in R_+^n$ into outputs $y = (y_1, y_2, \ldots, y_m) \in R_+^m$ can be modeled by the output set P(x). The output set contains all technically feasible output vectors for the input vector x, i.e., $P(x) = \{y \in R_+^m : x \text{ can produce } y\}$. It is assumed that the technology satisfies the maintained axioms of Färe (1988, p.6). In particular, outputs are assumed to be only *weakly disposable*. Conventionally, the assumption of *strong* (or free) *disposability* is made.[3] However, for the pulp industry, it is unlikely that the production of pulp is characterized by strongly disposable outputs given that regulations do not allow water pollution to be "thrown away" (freely disposed). The output distance function is defined on the output set P(x) as

$$(2.1) \quad D_o(x,y) = \min_{\theta}\{\theta:(y/\theta) \in P(x)\}$$

Equation (2.1) gives the largest radial expansion of the output vector, for a given input vector, which is consistent with that output vector belonging to P(x). The axioms regarding the output set P(x) impose a set of properties on the output distance function (for a detailed description, see Färe 1988, pp. 31-34). The value of the output distance function must be less than or equal to one ($D_o \leq 1$) for feasible output. Further, the value of the distance function is the reciprocal of the *Farrell output-based technical efficiency index* (Färe 1988).

The revenue function defined as (Shephard 1970 , Färe 1988),

$$(2.2) \quad R(x,r) = \max_{y}[ry: y \in P(x)]$$

can also completely describe the production technology, where the output price vector is denoted by $r = (r_1,....,r_m)$ and it is assumed that r can be nonpositive. The revenue function describes the maximum revenue that can be obtained from the given technology at output prices r. Shephard (1970) showed that the revenue function and output distance function are dual. Consequently, we can define the *revenue function* in terms of the distance function and vice versa. Formally,

$$(2.3a) \quad R(x,r) = \max_{y}\{ry: D_o(x,y) \leq 1\}$$

$$(2.3b) \quad D_o(x,y) = \max_{r}\{ry: R(x,r) \leq 1\}$$

The duality theorem shows that the revenue function can be derived from the output distance function by "maximizing" revenue over output quantities and that the output distance function is obtained from the revenue function by maximizing over output prices.

Following the analysis of Shephard (1970) and Färe et.al. (1993), it can be shown that the *revenue deflated* output shadow prices (r_m^s) for each observation can be derived as the derivative of the distance function (using dual Shephard's lemma), i.e.,

$$(2.4)\quad r_m^s(x,y) = \partial D_o(x,y)/\partial y_m = r_m^s(x,y)/R^*(x,r^s), \quad m = 1,2,......,M,$$

where $R^*(x,r^s)$ denotes shadow revenue. Färe et al. (1993) show that the *absolute* shadow prices can be computed when maximal revenue R(x, r) is known. In order to obtain these shadow prices, the following assumption is used (following Färe et. al. 1993): *One observed output price equals its absolute shadow price.* This assumption implies that at least one output market is efficient and it allows different plants to face different competitive markets. Alternatively, one could assume that in one output market observed revenue equals maximum revenue. Let output 1 denote the good output and assume that the observed good output price (r_1^o) equals its absolute shadow price (r_1^s), i.e., for m =1, $r_1^s = r_1^o$. The maximum revenue is then given by

$$(2.5)\quad R^*(x,r^s) = r_1^o / (\partial D_o(x,y)/\partial y_1)$$

The absolute shadow prices for each observation of undesirable outputs (m = 2,.....,M) can now be computed as

$$(2.6)\quad r_m^s = R^*(x,r^s) \bullet (\partial D_o(x,y)/\partial y_m) = r_1^o \bullet \frac{\partial D_o(x,y)/\partial y_m}{\partial D_o(x,y)/\partial y_1}$$

In equation (2.6) the ratio of output shadow prices reflects the relative opportunity cost of the outputs, i.e., they are equivalent to the marginal rate of transformation. It may be noted that the above expression does not require information on regulatory constraints. This is important, because we often do not have data on regulations, and even such data exists, the plants rarely operate exactly at the level of the constraint. Thus, "shadow prices reflect the tradeoff between desirable and undesirable outputs *at the actual mix of outputs,* which may or may not be consistent with the maximum allowable under regulation" (Färe et.al. 1993, p. 376). Further, the shadow prices do not require the plants to operate on the production frontier. Since the output distance function is homogeneous of degree +1 in outputs, the derivatives, which give the shadow prices, are homogeneous of degree zero with respect to proportional scaling of outputs. Since output distance function is such a proportional scaling of outputs, the shadow prices are independent of whether the observations are on the frontier (see Grosskopf and Hayes 1993).[4]

3. Empirical model

Let the production technology of the pulp plants be represented by the output distance function

(3.1) $D_O = f(X, Y; Y) \exp \varepsilon$

where D_0 is the distance measure, f(.) is the production technology, X is a matrix of inputs, Y is a matrix of outputs, Ψ is a vector of parameters to be estimated and ε is the error term. Estimation of a distance function, like (3.1), raises several econometric problems, and presumably because of these, there have been very few econometric applications of distance functions. The basic problem with distance functions, as concerns econometric estimation, is that one does not observe (have data on) the dependent variable. Further, if one sets the distance function equal to its efficient (frontier) value, D_0 = 1, the left-hand side of the distance function is invariant, an intercept cannot be estimated, and OLS parameter estimates will be biased. Further, if the distance function is expressed in logarithms, the left-hand side of the distance function will be zero for all observations (i.e., D = ln(1) = 0).

In order to avoid the above problems, Lovell et al. (1990) and Grosskopf et.al. (1992) utilize the property that the output distance function is homogeneous of degree +1 in outputs (see, also Grosskopf and Hayes 1993). Thus, for each observation to be used in estimating the distance function, a value that is unique to that observation can be used to multiply all output values on the right hand-side and the value of the distance function on the left-hand side. However, this transformation may cause estimation problems. After the transformation, the multiplicative variable appears on both the left -and right-hand sides of the equations, which may result in endogeneity on the right-hand side. Thus, one has to test whether the errors are correlated with the regressors and, if they are, instrumental variables estimation must be used.

Another problem in estimating the output distance function is that in theory the value of the distance function should never exceed that for plants operating on their frontier. However, in the estimation of equation (3.1) an error term with mean zero, but positive variance, is assumed. For some plants the forecasted value of the output distance function can therefore exceed the theoretically plausible value. To account for this problem, it is common (e.g., Lovell et.al. 1990, Grosskopf et al. 1992) to use the method known as corrected ordinary least squares (COLS). This amounts to calculating first the most negative residual from the estimated output distance function and then adding that residual to the intercept term so that the corrected estimates of the output distance function never exceeds the theoretically plausible value for any plant. In other words, this ensures that all observations are enveloped from above.

In the present study, a two-stage approach is used to estimate the output distance function. In short, the procedure consists of two steps: first, the measure for each plant's distance to the reference production frontier is computed using a deterministic nonparametric piecewise linear model, which treats desirable and undesirable outputs differently. Second, the distance measures computed in the first stage are used as a dependent variable in a parametric stochastic distance function model. By using this approach, one can relax the assumption that all plants are operating on the frontier. Existing studies have either assumed that all observations have an equal distance measure of 1, or inefficiency has been introduced afterwards by means of a composed error structure, as in Grosskopf and Hayes (1993). Naturally, the larger the dispersion of the distance measures from 1, the more biased the results obtained using methods that set the distance scores equal to 1. Thus, in order to not restrict the technology to being on the frontier, we compute the so-called *Farrel output efficiency measure*, *F*, which has been shown to be the reciprocal of the output distance function (e.g., Färe 1988, p. 135). Consequently, the dependent variable of the stochastic output distance function can be computed in a theoretically consistent way. From a slightly different perspective, the two-stage procedure consists of first constructing the production frontier using a nonparametric linear programming model and then approximating this frontier by a smooth, parametric functional form, which provides *additional and economically and statistically interpretable results*. Thus, the frontier is not dependent on a presupposed functional form. The advantages of complementary use of nonparametric linear programming and stochastic econometric models have recently been noted also by Banker and Cooper (1994).

