Estimation of shadow prices for undesirable outputs: an application to UK dairy farms*

Paper to be presented at the American Agricultural Economics Association Annual Meetings, Salt Lake City, Utah, August 2-5, 1998.

David Hadley
Division of Agriculture and Horticulture,
School of Biological Science,
University of Nottingham,
Sutton Bonington Campus,
Loughborough, Leics, LE12 5RD, UK.

May 1998

* This research was completed while the author was a PhD student at the University of Reading supported by a Ministry of Agriculture, Fisheries and Food studentship.

Abstract

Analysis of agricultural production generally ignores the undesirable outputs (such as nitrate, or pesticide contamination of water) that are produced alongside desirable, marketable outputs. This paper presents the results of research which integrates a simple physical model of nitrate leaching from dairy production into a multiple input/multiple output representation of the production technology: the output distance function. Estimation of the output distance function as a frontier allows for the derivation of shadow prices of the undesirable output which can be interpreted as the marginal abatement costs that each producer faces.

The study uses an unbalanced panel dataset derived from annual survey returns from 330 individual UK dairy farms which span the period 1982 to 1992 and totals to 2130 observations. The shadow price for the undesirable output evaluated at the mean of the data is estimated to be -£29.34.

ESTIMATION OF SHADOW PRICES FOR UNDESIRABLE OUTPUTS: AN APPLICATION TO UK DAIRY FARMS

INTRODUCTION

Generally agricultural production analysis is concerned with describing the relationships that characterise the transformation of inputs - land, labour, purchased materials, etc. - into marketable outputs - wheat, milk, meat and so on. Such outputs are desirable in the sense that they are demanded by consumers and yield utility in consumption. But, the process of modern, industrial agricultural production does not solely result in the provision of desirable products for the market. Within the process are also created outputs which society deems undesirable because they yield disutility in consumption. Among such outputs could be included the contamination of ground and surface waters due to runoff and leaching of nitrogenous fertilisers and pesticides, the emission of greenhouse gases to the atmosphere and the removal of hedgerows and subsequent 'prairiefication' of agricultural landscapes. All these 'bad' outputs impose costs, either in direct monetary terms - when, for example, water supply companies are faced with the cost of removing contaminants from supplies - or in a more indirect, but equally valid, loss of welfare - such as that suffered by consumers of rural landscapes whose 'enjoyment' is marred as a result of unaesthetic agricultural activity.

By their very nature these 'bad' outputs are nonmarketable and their prices are not observed. This paper uses the methodology developed in Färe, Grosskopf, Lovell and Yaisawarng (1993) (hereafter FGLY) to calculate shadow prices for a variable that represents a simple measure of nitrate leaching to groundwater. Whilst these shadow prices do not directly represent the costs to society that nitrate contamination imposes they do represent the cost to agricultural producers that reductions in emissions of nitrates would entail. In this sense they can be interpreted as a measure of the abatement costs that each farmer faces.

The work presented here differs from FGLY and from other related studies (Coggins and Swinton, 1996, and Piot-Lepetit and Vermersch, 1998) in that econometric methods are employed to estimate the parameters of an output distance function and differs from the one previous econometric approach to this analysis (Hetemäki, 1997) by using a systems estimation procedure.

THEORETICAL BACKGROUND

The output distance function

Consider the case of a farm using a vector of *N* inputs, denoted by $x = (x_1,...,x_n)$, $x \in \Re_+^N$ to produce a vector of *M* outputs, denoted by $y = (y_1,...,y_m)$, $y \in \Re_+^M$. The production technology of the farm can be defined by the output set, *P(x)*, which represents the set of all output vectors which can be produced with the input vector *x*. The output distance can be formally defined as (Färe and Primont, 1995):

$$D_o(x,y) = \inf_{\theta} \left\{ \theta > 0 : \left(\frac{y}{\theta} \right) \in P(x) \right\} \text{ for all } x \in \Re_+^N \quad (1)$$

The output distance function thus measures the reciprocal of the largest possible proportional increase in the output vector (the extent of which is indicated by the value of θ) given the input vector and given that the expanded output vector must still be a member of the output set.[1] Values of $D_O(x,y)$ for observations which are equal to one lie on the frontier and are technically efficient. Conversely, values which are less than one imply that production could be increased by y/θ given x and hence are technically inefficient.

The output distance function inherits a number of properties from the parent technology (see Färe, 1988; and Färe and Primont, 1995); including homogeneity of degree +1 in outputs and convexity in outputs. A further property that is fundamental to the analysis undertaken in this paper is its ability to enable the modelling of outputs as being weakly disposable. Conventionally, strong disposability of outputs is assumed; which is to say that application of an additional unit of input always yields some non-negative amount of additional output and outputs can be freely disposed of (McFadden, 1978). Weak disposability, however, allows for the fact that disposal of some outputs may be costly in terms of opportunity cost of foregone outputs of other commodities