For the model specification with strongly disposable desirable outputs, weakly disposable undesirable outputs and variable returns to scale (with respect to inputs), *F* can be expressed as (for a more detailed description, see Färe et al. 1994, pp. 105-106)

$$(3.2) \qquad F_{vrs}^{k}(v^k, w^k, x^k) = \max \pi$$

$$\pi v^k \leq \phi z^1 v^1 + \phi z^2 v^2 + \dots + \phi z^k v^k$$
$$\pi w^k = \phi z^1 w^1 + \phi z^2 w^2 + \dots + \phi z^k w^k$$
$$x^k \geq z^1 x^1 + z^2 x^2 + \dots + z^k x^k$$
$$\sum_{k=1}^{K} z^k = 1, z^k \geq 0, 0 \leq \phi \leq 1,$$

where v and w denote the desirable and undesirable output subvectors of y, respectively, and z the intensity or scaling vector and k (k=1,....,K) denotes the observation. The above problem can be linearized by setting $\phi = 1$. This procedure does not affect the maximizing π, z values. Model (3.2) is a very general model, entailing many different models as special cases. Furthermore, the computed distance measures are not dependent on any particular functional form (except linearity). The *F* index is computed by comparing each input-output combination (x,y) to a reference technology set formed from all observations. The frontier consists of linear facets, which are determined by the efficient units of the data. In accordance with the output distance function, the reciprocal of *F* gets values $0 \leq F \leq 1$.

The potential weakness of the two-stage approach is that it may cause endogeneity of the right hand side variables. Since the *F* measure is computed using the same quantity and input data that are used as exogeneous variables in the stochastic model, the explanatory variables may be correlated with the equation error and the least squares estimates of the coefficients may be biased and inconsistent. However, it should be stressed that for a number of reasons the potential endogeneity problem may not be important. First, in the stochastic model the homogeneity restriction, which is not used in the first stage linear programming problem, may allow for the identification of the model. Secondly, the stochastic model is treated as a panel with plant -and period-specific fixed or random effects, which do not enter the linear programming model. Thirdly, the stochastic model is based on a parametric functional form, while the linear programming problem is nonparametric. In any case, one should check whether endogeneity is a problem and, if it is, use the instrumental variables method to estimate the model. Finally, it may be noted that the potential measurement errors in the linear programming distance scores do not necessarily lead to serious problems, since as long as the regressors are measured properly, nonsystematic measurement error on the dependent variable can be absorbed in the disturbance term of the regression without affecting the estimated parameters. Obviously, a large measurement error would lead to a large variance of the disturbance term and this would show up as a poor fit of the model.

In order to be able to estimate the parametric stochastic distance function, a functional form has to be chosen. In principle, it would be desirable to use as flexible a functional form as possible. Initially, a translog form including 28 parameters was tried. However, the translog model could not be estimated consistently due to multicollinearity.[5] Consequently, the estimated output distance function is a special case of a translog function, with a first-order approximation in the input quantities (Cobb-Douglas technology) and second -order terms in the output quantities; it is shown as eq. 3.3 (the same functional specification was used, e.g., by Simar 1992).

$$(3.3) \quad \ln D_{opt} = \alpha_0 + \sum_{n=1}^{N} \beta_n \ln x_{npt} + \sum_{m=1}^{M} \gamma_m y_{mpt} + 1/2 \sum_{m=1}^{M} \sum_{m'=1}^{M} \gamma_{mm'} (\ln y_m)(\ln y_{m'}) + \varepsilon_{pt}$$

where x_n denotes inputs, i.e., capital (K), labor (L) and materials (MA); y_m denotes desirable and undesirable outputs, i.e., quantity of pulp production (Q), biological oxygen demand (BOD) and waste water flow (FLOW); p indexes the plants (p = 1,2,..,P); and t denotes the time period (t = 1,2,...T). For the estimation, the homogeneity restriction (i) and the symmetry restriction (ii) are set.

(i) $\sum_{m=1}^{M} \gamma_m = 1, \quad \sum_{m'=1}^{M} \gamma_{mm'} = 0, \quad m = 1,...,M$

(ii) $\gamma_{mm'} = \gamma_{m'm}, \quad m = 1,...M, \quad m' = 1....M$

4. Institutional environment and data

4.1 Institutional background

The Finnish pulp and paper industry is an example of an industrial sector that has been of central economic importance and at the same time a major polluter of the environment. The per capita export revenues earned by the Finnish pulp and paper industry are among the largest in the world (FIM 6191 in 1990). However, it has also been the single major water polluting sector as well as the major user of forest resources. Consequently, conflicts between environmentalists and industrialists have often found a battle ground in the pulp and paper industry.

The present study concentrates on one particular pollution type, namely, water pollution. In the present study we examine a homogeneous group from the whole industry sector. The motive for this is that when the plants can be assumed to operate in the same environment and can truly be considered to use approximately the same technology, one can hope to be able to estimate relatively satisfactory results for the production technology.[6] Since sulphate pulp plants account for the bulk of waste water effluents generated by the overall pulp and paper industry, this pulp process was chosen as the object of the study. Further, the sulphate pulp industry represents a typical process industry, whose inputs and end products are relatively homogeneous compared with most other industries. Thus, the inputs and outputs are also relatively precisely measurable.

Sulphate pulp mills are usually classified as integrated pulp and paper plants or non-integrated pulp plants. The first group consists of plants in which the production process is integrated with the production of paper or paperboard and the latter group represents plants that produce only sulphate pulp (for export or for sale to domestic paper plants). In 1990 there were 17 sulphate pulp plants in Finland, of which 7 were non-integrated. A major part of the output is used domestically; of total output, exports accounted for 34 % in 1972, 38 % in 1980 and 26 % in 1990. However, of the end product (paper/paperboard) approximately 90 % is exported. The main water effluents produced jointly with pulp are biological oxygen demand (BOD), suspended solids (SS), nitrogen (N), phosphorous (P), chemical oxygen demand (COD), and absorbable organic halogens (AOX).[7]

Up until the beginning of the 1970s water pollution control was not a very serious problem for the Finnish pulp and paper industry because it did not have to invest very much of its resources in the control of waste water. However, public concern over the quality of water resources and the fact that the Finnish pulp and paper industry was the main polluter of water resulted in tighter control measures. The Finnish Water Act was drafted in 1962 and the first waste water permits were issued at the end of the 1960s. The regulation of effluent loading is based to a large extent on discharge permits and Water Rights Court procedures, both of which differ from common practice in most other countries. This is mainly because most waters in Finland are in private hands. Thus, for example, a fishing community owning the property rights to waters near the pulp plants have been able to receive compensation payments from the plants due to the reduction in the quality of their waters. The local water authority regulates effluents by specifying effluent limits (such as kg/ton of product) for each individual plant. According to the water law (VL 10:24), a permit is issued for such effluents as cannot be eliminated with *reasonable* costs. The authority has to consider the economic benefits of the firm against its costs to the environment. Thus, acceptable standards are based on individual judgments and often on compromise.

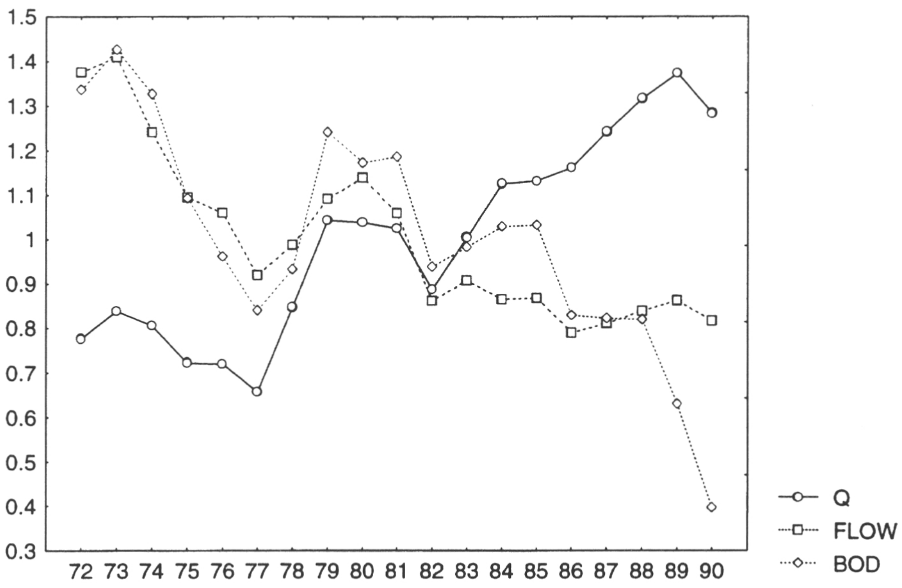
In practice, the regulations have been very heterogeneous across the plants. Documents from the water authorities indicate that it is difficult to find any clear evidence which would show that a common policy has been applied simultaneously to all plants. Permits have to be renewed typically after 3-10 years. In addition to standards, the regulatory authorities have controlled waste water effluents by giving tax credits for investment in pollution-control equipment and by ordering the plants to build special devices to clean the waters. However, the two latter regulation measures have had only a relatively small impact compared to the other measures (tax credits have amounted to a very small proportion of total investments in abatement capital).

4.2 Empirical data

In order to keep the sample as homogeneous as possible, only those plants were included which were operating during the whole period studied, in order to reduce the bias of comparing plants with different vintages of production technology. Moreover, the linear programming model requires a balanced panel. Consequently, the data sample contains annual data from 8 sulphate pulp plants observed over the period 1972-90 (a detailed description of the data is given in Appendix). All the plants are nonintegrated, except one, for which it was possible to separate the sulphate pulp production from the paper/paperboard production (in terms of the data needed). The plants in the sample have accounted for about one-half of the total production of the sulphate pulp industry during 1972-90. The pulp output and the effluents of the plants over time are shown in *Figure 4.1* (the series have been divided by their respective mean values). The figure shows that there has been a simultaneous increase in the pulp output and a decrease in effluents over time.

The data used for estimation consists of observations on quantity (Q) of sulphate pulp output, net fixed capital stock (K), hours worked (L), value of materials input (M), biological oxygen demand (BOD), total waste water flow (FLOW), and suspended solids (SS). Although, FLOW has not been regulated by the water authority, its reduction has nevertheless been one of the major means by which the plants have tried to reduce different water pollution substances. In particular, the reduction of FLOW describes the internal process changes adopted in order to reduce waste water. Moreover, the FLOW parameter is significantly correlated with a number of other effluents (phosphorous and nitrogen).

Figure 4.1 Pulp output (Q), waste water flow (FLOW), biological oxygen demand (BOD) (mean values)



The standard deviations, means, minimum and maximum values, skewness, and kurtuosis are shown in *Table 4.1.* The standard deviations for all variables are less than their mean values, indicating that the mills are a relatively homogeneous group. The skewness and kurtosis statistics are, with the exception of FLOW and SS, near 0 and 3, respectively. Indeed, the logarithms of the variables these statistics are very close to 0 and 3, indicating that the variables may be normally distributed, and therefore that the t-values could be used to make inferences concerning the statistical significance of the variables.

Table 4.1 Descriptive Statistics. 8 sulphate pulp mills observed annually between 1972 - 1990 (Sample Size = 152)

VARIABLE	UNIT	MEAN	ST. DEV	SKEW	KRT	MIN	MAX
PRODQ	1000 t	224.9	94.3	0.74	2.82	87.3	511.8
GVP90	mill. FIM	602.9	231	0.44	2.40	209.7	1200
MATER90	mill.FIM	420.2	165.4	0.33	2.28	137.6	848
WHOURS	1000 h	811.4	353.1	0.76	3.07	209.7	1803
CAPIT90	mill. FIM	974.2	372.3	-0.09	2.14	273.2	1797
FLOW	mill. m3	42.1	20.4	1.65	6.10	15.2	126.7
SS	t	1978.3	1569.3	2.55	8.23	277	9950
BOD	1000 t	6034.8	2988	0.57	3.22	554	15370

PRODQ = pulp output; GVP90 = gross value of output in 1990 prices; MATER90 = value of materials input in 1990 prices; WHOURS = hours worked (productive and non-productive workers); CAPIT90 = net fixed capital stock in 1990 prices; FLOW = waste water flow; SS = suspended solids; BOD = biological oxygen demand.

5. Results

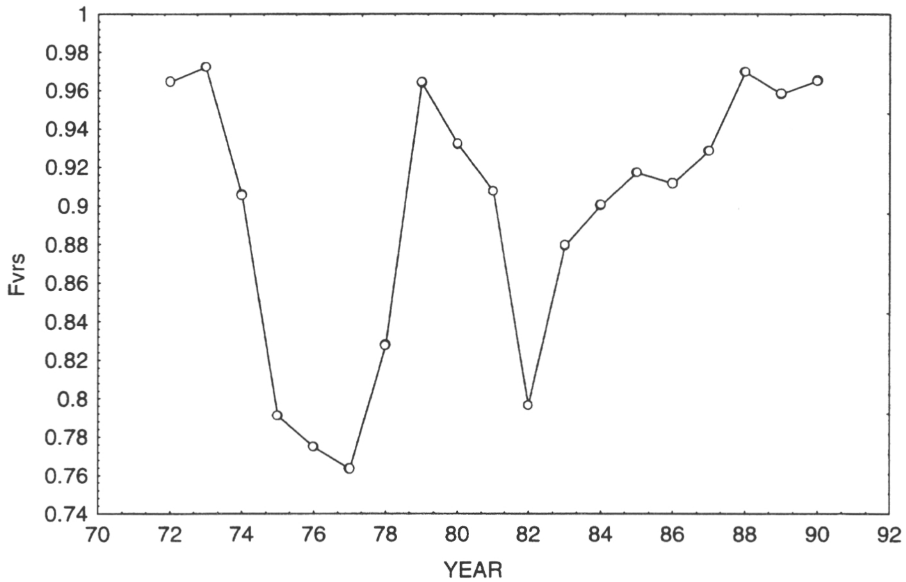
The nonparametric linear programming problem (3.2) was computed by treating the data set as a pooled sample, i.e. using all 152 observations at one time.[8] The results for the mean value for the whole sample and the mean values for each plant over time are summarized in *Table 5.1*. Besides providing the measures for the distances from the frontier, the results are interesting in their own right, since they can be interpreted as efficiency measures (see, e.g., Färe et al. 1989). The distance scores were computed for both the constant-returns-to-scale (F^{crs}) and variable returns to scale (F^{vrs}) models. The differences in the distance scores between the two models are small, the Pearson correlation coefficient between the two series is 0.94 and the correlation is significant at the 1 % level. The mean efficiency in the variable returns to scale is 0.90. Thus, the sulphate pulp production for plants in the sample could be increased by about 10 percent on average if all plants were to operate on the production possibility frontier. In a constant-returns-to-scale case, the mean efficiency is 0.87, implying that the current output could be produced with 13% lower cost. These results indicate that there are considerable gains to be made by improving the efficiency.

Table 5.1 Distance measures from nonparametric linear programming model

VARIABLE	MEAN	ST. DEV.	MIN	MAX	SKEW.	KURT.
F^{vrs}	0.90	0.12	0.54	1	-0.99	2.85
F^{crs}	0.87	0.13	0.51	1	-0.87	2.76
F^{vrs} plant 1	0.91	0.14	0.58	1		
F^{vrs} plant 2	0.92	0.10	0.60	1		
F^{vrs} plant 3	0.75	0.12	0.60	1		
F^{vrs} plant 4	0.94	0.10	0.73	1		
F^{vrs} plant 5	0.86	0.12	0.54	1		
F^{vrs} plant 6	0.87	0.11	0.70	1		
F^{vrs} plant 7	0.94	0.08	0.76	1		
F^{vrs} plant 8	0.97	0.06	0.81	1		

The variation of the mean efficiency over time is shown in *Figure 5.1*. The changes in efficiency over time can be roughly divided into two periods. The average inefficiencies have been the largest in 1973-77 and 1980-82, which are periods that coincide with economic slumps. In these years the plants' outputs were below their long-run trends. Thus, it appears that the adjustment to changes in the market environment happens partly through changes in efficiency. However, it should be stressed, that the above efficiency scores are here mainly a means to an end, i.e., they are used as the dependent variable in the estimation of stochastic distance function. From this standpoint, the results show that had we set the dependent variable equal to 1 in the stochastic output distance function model (i.e., used the frontier approach), it would have been a rather poor approximate of the actual distance measures (59 of the 152 observations have distance measure equal to 1).