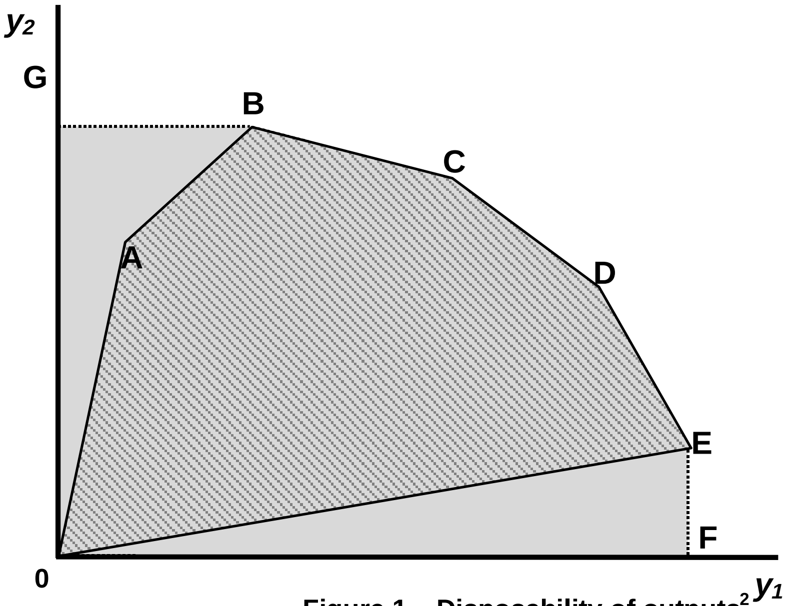

Figure 1 – Disposability of outputs[2]

These concepts are best explained diagrammatically. Figure 1 shows output sets (constructed here using a piecewise linear frontier) which feature strong ($P^S(x)$) and weak disposability ($P^W(x)$) of outputs. $P^S(x)$ is bounded by the line segments 0G, GB, BC, CD, DE, EF and F0, whilst $P^W(x)$ is bounded by 0A, AB, BC, CD, DE and E0 - and is obviously a subset of $P^S(x)$. If y_1 is an undesirable output and y_2 a desirable output and all outputs are strongly disposable then from point B it would be possible to reduce the level of y_1 along the dotted line segment towards point G with no cost in terms

---

[1] The value of the output distance function is thus the inverse of the Farrell output measure of technical efficiency (Färe, Grosskopf and Lovell, 1985).

[2] Adapted from Färe, Grosskopf and Pasurka (1986).

of reduced output of y_2. Alternatively, operating within $P^W(x)$ where disposal of the undesirable output is not free then the segment GB is no longer feasible; reduction of y_1 requires an accompanying reduction in y_2 along the segment AB or 0A.

To summarise, weak disposability of outputs, for this example, implies that reduction in the undesirable output has an associated opportunity cost of reduced desirable output (or an associated cost in terms of increased input use).

Derivation of shadow prices

Denoting output prices by $r = (r_1,....,r_M)$ the revenue function can be defined as $R(x,r) = \sup_y \{ry : y \in P(x)\}$. Shephard (1970) proves that the revenue function and the output distance function are dual and hence each can be expressed in terms of the other:

$$R(x,r) = \sup_y \{ry : D_0(x,y) \leq 1\}, \tag{2}$$

$$D_o(x,y) = \sup_r \{ry : R(x,r) \leq 1\}, \tag{3}$$

FGLY show that if these functions are differentiable then the maximal solution vector to the Lagrangian problem for (2) will satisfy:

$$r = R(x,r) \bullet \nabla_y D_o(x,y), \tag{4}$$

where ∇ is the gradient vector. Denoting the vector of revenue maximising output prices obtained from (2) as $r^*(x,y)$, FGLY (1993) apply Shephard's dual lemma to the revenue maximisation problem to yield:

$$\nabla_y D_o(x,y) = r^*(x,y). \tag{5}$$

Substituting (5) into (4) gives:

$$r = R(x,r) r^*(x,y). \tag{6}$$

The vector $r^*(x,y)$ can be interpreted as revenue deflated shadow prices for outputs (FGLY, 1993). Calculation of absolute shadow prices requires knowledge of maximum revenue ($R(x,r)$), this in turn requires computation of the shadow prices which are to be derived and hence some assumption regarding the values of shadow prices or the value of maximum revenue must be made to make the calculation operational. FGLY employ the following assumption in their analysis: *one observed output price equals its absolute shadow price* and note that it may also be appropriate to assume that maximum revenue is equal to observed revenue. Using the former assumption and denoting the observed market price of 'good' output one by r_{g1}^o and its revenue deflated shadow price by r_{g1}^* allows calculation of maximum revenue as follows:

$$R(x,r) = \frac{r_{g1}^o}{r_{g1}^*(x,y)} = \frac{r_{g1}^o}{\left(\dfrac{\partial D_o(x,y)}{\partial y_{g1}}\right)}. \tag{7}$$

Absolute shadow prices for 'bad' outputs (identified here with a 'b' subscript) with no observable market prices can then be derived as:

$$r_{b1} = R(x,r) \bullet \left(\frac{\partial D_o(x,y)}{\partial y_{b1}} \right) = r_{g1}^{o} \bullet \left(\frac{\left(\frac{\partial D_o(x,y)}{\partial y_{b1}} \right)}{\left(\frac{\partial D_o(x,y)}{\partial y_{g1}} \right)} \right) \quad (8)$$

DATA

The central core of the farm production data is drawn from the Farm Business Survey (FBS) for England and Wales (Ministry of Agriculture, Fisheries and Food, Economic (Farm Business) Division, Welsh Office, (1994)) covering the production years from 1982 to 1992. The FBS is an annual survey of more than 2,800 farms that are selected from a random sample of census data that is stratified according to region, economic size of farm and type of farming. A sub sample of 330 dairy farms (defined here as those farms where 60% or more of revenue is derived from milk or milk products) observed for varying numbers of years (the mean duration being 6.45 years) are extracted from this dataset to form an unbalanced panel. Analysis is restricted to 330 farms by the availability of soil and annual precipitation information. Dairy farms are chosen because milk is a relatively homogenous product (hence aggregation problems are minimised) which generally accounts for a large proportion of farm revenue and, because they are the most widely represented farm types in the FBS (both in terms of geographical distribution and in numbers of surveyed farms). In total, 2130 observations form the final panel dataset.

The 'bad' output

The 'bad' output (Y_{b1}) is represented within the model as an index - the Groundwater Vulnerability Index (GWVIN) as described in Kellogg, Maizel and Goss (1992). This index is constructed from two components:

- a nitrogen leaching index (LI), the value of which is dependent upon the soil characteristics of each farm and annual rainfall patterns, and
- an estimate of excess nitrogen (N_e) for each farm, involving the calculation of individual farm nitrogen budgets, i.e. the balance between nitrogen inputs and outputs.

The LI has been developed by Williams and Kissel (1991) as an indicator of the potential for nitrate to be leached and its calculation requires data on the soil hydrologic group and monthly rainfall at each farm and for each year in which that farm appears in the sample. It is itself the product of two other indices; a percolation index - an estimate of average annual percolation through and below a crop's root zone which is a function of annual precipitation and the water transmission properties of the soil - and a seasonal index - which attempts to account for increases in percolation occurring outside of crop growing seasons.[3]

[3] Rain falling during the autumn and winter months is more likely to percolate below the root zone than that which falls in the growing season when rainfall is, to some extent, taken up by crops and evapotranspiration is higher, (Williams and Kissel, 1991).

Individual farm nitrogen budgets are constructed using estimates of flows of nitrogen across the farm boundary (using a simplified version of the accounting procedure outlined in Meisinger and Randall, 1991). Annual quantities of nitrogen contained in purchased inputs (fertilisers and animal feeds) are summed along with an estimate of nitrogen deposited from the atmosphere, subtracted from this is the annual quantity of nitrogen contained in the harvested part of crops and in livestock (and livestock products) sold off farm. Any product of this calculation that is greater than zero is termed excess nitrogen. Formally, this calculation is made as follows:

$$N_e = (N_{fert} + N_{feed} + N_{atmos}) - (N_{crop} + N_{lstk}) \quad (9)$$

where: N_e = excess nitrogen, N_{fert} = quantity of nitrogen contained in fertiliser, N_{feed} = quantity of nitrogen contained in animal feed, N_{atmos} = quantity of nitrogen deposited from the atmosphere[4], N_{crop} = quantity of nitrogen contained in harvested crops, and N_{lstk} = quantity of nitrogen contained in livestock and livestock products sold. All elements are measured in tonnes.

The GWVIN for each farm in each year is then computed as:

$$GWVIN_{it} = LI_{it} \times N_{e\ it}, \quad (10)$$

where: *i* indexes individual farms, and *t* indexes years.

In essence the GWVIN represents an estimate of excess nitrogen which is weighted by is potential to be leached. Whilst its virtue is its relative ease of computation it cannot be directly interpreted as an estimate of nitrate emissions at any site and it is used here simply to illustrate the analysis employed, work continues to integrate a more ‘realistic’ model of the leaching process into the production technology model.

Conventional outputs

Two conventional, ‘good’, outputs are also included in the analysis; milk and milk products (Y_{g1}) and an other outputs variable (Y_{g2}). The former is defined in terms of quantity of milk and milk products (in hectolitres) sold off farm, plus quantity of farmhouse consumption and benefits in kind, less quantity used on farm. The latter is a catch-all variable defined as revenue derived from all other livestock and crop enterprises (livestock and crop products used on farm – e.g. for feed or seed – are deducted to reflect their deduction from the appropriate input variables).

Conventional inputs

Six input aggregates are considered. Livestock inputs (X_1) represents the flow of incurred livestock expenses and includes; livestock variable costs (feed, veterinary expenses, etc.), expenditure on purchased animals and interest on the animal capital stock (included in order to account for the

[4] Substantial amounts of nitrogen are deposited on all land surfaces in the UK through wet and dry deposition from the atmosphere and any calculation of nitrogen inputs to agricultural systems must take account of this (Goulding, 1990). Here it is assumed that annual deposition has been constant over the period 1982-92 and estimates for nitrogen deposition for the three geographical zones described in Scholefield, Lockyer, Whitehead and Tyson (1991) are used.

opportunity cost of capital tied up in livestock).[5] Labour (X_2) is a simple sum of hours worked annually by all classes of labour (family, hired and casual). Miscellaneous costs, electricity, heating fuel, etc. are aggregated to form a general costs variable (X_3). Crop inputs (X_4) is made up of annual expenditure on seeds and young plants, fertilisers, crop protection and other miscellaneous variable crop costs. The flow of services emanating from capital stock items such as machinery, buildings and land improvements (X_5) is measured by summation over these elements of maintenance and running costs, depreciation charges and interest on the capital stock (calculated according to the logic employed for the interest on the animal capital stock item). Finally, agricultural land (X_6) is simply defined as total utilised agricultural area (in hectares) for each farm.