Figure 5.1 Distance measures from the nonparametric linear programming model



Since the distance measure gets values between 0 and 1, with bulk of the values lying near 1, the distribution of the distance measure is truncated normal distribution. However, in the stochastic distance function model the homogeneity assumption is imposed, which in effect also transforms the dependent variable. The histogram of the transformed variable with the respective Gaussian bell curve and the Shapiro-Wilkins W-statistics indicated that the transformed variable is normally distributed.

In order to be able to choose the most appropriate specification for the stochastic output distance function, the estimations were computed for pooled model and for four different panel data models. The panel data specifications were one-factor and two-factor fixed effects models estimated using ordinary least squares (OLS) and one-factor and two-factor random effects models estimated using generalized least squares (GLS) (see Appendix). The five models were in turn estimated using different specifications for the dependent variable (constant or variable returns to scale) and for autocorrelation specifications (with and without Cochrane-Orcutt AR1 transformation). The estimation results were not sensitive to whether the AR1 correction or whether DF^{vrs} or DF^{crs} was used as a dependent variable.

The test statistics (LM and Hausman tests) used to discriminate between the different models indicate that the two-factor random effects (TWREM) model described the data generating mechanism best (see *Table 5.3*).[9] The error component in the model is specified as $\nu_{pt} = \varepsilon_{pt} + \mu_p + \varpi_t$, where ε_{pt} is the error term, assumed to have mean zero and constant variance; μ_p is the random effect characterizing the *p*th observation and constant through time, i.e., it represents the collection of factors not in the regression that are specific to a plant; ϖ_t is the random effect, which varies through time but is constant across plants. The test statistics indicate that the TWREM model specification provides a good fit. Moreover, the explanatory variables do not suffer from the possible endogeneity bias (see Appendix). The results in *Table 5.2* also show that the differences between the various model specifications are small. This result is in accordance with the fact that when P (the number of plants) is fixed and T (number of time periods) tends to infinity, the different estimation methods converge (Hsiao 1986). Consequently, whether the individual -and time-specific effects are treated as fixed or random does not change the results significantly.[10]

Parameter estimates of the TWREM model were used to compute the value of the output distance function and the shadow prices for each plant. The estimated fitted values for the output distance function varied between 0.55 and 1.32. Thus, as would be expected, the results violate the theoretical upper limit value ($D_0 \leq 1$) due to a number of positive residuals. Grosskopf et al. (1992) and Lovell et al. (1990) use the COLS correction to yield theoretically consistent values. In addition, they interpret the residuals (or deviations from the maximum residual) as a measure of inefficiency. However, since the measured dependent variable (and thus the stochastic model) in the present study already includes inefficiency, it is appropriate to interpret the fitted values which are above 1 as reflecting purely random errors and not including any inefficiency component.

The t-statistics show that all the parameters, except the ones related to BOD, are significant at the 5 percent level (*Table 5.2*). The shadow prices of undesirable outputs were computed using equation (2.6) and the assumption that the absolute shadow price of pulp (r_Q) is equal to its observed market price (r_Q^o). Because the shadow prices for BOD were insignificant, they are not presented.

Table 5.4 shows the means of the plant-specific absolute shadow prices for FLOW and the value of output (per ton of pulp). In interpreting the shadow prices, one should bear in mind the following facts. First, it is important to note that the figures give the shadow prices at the actual levels of emissions, not at the level of some pollution constraint. This is indeed a desirable feature, since in practice the plants have rarely exactly met the constraints.[11] Usually the plants have been below the constraint, probably due to the fact that they start to adjust to new regulation when they receive information about a forthcoming regulation, even though it may actually come into force after 1-5 years. Thus, the shadow prices do not measure directly the effect of the regulation, but rather the marginal rate of transformation between the good and bad outputs.

Table 5.2 Parameter Estimates (152 obs.)

Parameter	1. OLS Pooled	2. OLS Fixed Effect	3. GLS REM	4. OLS TwoFactor Fix	5. GLS Two Factor REM
α_0	12.22	12.85	12.93	14.54	13.51
	(31.03)	(20.14)	(20.57)	(18.50)	(20.64)
β_K	-0.22	-0.22	-0.22	-0.24	-0.23
	(7.05)	(5.81)	(6.40)	(5.74)	(6.43)
β_L	- 0.08	- 0.16	- 0.13	- 0.19	- 0.14
	(2.99)	(3.45)	(3.24)	(3.58)	(3.44)
β_M	- 0.64	- 0.64	- 0.64	- 0.68	- 0.66
	(17.26)	(13.42)	(14.91)	(11.84)	(14.70)
γ_Q	0.77	0.72	0.74	0.68	0.72
	(25.17)	(15.79)	(18.71)	(11.06)	(15.78)
γ_{BOD}	0.07	-0.007	-0.04	0.06	0.07
	(2.43)	(0.14)	(1.00)	(1.03)	(1.49)
γ_{FL}	0.10	0.25	0.18	0.24	0.17
	(3.80)	(3.87)	(3.59)	(3.38)	(3.30)
γ_{BODFL}	-0.08	-0.13	-0.09	-0.13	-0.10
	(1.42)	(1.77)	(1.33)	(1.72)	(1.44)
γ_{BODQ}	-0.03	0.02	-0.02	-0.001	-0.02
	(0.71)	(0.36)	(0.32)	(0.01)	(0.39)
γ_{BOD2}	0.02	0.003	0.008	0.01	0.01
	(1.04)	(0.14)	(0.35)	(0.48)	(0.47)
γ_{FLQ}	-0.30	-0.33	-0.32	-0.46	-0.39
	(4.04)	(4.23)	(4.19)	(5.22)	(4.77)
γ_{FL2}	0.33	0.33	0.32	0.37	0.34
	(7.45)	(6.43)	(6.54)	(6.70)	(6.78)
γ_{Q2}	0.10	0.10	0.11	0.21	0.17
	(2.26)	(2.16)	(2.48)	(3.74)	(3.35)
adj R2	0.98	0.98	0.98	0.98	0.98

* t-ratios in parentheses

Table 5.3 Test Statistics for Model Selection

Model	Log-Likelihood	Sum of Squares	R^2
(i) Constant term only	- 157.41	0.007	0.00
(ii) Group dummies only	-95.71	0.003	0.55
(iii) X,U variables + α_0	146.21	0.01	0.98
(iv) Full Fixed Effects	157.05	0.01	0.98
(v) Full Two Factor Fixed	174.49	0.89	0.99

LM-test: Model (v) vs. Model (iii) = 1.11 (2df)
Hausman-test: Model (v) vs. Model (iv) = 10.67 (12df)

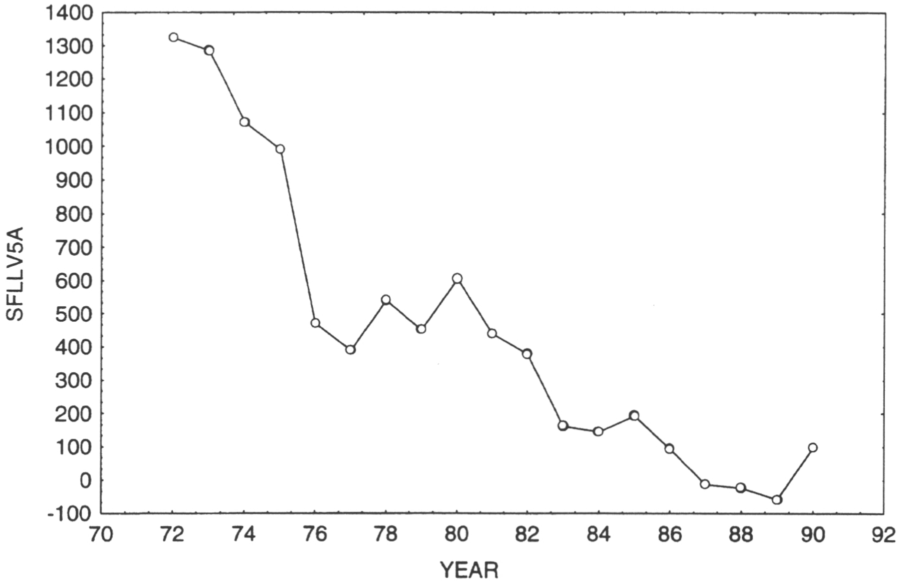
Table 5.4 Shadow Prices (mean values, FIM/ton)