All output and input variables defined in value terms are deflated using the appropriate annual price indices published by the UK Ministry of Agriculture, Fisheries and Food. Further detail and discussion regarding the construction of these aggregate variables can be found in Hadley, 1997. Summary statistics for these variables and others of interest are detailed in Table 1 below.

Variable	Description	Unit	Mean	Standard Deviation	Minimum	Maximum
Y_{g1}	Milk output	hl	5264.68	3525.15	396.97	32091.00
Y_{g2}	Other output	£	28412.06	23839.02	467.29	199701.14
Y_{b1}	GWVIN (bad output)	index	113.46	135.66	0.01	789.47
X_1	Livestock inputs	£	41222.80	29498.12	2834.27	256283.61
X_2	Labour	hours	6397.87	3204.14	1900.00	29000.00
X_3	General costs	£	16982.33	13868.12	952.68	105586.13
X_4	Crop costs	£	9850.16	8358.29	99.39	71738.53
X_5	Capital	£	22341.06	28720.22	759.44	1152144.65
X_6	Land	ha	78.99	59.33	9.15	742.16
N_e	Excess nitrogen	t	22.36	15.14	1.17	103.81
S_{yg1}	Revenue share Y_{g1}	%	74.82	7.86	59.58	97.88
S_{x1}	Cost share X_1	%	34.02	8.12	7.76	64.51
S_{x2}	Cost share X_2	%	19.39	6.46	4.52	53.20
S_{x3}	Cost share X_3	%	13.60	5.27	2.80	37.10
S_{x4}	Cost share X_4	%	7.67	3.09	0.33	22.72
S_{x5}	Cost share X_5	%	18.15	5.66	3.16	81.82
S_{x6}	Cost share X_6	%	7.18	2.45	1.21	17.42
	Milk price	£ hl^{-1}	15.51	1.42	12.47	30.41
Farms	No. of farms		330			
N	No. of observations		2130			

Table 1 - Summary statistics for sample data 1982-1992

5 Interest on the animal capital stock is calculated as the average annual value of animal stocks multiplied by the real rate of interest on 91 day UK Treasury Bills (chosen to represent a real after-tax return on a safe asset and following Paul and Abey, 1984).

ESTIMATION

The econometric estimation of distance functions is fundamentally hampered by the fact that the function has no observable dependent variable, a problem that is compounded by issues surrounding the endogeneity of the independent variables. Two differing strategies have been used to combat these problems in previous work:

- by setting the value of the dependent variable equal to one for all observations, using instruments for the endogenous independent variables and correcting resulting estimates for the biases the estimation procedure necessarily includes (see, for example, Grosskopf and Hayes, 1993),
- by exploiting the fact that the output distance function is homogenous of degree +1 in outputs and transforming the dependent and endogenous right-hand side output variables by division by some arbitrarily chosen output, and subsequently assuming the transformed right-hand side variables are exogenous, output-mix, variables (see Lovell, Richardson, Travers and Wood (1994) and Coelli and Perelman (1996)).[6]

Specifying the output distance function as translog (since we have no *a priori* expectations regarding its functional form) and applying the latter of these two strategies and using the 'bad' output as a numeraire the function to be estimated can be written:

$$\begin{aligned}\ln\left[\frac{D(x,y,t)_{it}}{y_{b1it}}\right] &= \alpha_0 + \sum_{m=1}^{2}\alpha_m \ln\left(\frac{y_{gmit}}{y_{b1it}}\right) + \frac{1}{2}\sum_{m=1}^{2}\sum_{n=1}^{2}\alpha_{mn}\ln\left(\frac{y_{gmit}}{y_{b1it}}\right)\ln\left(\frac{y_{gnit}}{y_{b1it}}\right)\\ &+ \sum_{k=1}^{6}\beta_k \ln x_{kit} + \frac{1}{2}\sum_{j=1}^{6}\sum_{k=1}^{6}\beta_{jk}\ln x_{jit}\ln x_{kit}\\ &+ \sum_{k=1}^{6}\sum_{m=1}^{2}\delta_{km}\ln x_{kit}\ln\left(\frac{y_{gmit}}{y_{b1it}}\right) + \alpha_t t + \frac{1}{2}\alpha_{tt}t^2 + \sum_{m=1}^{2}\alpha_{mt}\ln\left(\frac{y_{gmit}}{y_{b1it}}\right)t\\ &+ \sum_{k=1}^{6}\beta_{kt}\ln x_{kit}t,\end{aligned} \tag{11}$$

where: *i* indexes farms, *t* indexes time, y_{b1} is the bad output, y_{gm} are good outputs, x_k are inputs, and ln denotes natural logarithms (first and second-order time and output/input/time interaction terms are also included to account for non-neutral technical change over the period). Exploiting the properties of logarithms then:

$$\ln\left[\frac{D(x,y,t)_{it}}{y_{b1it}}\right] = \ln[D(x,y,t)_{it}] - \ln(y_{b1it}). \tag{12}$$

Subtracting $\ln[D(x,y,t)_{it}]$ from both sides of (11) yields $-\ln(y_{b1\,it})$ as the dependent variable and adds the negative unobservable value of the distance function to the right hand side of the equation. Interpreting this as a one sided error term which is assumed to account for firm-specific effects and adding a symmetric error term (which accounts for firm- and time-specific statistical noise) to (11)

---

[6] Coelli and Perelman (1996) argue that outputs appear on the right-hand side of the transformed function as ratios and since the output distance function is defined for radial expansion of all outputs, given observed input levels, output ratios are held constant for each observation by definition. Hence these output mix ratios can be considered as exogenous variables in the model.

produces a function which is amenable to estimation within the standard stochastic error component frontier function model introduced by Aigner, Lovell and Schmidt (1977). Furthermore, the presence of panel data allows for the value of the one-sided error to be predicted using the panel data frontier estimation methods outlined in Schmidt and Sickles (1984).

Single equation estimates of the modified version of (11) are likely to severely suffer from the effects of multicollinearity. Färe and Primont (1995) summarise the duality relationships between the output and input distance functions and the revenue and cost functions respectively and also note the equivalence between the output distance function and the reciprocal of the input distance function (assuming the production technology exhibits constant returns to scale). Exploiting these relationships allows for the parameters of the output distance function, in logarithmic form, to be estimated more efficiently as a system along with derived revenue share and negative cost share equations.[7] In the context of panel data this system can be estimated using the iterative seemingly unrelated regression (ISUR) technique where the variables entering the output distance function are transformed according to the usual panel data within and generalised least squares (GLS) estimation methods, whilst the variables appearing in the derived share equations remain untransformed.[8]

A number of issues must be considered regarding the appropriate choice of panel data estimator. GLS estimation may be preferable over the within estimator when *N* (number of firms) is large and *T* (the length of the time dimension of the panel) is small (as is the case here) and it has an important advantage in its ability to incorporate time-invariant regressors (which are wiped out by the within transformation) (Schmidt and Sickles, 1984). Mundlak (1978) also argues that the fixed firm effects derived from within estimation can be considered random, but that inference is conditional upon the sample. The GLS estimator makes specific distributional assumptions about the random firm effects (in terms of their variance) and allows unconditional inference to be made. However, GLS is biased if the regressors are correlated with the firm-specific effects. A Hausman test (Hausman, 1978) of the within against the GLS estimates of the parameters of the transformed version of (11) produced a χ^2_{55} statistic of 36.58 (the critical value at 5% being 73.31) leading to an acceptance of the null hypothesis that firm effects and regressors are uncorrelated and that GLS estimates are consistent. It is the results from the GLS procedure that are therefore presented in the following section.

---

[7] Grosskopf, Hayes, Taylor and Weber (1997) and Bosco (1996) describe applications that, respectively, derive budget and cost share equations from an indirect output distance function and an input distance function.
[8] See Hsiao (1986) and Baltagi (1995) for expositions of panel data estimation techniques. The procedure adopted here is an adaptation of Kumbhakar (1997) – the short time dimension used prevents the application of the heteroscedasticity correction the original employs.

RESULTS

The GLS estimates of the parameters of the output distance function are detailed in Table 2 below.[9] Goodness-of-fit as measured by R^2 for each equation of the system are: 0.986 for the output distance function, 0.931 for the dairy revenue share equation and 0.956, 0.836, 0.941, 0.888, and 0.945 for the five negative cost share equations (X_1 to X_5 respectively). The McElroy (McElroy, 1977) system wide measure of goodness-of-fit is valued at 0.963.