Variable	Overall	Plant 1	Plant 2	Plant 3	Plant 4	Plant 5	Plant 6	Plant 7	Plant 8
r_{GVP}	2728	2739.5	2877.7	3078.6	2549.5	2655.2	2886.2	2691.8	2355.1
	(373)	(371)	(406)	(315)	(217)	(252)	(226)	(446)	(205.8)
r_{FLOW}	449.2	987.6	452.8	989.7	786.1	-681.2	216.7	1009	-167.2
	(1002)	(951.3)	(616.6)	(763.3)	(343.9)	(1616)	(415.4)	(634.2)	(613.4)

Standard deviations in parentheses

The most significant feature of the above shadow price estimates is that for bulk of the observations, the FLOW shadow prices (r_{FLOW}) are positive (more about this below). The overall mean value of the shadow price for FLOW is 449.2 FIM/ton (the mean value of the FLOW shadow price is around 16 % of the mean value of pulp output (r_{GVP})). The differences in the mean values of the FLOW shadow prices across the plants are rather large (they vary between -681.2 and 1009). When the shadow prices are examined over time, one can see a clear downward pattern (*Figure 5.2*). The average shadow price in 1972 is 1323.6 and it goes down to 98.5 in 1990 (the shadow price is negative in 1987, 1988 and 1989). The decreasing gains probably reflect the general change in the production process towards a closed-loop water system and the fact that the gains associated with this process gett smaller at the margin (see below).[12]

Figure 5.2 Flow shadow prices (means across plants)



The fact that the variations in the mean values of shadow prices between the plants are rather large would seem to indicate that the regulatory system has not been cost efficient. However, it should be noted that this result holds only if the environmental benefits from marginal reduction in effluents are equal across the plants. Since the plants for the most part pollute different waters, it is difficult to make judgements as to the benefits of reduced emissions and hence about the cost effectiveness of the regulatory system.

It is interesting to see how the above results would change, if the forntier approach had been used, rather than the nonfrontier one. Thus, a model in which the observations were restricted to be on the frontier (i.e., the dependent variable was set equal to 1) was estimated.

According to the Hausmann test statistics, the two-way random effects specification was also the preferred specification for the frontier model. The explanatory power of the model is slightly smaller and the t-values lower than for the non-frontier model used in the present study. However, the most interesting result is that the derived shadow prices for FLOW are different for the two models; the prices are lower in the frontier model (the mean value is FIM 365.8) and they are all positive. Consequently, the results are clearly sensitive to whether one assumes that the plants are operating on the frontier or, alternatively, that they are distributed according to the actual distance measures.

The positive shadow prices for FLOW appear to be in conflict with the theoretical results from environmental economics literature (see, e.g., Oates et al. 1993). Conventional economic wisdom would suggest that regulation diverts the plant's resources away from the production of pulp, which in turn results in increasing costs, declining production, reduced employment and decreased profits for the firms. In contrast, the above results appear to indicate that environmental regulations have either enhanced the revenues of the plants or have had no effects at all. What are the possible reasons for the positive shadow prices?

There are basically two ways for the plants to reduce waste water: by modifying the effluents using external treatment measures (e.g., building an aerated pond or activated sludge plant), and/or by reducing the generation of effluents by implementing internal production process changes. External treatment measures clearly add an extra costs to the plants and do not generate any additional benefits in terms of higher efficiency or productivity. Although the plants have received investment credits from the state for pollution abatement equipment, they have been a small proprotion of the total abatement costs. Consequently, the positive shadow prices for FLOW are probably due to the fact that the internal process changes in the production of pulp have simultaneously decreased the amount of waste water effluents and improved productivity. The long-run strategy of the pulp plants in developing the production process has been to aim at closed-loop water systems, which simultaneously improve the efficiency in the control of production systems and reduce water pollution. As a result of this strategy, the production of one ton of pulp in 1990 required on average ten times less water than in the 1950s.[13]

It is important to note that the changes in the production process may have occurred either independently of the regulation or as a result of the regulation. In the latter case, productivity improvements have emerged as serendipitous by-products of waste water reduction. Thus, there may be potentially significant *"learning by doing"* effects associated with environmental regulations. Further, it is likely that gains generated by learning spill over across the plants. This kind of argument has recently been put forward by Porter (1990). Indeed, in their analysis of the "Porter hypothesis", Oates et al. (1993) argue that the most likely reason that regulations might generate positive effects on firms profits is that there has existed inefficiency and unrealized opportunities for cost-savings and product enhancement before the regulation, and that the regulations induce the realization of these opportunities.

If the above argument is correct, one may ask why the firms utilized the positive spillover effects only after the regulations forced them to do so. One possible answer to this question is related to the information cost argument. There are potentially many ways by which production efficiency could be enhanced, and firms are uncertain about which of the possibilities will result in benefits that exceed the costs (research and development costs). However, analyzing a wide range of possible new ways to increase productivity is costly and the firms may not utilize these possibilities until regulations force them to do so. Oates et al. 1993 list a number of other possible reasons why firms do not realize the potential gains in the absence of regulation.

The results obtained in the present study are in accordance with those obtained in Hetemäki (1994). Using the same theoretical framework and data set, but applying the conventional stochastic distance function estimation procedure rather than the two-stage method, Hetemäki (1994) obtained shadow prices for undesirable outputs which were very similar to the present ones. In contrast, the present results are rather different from those obtained by Färe et al. (1993) using a deterministic output distance function and cross-section data for pulp mills operating in Michigan and Wisconsin in 1976. The results obtained by Färe et al. showed that the absolute shadow prices of different measures of water pollution were large and negative and that there are large variations in shadow prices across the different mills. For example, the mean of the plant -specific absolute shadow price of BOD indicated that reducing one ton of BOD emissions diverts enough resources to have produced over two tons of paper, and the standard deviation of the plant -specific shadow prices was higher than their mean. If these differences in the results of the two studies could be regarded as reflecting purely country -specific differences, it would indicate that the Finnish and US pulp mills are using very different production technologies or/and that they are operating in a strikingly different environment. However, it would seem more plausible to consider that a significant amount of the difference is a result of the differences in data base (homogeneous panel data vs. heterogeneous cross-section), estimation method (stochastic vs. deterministic) and functional form (restricted translog vs. translog), and in the fact that Färe et al. restrict the prices of undesirable outputs to be negative or zero.

6. Conclusions

The distinguishing feature of the present approach is the combination of nonparametric linear programming with the stochastic distance function and with plant level panel data. This framework provides important improvements on previous deterministic and stochastic distance function studies. First, the estimation procedure does not require the plants to operate on the frontier of the production technology but is instead based on the actual distance measures. Secondly, no a priori restrictions on the values of the undesirable output parameters are set. Furthermore, plant level panel data from a relatively homogeneous industry sector provide more informative and robust results than the commonly used aggregate cross-section or time series data.

The results of the present study show that water pollution reduction by Finnish pulp plants has, for most of the plants and for most of the period studied, enhanced the revenues of the plants. For some plants and some years, the effects have been either slightly negative or close to zero. This result should not, however, be interpreted unambiguously to show that environmental regulations cause plants' profits to increase. Rather, the result indicates that control of emissions is part of the control of the whole pulping process. Recycling waste water and closing the water circulation simultaneously reduces material waste, improves the production process and reduces water pollution. In other words, environmental regulation is not the only factor which has caused these plants to reduce water pollution, but also the fact that pollution control measures and improvements in the production process appear to be strongly positively correlated.