Parameter	Variable	Estimate	Standard Error	*t*-statistic
α_0	Constant	0.62129	0.28238	2.200
α_1	$\ln Y_{g1}$	1.02376	0.00657	155.862
α_2	$\ln Y_{g2}$	-0.14669	0.01996	-7.348
α_3	$\ln Y_{b1}$	0.12293	0.01886	6.519
β_1	$\ln X_1$	-0.05533	0.00641	-8.634
β_2	$\ln X_2$	-0.19345	0.01017	-19.026
β_3	$\ln X_3$	-0.15654	0.00490	-31.971
β_4	$\ln X_4$	-0.14886	0.00447	-33.301
β_5	$\ln X_5$	-0.19606	0.00546	-35.930
β_6	$\ln X_6$	-0.32223	0.12669	-2.543
α_{11}	$(\ln Y_{g1})^2$	0.16676	0.00100	165.998
α_{12}	$\ln Y_{g1} \times \ln Y_{g2}$	-0.16707	0.00099	-169.219
α_{13}	$\ln Y_{g1} \times \ln Y_{b1}$	0.00031	0.00025	1.278
α_{22}	$(\ln Y_{g2})^2$	0.17718	0.00157	112.496
α_{23}	$\ln Y_{g2} \times \ln Y_{b1}$	-0.01011	0.00115	-8.768
α_{33}	$\ln(Y_{b1})^2$	-0.00979	0.00112	-8.730
β_{11}	$\ln(X_1)^2$	-0.21439	0.00088	-243.246
β_{12}	$\ln X_1 \times \ln X_2$	0.06209	0.00097	64.119
β_{13}	$\ln X_1 \times \ln X_3$	0.04877	0.00052	94.347
β_{14}	$\ln X_1 \times \ln X_4$	0.02504	0.00047	52.943
β_{15}	$\ln X_1 \times \ln X_5$	0.06277	0.00056	111.770
β_{16}	$\ln X_1 \times \ln X_6$	0.01802	0.00099	18.238
β_{22}	$\ln(X_2)^2$	-0.14316	0.00187	-76.501
β_{23}	$\ln X_2 \times \ln X_3$	0.02263	0.00075	30.281
β_{24}	$\ln X_2 \times \ln X_4$	0.01139	0.00066	17.209
β_{25}	$\ln X_2 \times \ln X_5$	0.03642	0.00082	44.520
β_{26}	$\ln X_2 \times \ln X_6$	0.00166	0.00147	1.130
β_{33}	$\ln(X_3)^2$	-0.11411	0.00057	-201.066
β_{34}	$\ln X_3 \times \ln X_4$	0.01034	0.00038	27.159
β_{35}	$\ln X_3 \times \ln X_5$	0.02617	0.00046	56.780
β_{36}	$\ln X_3 \times \ln X_6$	0.00756	0.00076	9.989
β_{44}	$\ln(X_4)^2$	-0.05692	0.00050	-114.431
β_{45}	$\ln X_4 \times \ln X_5$	0.01259	0.00043	29.348
β_{46}	$\ln X_4 \times \ln X_6$	-0.00184	0.00068	-2.705
β_{55}	$\ln(X_5)^2$	-0.13906	0.00071	-196.550
β_{56}	$\ln X_5 \times \ln X_6$	0.00586	0.00082	7.167
β_{66}	$\ln(X_6)^2$	0.00073	0.03041	0.024

continued

Table 2 - Generalised least squares parameter estimates[10]

[9] Parameter estimates are obtained from GLS estimation of the homogeneity transformed output distance function given in (11) and *m*-1 and *k*-1 revenue and negative cost share equations.
[10] Shaded entries refer to parameter estimates recovered using the homogeneity conditions.

Table 2 - continued

Parameter	Variable	Estimate	Standard Error	*t*-statistic
δ_{11}	$\ln X_1 \times \ln Y_{g1}$	0.00058	0.00065	0.886
δ_{21}	$\ln X_2 \times \ln Y_{g1}$	-0.00059	0.00097	-0.611
δ_{31}	$\ln X_3 \times \ln Y_{g1}$	0.00229	0.00052	4.439
δ_{41}	$\ln X_4 \times \ln Y_{g1}$	0.00107	0.00047	2.286
δ_{51}	$\ln X_5 \times \ln Y_{g1}$	-0.00155	0.00057	-2.739
δ_{61}	$\ln X_6 \times \ln Y_{g1}$	-0.00528	0.00109	-4.864
δ_{12}	$\ln X_1 \times \ln Y_{g2}$	0.00146	0.00065	2.242
δ_{22}	$\ln X_2 \times \ln Y_{g2}$	-0.00271	0.00097	-2.793
δ_{32}	$\ln X_3 \times \ln Y_{g2}$	-0.00113	0.00051	-2.204
δ_{42}	$\ln X_4 \times \ln Y_{g2}$	-0.00105	0.00046	-2.273
δ_{52}	$\ln X_5 \times \ln Y_{g2}$	0.00253	0.00056	4.502
δ_{62}	$\ln X_6 \times \ln Y_{g2}$	0.00544	0.00406	1.338
δ_{13}	$\ln X_1 \times \ln Y_{b1}$	-0.00204	0.00021	-9.636
δ_{23}	$\ln X_2 \times \ln Y_{b1}$	0.00330	0.00032	10.291
δ_{33}	$\ln X_3 \times \ln Y_{b1}$	-0.00116	0.00016	-7.290
δ_{43}	$\ln X_4 \times \ln Y_{b1}$	-0.00002	0.00013	-0.165
δ_{53}	$\ln X_5 \times \ln Y_{b1}$	-0.00097	0.00016	-5.957
δ_{63}	$\ln X_6 \times \ln Y_{b1}$	-0.00016	0.00394	-0.040
α_t	Time	0.02241	0.00760	2.949
α_{tt}	Time^2	-0.00089	0.00057	-1.561
α_{1t}	$\text{Time} \times \ln Y_{g1}$	-0.00024	0.00017	-1.387
α_{2t}	$\text{Time} \times \ln Y_{g2}$	0.00083	0.00065	1.275
α_{3t}	$\text{Time} \times \ln Y_{b1}$	-0.00059	0.00063	-0.940
β_{it}	$\text{TIme} \times \ln X_1$	0.00365	0.00014	25.669
β_{2t}	$\text{Time} \times \ln X_2$	-0.00875	0.00022	-39.882
β_{3t}	$\text{Time} \times \ln X_3$	0.00177	0.00011	16.144
β_{4t}	$\text{Time} \times \ln X_4$	0.00076	0.00009	8.727
β_{5t}	$\text{TIme} \times \ln X_5$	0.00157	0.00011	14.126
β_{6t}	$\text{Time} \times \ln X_6$	-0.00609	0.00164	-3.717

Individual firm effects (α_i) are recovered from the residuals of the output distance function (e_{it}) as follows (Schmidt and Sickles, 1984):

$$\alpha_i = \frac{1}{T}\sum_i e_{it}. \qquad (13)$$

The maximum value of the vector of these firm effects is interpreted as indicating the best practice farm (i.e. which is 100% technically efficient and for which the output distance function value is equal to one) and the value of the output distance function for all other firms is measured relative to this according to:

$$D(x,y,t)_i = \exp(\alpha_i - \alpha), \qquad (14)$$

where, $\alpha = \max(\alpha_i)$. By definition values are bounded by 0 and 1.

Predicted values of the output distance function range from a minimum of 0.893 to 1.000 with a mean of 0.936 indicating that on average outputs could be increased by approximately 7% if all farms operated according to the best practice in the sample.[11]

Absolute shadow prices for the 'bad' output are computed at each observation using (8) and assuming that the observed price of Y_{g1} is equal to it absolute shadow price. The 'bad' shadow price evaluated at the mean of the data has a value of -£29.34 indicating the loss in revenue the average farm would incur for a reduction of one unit of the 'bad' output. Over all 2130 observations 1989 shadow prices are negative and of the 141 positive values 48% of them occur for levels of Y_{b1} which are less than one. Very large positive shadow prices occur where Y_{b1} is very small (i.e. below 0.3) and these large values skew the mean shadow price evaluated over all observations to be a positive value (£3640.82). Given this very skewed distribution a more appropriate summary statistic is the median value which evaluates to -£33.77 (inter-quartile range = 89.88).

The relationship between shadow prices and levels of the 'bad' output was explored using shadow price values computed at the mean for each farm in a regression of the log of farm mean 'bad' shadow price (SY^*_{b1}) (multiplied by –1) on a constant and the farm mean log of Y_{b1}. Estimated parameters are as follows (*t*-statistics are shown in parentheses below each parameter);

$$\ln(-SY^*_{b1i}) = \underset{(34.57)}{6.075} - \underset{(-12.90)}{0.5453} * \ln(Y_{b1i}). \qquad (15)$$

$$\bar{R}^2 = 0.34$$

This relationship is illustrated graphically in Figure 2:

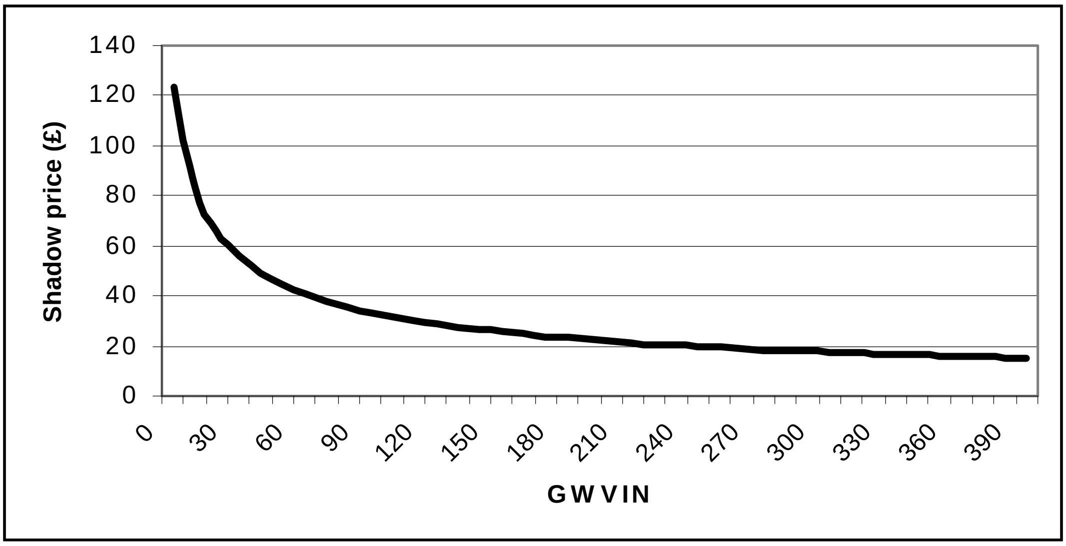

Figure 2 – Abatement costs

[11] Note that in order to promote simplicity output distance function values are assumed constant over the period covered by the sample. Also note that the inclusion of the 'bad' output in the function specification effectively credits farms for high levels of production of that output (*cf.*, for example, Färe, Grosskopf, Lovell and Pasurka (1989) for non-parametric programming methods which evaluate efficiency where 'good' and 'bad' outputs are treated asymmetrically).

This estimated curve can be interpreted as representing the abatement costs, in terms of diminished revenue, that farms would face for reductions in levels of Y_{b1}. Its shape shows that the marginal cost of abatement for producers with high levels of undesirable emissions is significantly lower than for those producers whose level of nitrate emission is low.