How legitimately the above result can be generalized to other production processes is an empirical question. Nevertheless, the result indicates that one should not a priori rule out the possibility that pollution control may be positively correlated with an increase in frims' revenues. Indeed, it may be, that the conventional wisdom, which suggest that pollution control results in additional burdens to firms and thus reduces revenues and productivity, is perhaps not as good an approximation of reality in the future as it may have been in the past. For example, there is a clear tendency in the markets of North America and Europe to value "green" products more highly over more polluting ones, and the environmental reputation of the firms is becoming an increasingly important marketing strategy. If consumers are willing to pay a premium for green products and for the good environmental reputation of the suppliers, and if this premium more than offsets the abatement costs, the net effect is that the profits of the firms are increased simultaneously with the reduction of pollution, ceteris paribus. The essence of this issue is also captured in the words of the director of environmental affairs: "It has always been our belief that good environmental practices are good business practices. It is not altruism which motivates us, but pragmatism" (Gordon Wallace, director of environmental affairs, James River Fine Papers).[14] Naturally, this type of change in markets may also have important implications for the theoretical and empirical models used to analyze the effects of pollution control costs.

Notes

1. According to Shephard (1974, p. 205), "... , for the future where unwanted outputs of technology are not likely to be freely disposable, it is inadvisable to enforce free disposal of inputs and outputs. Since the production function is a technological statement, all outputs, whether economic goods are wanted or not, should be spanned by the output vector y."

2. For example, there are reasons related to positive spillover effects of pollution control which suggest that shadow prices of undesirable outputs may also be positive. Gray and Shadbegian (1993) note that, "In some cases regulations may increase productivity. In response to pressures to reduce waste water discharges, some plants adopted "closed-loop" production processes and discovered after doing so that the cost savings from recycling raw materials reduced total costs. .....New equipment, installed to reduce pollution, may also be more productive than the old equipment it replaces". Similar evidence can also be found in the classical study by Kneese and Bower (1968) (see also Oates et.al. 1993).

3. Outputs are called *weakly disposable*, if $y \in P(x)$ and $\theta \in [0,1]$ and $\theta y \in P(x)$; and *strongly disposable*, if $v \leq y \in P(x)$ then $v \in P(x)$. Basically, weak disposability implies that radial (equiproportional) reduction in outputs is possible, but reduction of some outputs may not be feasible without altering inputs. On the other hand, strong disposability implies that outputs can be disposed without any resource use. For a more detailed description of the concepts, see Färe et al. (1994).

4. Although an inefficient observation has no supporting hyperplane (since there is no supporting frontier), the shadow prices for inefficient observation still make sense in the present context. The way to think this is that, although the observation is inefficient, we calculate the shadow price that would have been, if the observation had been efficient. The inefficient observation is scaled *proportionally* up to the frontier. Then the derivatives, which give the shadow prices, yield the same mutual relation as the derivatives evaluated at the optimal (efficient) point. Thus, as long as the scaling (inefficiency) is proportional, it does not effect the relations between the shadow prices.

5. A number of other specifications were also tested. For the nonrestricted translog specification, the bulk of the parameters were insignificant and very sensitive even to minor changes in the model specification and to changes in the estimation method. The specification used in the present study was retained because it appeared to be the least affected by multicollinearity of those specifications which included second order and cross product terms of the output variables.

6. Pittman (1981) notes that, "The different pulping processes represented in the sample -and the associated differences in end product characteristics- are troublesome for purposes of estimation; technological homogeneity is desirable in production function analysis. In particular, Bower et al. (1971), Bower (1975), and Krutilla and Smith (1979) have noted the importance of end product characteristics for pollution control requirements in this industry (footnote 7, p.4)."

7. *Biological Oxygen Demand* is a parameter which describes the amount of oxygen the micro-organism uses within a specific time period (usually 7 days demand) for biologically dissolving waste water and it gives an indication of the amount of easily dissolvable matters in waste water. *Suspended Solids* are the fibers and particles in the waste water. For a more detailed description of the waste water effluents (see, James 1985).

8. When the distance measures were computed by taking each year separately (i.e., 19 sets with 8 observations each) almost all the distance measures were equal to 1, due to the small number of observations. On the other hand, when the *sequential approach* (see

Lovell 1993, pp. 47-49), which allows progressive technical change, was used to compute the distance measures, the results did not differ greatly from those presented in *Table 5.1*. However, we chose to use the pooled data in the nonparametric analysis, and to allow the period -and plant -specific effects to enter the stochastic model. This has the advantage of providing parameter values for time and plant specific effects with the respective significance levels for the obtained parameters.

9. The model test statistics in *Table 5.2* are the Breusch-Pagan's Lagrange (LM) multiplier statistic for testing the TWREM model against Model (iii), i.e., the model without plant effects. Large values for the LM statistic argue in favor of the fixed-effects models against the regression without plant -specific effects. The Hausman test (m) is a chi-square test of whether the GLS estimator is an appropriate alternative to the fixed-effects estimator. If the individual effects are not correlated with the other regressors, OLS and GLS are consistent but OLS is inefficient. In the opposite case, OLS is consistent, but GLS is not. It should be noted that when P is fixed and T tends to infinity, β_{FE} and β_{GLS} become identical. Consequently, the Hausman test,

$$m = \hat{q}'\hat{V}ar(\hat{q})^{-1}\hat{q}, \qquad \text{where } \hat{q} = \hat{\beta}_{FE} - \hat{\beta}_{GLS}, \hat{V}ar(\hat{q}) = Var(\hat{\beta}_{FE}) - Var(\hat{\beta}_{GLS})$$

approaches zero. Thus, the Hausman test cannot be used to test for misspecification. However, in this case the fixed-effects and random-effects models are indistinguishable for all practical purposes (Hsiao 1986). It should also be stressed that when P is fixed and $T \to \infty$, the maximum likelihood estimates of μ, β, and σ^2_μ converge to the fixed-effects estimator and are consistent, but the MLE of the variance of the plant effects is inconsistent. This is due to the fact that when P is fixed, there is not enough variation in the plant effects no matter how large T is (Hsiao 1986).

10. In the estimation, the homogeneity restriction turned out to be problematic. The homogeneity restriction consists of four different parts since there are three outputs, i.e., one part imposes $\sum_{m=1}^{M} \gamma_m = 1$, and three parts set $\sum_{m'=1}^{M} \gamma_{mm'} = 0$. The first restriction is maintained hypothesis, since it also transforms the dependent variable and imposes the weak disposability of outputs. On the other hand, the latter three restrictions were clearly rejected on the basis of the F-test. However, the restriction that $\gamma_{BODFL} + \gamma_{BODQ} + \gamma_{BOD2} + \gamma_{FLQ} + \gamma_{FL2} + \gamma_{Q2} = 0$ could be accepted even at the 1% level according to the F-test. Consequently, the results used to compute the shadow prices are based on the model, in which the latter restriction is set.

11. Pittman (1981, p.9) states that "given the imprecise nature of production and pollution control technologies, it seems unlikely that one would ever find such a constraint exactly met. This would not seem to imply, however, that such plants are not behaving under constraint, nor that they would attach zero value to a relaxation of the constraint." Similarly, Brännlund and Löfgren (1994, pp.1-2) note that in the case of the Swedish pulp industry, "Most plants ...,have an average emission level far below the allowed level. We do not think, however, that this means that almost all kinds of regulations are ineffective. Instead, we believe that the main reason for this inequality is that the firm cannot control its emissions exactly..., the plant's waste load is subject to stochastic fluctuations. ...,if the firm wants to avoid a violation of the regulation,..., it will on average pollute less than the allowed level.

12. Since the BOD parameters are not significant, the BOD shadow prices are not necessarily robust. However, using the TWRE model parameters, the derived mean BOD shadow price is FIM 171, with minimum of -470.2 and maximum of 772.3. Similar to the FLOW shadow price, the avearge BOD shadow price gets smaller over time.

13. The fact that the internal process measures which reduce waste water also have positive effects on the revenues of the plants is recorded also in the data collected from the plants by the water authority. The data concerning the net costs of the internal process changes undertaken due to pollution regulation show that for some plants and some years the net effects are positive. Since the plants have incentives to overestimate the true costs, the above result may be even more general than the data appears to indicate.

14. *Pulp and Paper International*, May 1994, p. 69.

References

Aigner, D. and S.F. Chu (1968). On Estimating the Industry Production Function, *American Economic Review*; 826-839.

Banker, R.D. and W.W. Cooper (1994). Validation and Generalization of DEA and its Uses, to be published in *Trabajos de Investigacion Operativa.*

Barbera, A.J. and McConnell, V.D. (1990). The Impact of Environmental Regulations on Industry Productivity: Direct and Indirect Effects, *Journal of Environmental Economics and Management* 18, 50-65.

Brännlund, R. and K.-G., Löfgren (1993). Emission Standards and Stochastic Waste Load, unpublished mimeo, Department of Economics, University of Umeå.

Conrad, K. and Morrison, C.J. (1989): The Impact of Pollution Abatement Investment on Productivity Change: An Empirical Comparison of the U.S., Germany, and Canada. *Southern Economic Journal*, January: 684-698.

Fried, H.O., C.A.K. Lovell & S.S. Schmidt. (eds.) (1993). *The Measurement of Productive Efficiency*, Oxford University Press.

Färe, R. (1988): *Fundamentals of Production Theory*, Lecture Notes in Economics and Mathematical Systems, vol. 311, Berlin: Springer Verlag.

Färe, R., S. Grosskopf, C.A.K. Lovell and C. Pasurka (1989). Multilateral Productivity Comparisons when Some Outputs are Undesirable: A Nonparametric Approach. *Review of Economics and Statistics*, February; 90-98.

Färe, R., S. Grosskopf, C.A.K. Lovell and S. Yaisawarng (1993). Derivation of Virtual Prices for Undesirable Outputs: A Distance Function Approach, *Review of Economics and Statistics*, May; 374-380.

Färe, R., S. Grosskopf, C.A.K. Lovell (1994). *Production Frontiers*, Cambridge University Press.

Gollop, F.M. and M.I. Roberts (1983). Environmental Regulations and Productivity Growth: The Case of Fossil-fueled Electric Generation, *Journal of Political Economy*: 654-674.

Gray, W.B. & S.J. Shadbegian (1993). Environmental Regulation and Manufacturing Productivity at the Plant Level, National Bureau of Economic Research, Working Paper No. 4321.

Grosskopf, S., K. Hayes, L. Taylor and W. Weber (1992). Budget Constrained Frontier Measures of Fiscal Equality and Efficiency in Schooling, unpublished mimeograph.

Grosskopf, S. and K. Hayes (1993). Local Public Sector Bureaucrats and Their Input Choices, *Journal of Urban Economics*, 33, 151-166.

Gulledge, T.R. and C.A.K. Lovell (1992 eds.). Special Issue: International Applications of Productivity and Efficiency Analysis, *Journal of Productivity Analysis,* vol.3, no. 1/2.

Hetemäki, L. (1994). The Impact of Pollution Control on a Firm's Production Technology and Efficiency: A Stochastic Distance Function Approach, Proceedings of Environmental Economics Conference held at Ulvön (eds. Brännlund, R., B. Kriström & K.-G. Löfgren), Sweden, June 10-13, 1993. Swedish University of Agricultural Sciences, Department of Forest Economics, Report no. 106.

Hsiao, C. (1986). *Analysis of Panel Data*, Cambridge University Press.

James, D. (1985). Environmental Economics, Industrial Process Models, and Regional-Residuals Management Models, in *Handbook of Natural Resource and Energy Economics*, vol.I, ed. Kneese, A.V. and J.L. Sweeney.

Jorgenson, D. and P.J. Wilcoxen (1990). Environmental Regulation and US. Economic Growth, *RAND Journal of Economics*, vol.21,no.2; 314-340.

Kneese, A.V. and B.T. Bower (1968). *Managing Water Quality: Economics, Technology, Institutions.* Resources for the Future, John Hopkins Press.
Lovell, C.A.K., S. Richardson, P. Travers and L.L. Wood (1990). Resources and Functionings: A New View of Inequality in Australia, Department of Economics, University of North Carolina, Working Paper Series, no. 90-8.
Lovell, C.A.K. (1993). Production Frontiers and Productive Efficiency, in Fried, H.O., C.A.K. Lovell & S.S. Schmidt. (eds.)
Myers, J.G. and L. Nakamura (1980). Energy and Pollution Effects on Productivity: A Putty-Clay Approach, in *New Developments in Productivity Measurement and Analysis, Kendrick* (ed.), University of Chicago Press.
Oates, W., K. Palmer and P. R. Portney (1993). Environmental Regulation and International Competitiveness: Thinking About the Porter Hypothesis, Resources for Future, Discussion Paper 94-02.
Pittman, R.W. (1981). Issues in Pollution Interplant Cost Differences and Economies of Scale, *Land Economics*, 1-17.
Porter, M. (1990). *The Competitive Advantage of Nations,* New York: The Free Press.
Shephard, R.W. (1970). *Theory of Cost and Production Functions*, Princeton University Press.
Shephard, R.W. (1974). Comment on Diewert, E. "Applications of Duality Theory", in Intriligator, M.D. and Kendrick, D.A. (eds.), *Frontiers of Quantitative Economics vol. II*, North-Holland.
Simar, L. Estimating Efficiencies from Frontier Models with Panel Data: A Comparison of Parametric, Non-Parametric and Semi-Parametric Methods with Bootstrapping, *Journal of Productivity Analysis*, no. 3; 171-203.

APPENDIX I: Data and variables

The data was collected from two different sources; Industrial Statistics collected by the Central Statistical Office of Finland (Teollisuuden Yleislomake ja Energialomake) and water pollution statistics collected by the National Board of Waters and Environment (Vesiensuojelun A ja B lomake). Both type of statistics are based on annual questionnaires sent to all plants. Because the questionnaires sent by two different authorities for collection of different information were not necessarily coherent, some of the figures were checked and corrected by directly contacting the plants involved. Since the data is confidential and its collection requires permission from each of the firms, code numbers are used for the plants in order to make them unidentifiable.

OUTPUT. The pulp and paper output series include information on the value and quantity (tons) for sulphate pulp. The implicit price index for output is derived by dividing the value of output by the quantity of output.

WATER POLLUTION. The principal aim of water pollution monitoring is to assess the waste water ingredients, their quantity and toxicity, to control the compliance with permit conditions and to assess treatment efficiency and factors affecting efficiency. The monitoring is done according to a program approved by the supervising authority, i.e., the local water authority. The monitoring is carried out by both the official water laboratory (of which there are around 20 in the whole country) and by the pulp plants themselves. The local water authority gives limiting values at the plant level for the discharge in terms of total load per time unit or specific load per ton of product. In general, the limits must be attained as mean values for 1,3, or 6 months, depending on the size and type of plant. The water pollution statistics concerning the quantities of effluents of the pulp plants is considered to be of good quality by the National Board of Waters and Environment (see Enckell-Sarkola et al. [8]).

The water pollution statistics used in the present study consists of information on the flow of waste water (m3/a), biological oxygen demand (BOD7) (t/a) and suspended solids (t/a).

LABOR. The information on labour input consists of data on both production and non -production (white collar) workers total numbers, hours worked, and wages and social security costs. Social security costs are not available for 1972-73 and so were estimated using the procedure outlined in Mäisti (Tulonjako paperiteollisuudessa vuosina 1955-1977, Työväen Taloudellinen Tutkimuslaitos, tutkimusselosteita 8:1979). The quantity of labour input is measured as the hours worked. Since there may be differences between production and non-production workers that is not reflected in the number of hours worked, the Divisia (or discrete time Törnqvist) index was used to compute an aggregate index of hours worked.

CAPITAL. As is well known, the construction of data series for capital stock and price (user cost) of capital poses fundamental difficulties. For a clear exposition of these issues, see, e.g., Berndt, E. (1991). *The Practice of Econometrics*, Addison-Wesley.

The capital series consists of information on annual (1974-1990) purchases of capital goods (a), basic improvement costs (b), sales (c), and rented capital goods (d) of 6 different classes of capital assets (1. residential buildings, 2. non-residential buildings, 3. machinery, instruments and tools, 4. transportation equipment, 5. land and water structures, 6. other material investments). The gross investment series (e) is constructed as $e = a + b + d - c$. From 1972-73 there are also data on the fire insurance values of the different classes of capital assets. The 6 different classes of capital assets were first aggregated into two groups, namely, buildings = 1 + 2 + 5 and equipment and machinery = 3 + 4 + 6.

The replacement cost values of fixed capital assets were calculated from the perpetual inventory formula, $K_t = (1-\delta_t)K_{t-1} + I_{t-1}$,where K_t is the capital stock at the beginning of time t, δ is the constant rate of depreciation, and I_{t-1} is investment in period t-1. In order to obtain the starting (or benchmark) values for the capital stock we assumed equality of fire insurance cost and historic cost valuations of the capital stock in the first year of the data (1972) (Nickell et. al. 1992) have noted that "the choice of an accurate benchmark may be largely irrelevant" in a fixed effects panel data model).

In order to calculate the constant exponential rate of depreciation, the procedure given in Kuh, E. and R. Schmalense (*An Introduction to Applied Macroeconomics*, North-Holland, 1973) was used. According to this procedure the depreciation rate is calculated using the equation, $(1-\delta)L = X$,
where L is the average service lives of capital assets and X is the value of capital assets as a percentage of their initial values at the end of their average service lives. It was assumed that, of the initial value of equipment and machinery, 10 percent is left after 32 years in the paper industry and after 25 years in the pulp industry. The corresponding figure for buildings was assumed to be 65 years for both industries. These figures for the service lives of capital assets are higher than those reported in the National Accounts. The figures used here are based on Simula (Tuottavuus Suomen metsäteollisuudessa. Licentiate thesis, University of Helsinki, Department of Social Economics of Forestry, 1979) rather than the more simple calculations of the Central Statistical Office. However, the figures should still be regarded as crude approximations. The above assumptions imply values of δ of 8.8% for equipment and machinery and 3.5% for buildings (For comparison, e.g., Nickell et. al. 1992 use the values 8.19% and 2.5%, respectively for the UK manufacturing industry). Finally, the replacement cost valuation of *total* fixed capital assets is calculated as the simple sum of the fixed capital assets of plant and machinery and buildings.

MATERIALS. The data on intermediate materials consists of information on the value of materials (including electricity). This is a "catch-all" variable which includes data on various inputs with different units. The important problem in constructing a materials input variable is that, as usual, there is no data on the quantity or price of this "input". Since it is essential to determine how much change in value can be considered a result of changes in quantity over time and across plants rather than in the prices, relevant price indexes or deflators must be found. In the present study the production price index for manufacture of paper and paper products was used as a deflator.

APPENDIX II: Panel data models

In the present study, it has been assumed that the slope coefficients are constant across plants, and differences between plants are captured through intercept or disturbance terms. The five different model specifications, outlined below, were used to estimate the econometric model (see Hsiao1986).

The pooled cross-section time-series model (eq. 1), in which the constant terms are the same across plants (i.e., only one common constant term, α_0), was estimated using Ordinary Least Squares (OLS).

(1) $D_{Opt} = \alpha_0 + \beta' X_{pt} + \varepsilon_{pt}$,

In equation 2, it is assumed that differences across plants can be captured in differences in the constant term. Thus, α_p is an unknown parameter to be estimated using dummy variables (d_p) indicating the *p*th plant. This model is the Fixed Effects (FE) model and it is estimated using OLS.

(2) $D_{Opt} = \alpha_p + \beta' X_{pt} + \varepsilon_{pt}$

The FE model can be extended to include a time-specific effect as well, i.e. it is assumed that similar effects "hit" every plant in each particular period. The extended model is known as the Two Factor Fixed Effects (TFFE) model. This model is obtained by the addition of an additional T-1 dummy variable to equation (2) and is also estimated by OLS.

(3) $D_{Opt} = \alpha_0 + \alpha_p + \lambda_t + \beta' X_{pt} + \varepsilon_{pt}$

Often the fixed effects models are viewed as applying to the cross-sectional units in the study, not to additional ones outside the sample. Since the present sample of sulphate pulp plants is not an exhaustive sample (there are between 12 to 17 plants altogether, depending on the definition and time period one is using), it may be more appropriate to model plant-specific constant terms as randomly distributed across cross-sectional plants. The Random Effects (RE) model is shown in (4), where the term μ_p is the random disturbance characterizing the *p*th observation and is constant through time. In other words, the assumption that the μ_p are random variables implies that the P plants can be regarded as a random sample from some larger population, and it also implies that the μ_p and X_p are uncorrelated. The efficient estimator for model (4) is Generalised Least Squares (GLS). First, the variance components are estimated using the residuals from the OLS regression. In the second stage, feasible GLS estimates are computed using the estimated variances.

(4) $D_{Opt} = \alpha_0 + \beta' X_{pt} + \mu_p + \varepsilon_{pt}$

It is possible that the random effects also vary through time and that the correct model is the Two Factor Random Effects (TFREM) model (eq. (5)). The term ν_t captures the time varying random effects. The TFREM model is estimated consistently and efficiently by feasible GLS.

(5) $D_{Opt} = \alpha_0 + \beta' X_{pt} + \mu_p + \nu_t + \varepsilon_{pt}$

Appendix III: Model diagnostics

Autocorrelation. In order to check the residual autocorrelation in the TWREM-model, the autocorrelation functions, Box-Pierce (BP) Q-statistic and Ljung-Box (LB) Q-statistic were computed. The Q statistics test the hypothesis that all of the autocorrelations are zero (three lags in the present case). The Ljung-Box statistic has the same large sample distribution as BP statistic, but it has been found to have better small sample properties. The residual analysis was carried out for each plant separately. The results showed that the null hypothesis of no autocorrelation could be accepted.

Normality. The Shapiro-Wilk W-test of normality was carried for each model specification. The W-statistic for TWREM model was 0.97 and the null hypothesis of normality could be accepted at the 1.5% significance level. The results from testing for normality for each plant separately showed that normality could be accepted for all of them, except for two plants, at the 5 % significance level.

Orthogonality. The results in *Table 5.2* maintain the assumption of orthogonality between the error term and the regressors. If orthogonality fails to hold, the estimates of all the coefficients could be affected. In order to check whether the potential correlation of the right-hand side variables with the error term is great enough for the results of the model 5 to be biased, the *Pearson correlation coefficients (r)* and the significance level (or probability level, *p*) of the respective correlations were computed. In addition the relationship between the residuals and the right-hand side variables was examined by looking the slopes of the regression lines from regressing the residuals on each of the exogeneous variable in turn. Variables that are not correlated are not necessarily independent, except for the case of the joint normal distribution, in which lack of correlation does imply independence. However, Monte Carlo studies suggest that meeting the normality assumption is not absolutely crucial if the sample size is not very small (>100) and when the departure from normality is not large.

Pearson correlation coefficients

	β_K	β_L	β_M	γ_Q	γ_{BOD}	γ_{FL}	γ_{BODFL}	γ_{BODQ}	γ_{BOD2}	γ_{FLQ}	γ_{FL2}	γ_{Q2}
RESID *r*[x,e]	0.19	0.20	0.12	-0.04	-0.01	-0.12	0.15	0.11	0.11	0.03	0.10	-0.02
p-level	0.02*	0.02*	0.13	0.59	0.88	0.14	0.06	0.18	0.17	0.76	0.24	0.79

The above correlation coefficients show that the correlation is rather low, the maximum value being 0.20, and the correlation being significant at the 5% level only for the labor and capital inputs. Even in the latter two cases, the coefficient of determination, r^2, is only 0.04, indicating that the possible endogeneity bias is very small. Moreover, for the two-factor fixed effects model, none of the correlation coefficients were significant.

ISBN 951-40-1375-1
ISSN 0358-4283

Helsinki 1994
Yliopistopaino
Pikapaino