CONCLUSION

This paper applies FGLY's (1993) methodology to evaluate shadow prices for nitrate emissions to groundwater from a sample of UK dairy farms. An output distance function incorporating a variable representing nitrate emissions is estimated with unbalanced panel data on 330 farms observed over the period 1982 to 1992. The parameters of this function are evaluated using a GLS panel data estimator within a system incorporating the function itself and derived revenue share and negative cost share equations. From the estimated parameters shadow prices for the undesirable output are computed and an abatement cost curve is estimated.

The shape of the abatement cost curve suggests that policy designed to curb emissions of nitrate to groundwater will be most effective if it is possible to target producers whose emissions are relatively high and whose abatement costs are significantly smaller relative to farms where emissions are low. Policy instruments which do not differentiate between producers according to their level of emissions will punish producers with low emission rates disproportionately.

REFERENCES

Aigner, D., Lovell, C.A.K. and Schmidt, P. (1977) Formulation and estimation of stochastic frontier production models, *Journal of Econometrics*, **6**, 21-37.

Coelli, T.J. and Perelman, S. (1996) Efficiency measurement, multiple-output technologies and distance functions: with application to European railways, CREPP Working Paper No:96/05, Universite de Liege.

Coggins, J.S. and Swinton, J.R. (1996) The price of pollution: A dual approach to valuing SO_2 allowances, *Journal of Environmental Economics and Management*, **30**, 58-72.

Färe, R. (1988) *Fundamentals of Production Theory*, Springer-Verlag, Berlin.

Färe, R., Grosskopf, S. and Lovell, C.A.K. (1985) *The Measurement of Efficiency of Production*, Kluwer-Nijhoff, Boston.

Färe, R., Grosskopf, S., Lovell, C.A.K. and Yaisawarng, S. (1993) Derivation of shadow prices for undesirable outputs: a distance function approach, *The Review of Economics and Statistics*, **75,** 374-380.

Färe, R., Grosskopf, S. and Pasurka, C. (1986) Effects on relative efficiency in electric power generation due to environmental controls, *Resources and Energy*, **8**, 167-184.

Färe, R. and Primont, D. (1995) *Multi-output Production and Duality: Theory and Applications*, Kluwer Academic Publishers, Boston.

Goulding, K.W.T. (1990) Nitrogen deposition to land from the atmosphere, *Soil Use and Management*, **6**, 61-63.

Grosskopf, S. and Hayes, K. (1993) Local public sector bureaucrats and their input choices, *Journal of Urban Economics*, **33**, 151-166.

Hadley, D. (1997) *Estimating shadow prices of undesirable outputs: an application to UK dairy farms*, Unpublished PhD thesis, Department of Agricultural and Food Economics, University of Reading.

Hausman, J.A. (1978) Specification tests in econometrics, *Econometrica*, **46**, 1251-1271.

Hetemäki, L. (1996) *Essays on the Impact of Pollution Control on a Firm: A distance function approach*, Finnish Forest Research Institute, Research Papers 609, Helsinki.

Kellogg, R.L., Maizel, M.S. and Goss, D.W. (1992) *Agricultural Chemical Use and Ground Water Quality: Where are the potential problem areas?*, United States Department of Agriculture, Washington, D.C.

Lovell, C.A.K., Richardson, S., Travers, P. and Wood, L. (1994) Resources and functionings - a new view of inequality in Australia, *Models and Measurement of Welfare and Inequality*, (Ed. Wolfgang Eichorn), Springer-Verlag, Berlin.

McElroy, M.B. (1977) Goodness of fit for seemingly unrelated regressions: Glahn's $R^2_{y,x}$, and Hooper's $\bar{r}^2$, *Journal of Econometrics*, **6**, 381-387.

McFadden, D. (1978) Duality of production, cost and profit functions, Part I in, *Production Economics: A dual approach to theory and applications*, (Eds. Melvyn Fuss and Daniel McFadden), North-Holland, Amsterdam.

Meisinger, J.J. and Randall, G.W. (1991) Estimating Nitrogen Budgets for Soil-Crop Systems, Chapter 5 in *Managing Nitrogen for Groundwater Quality and Farm Profitability*, (Eds. R.F. Follett, D.R. Keeney and R.M. Cruse), Soil Science Association of America, Madison, Wisconsin.

Ministry of Agriculture, Fisheries and Food Economic (Farm Business) Division, Welsh Office (1994) *Farm Business Survey 1982-1993* [computer files & associated documentation], Colchester: ESRC Data Archive.

Mundlak, Y. (1978) On the pooling of time series and cross section data, *Econometrica*, **46**, 69-85.

Paul, P. and Abey, R. (1984) *Conceptual issues in measuring productivity growth*, paper presented at the 28th Annual Conference of the Australian Agricultural Economics Society, Sydney, February.

Piot-Lepetit, I. and Vermersch, D. (1998) Pricing inorganic nitrogen under the weak disposability assumption: an application to the French pig sector, *Journal of Agricultural Economics*, **49**, 85-99.

Schmidt, P. and Sickles, R.C. (1984) Production frontiers and panel data, *Journal of Business and Economic Statistics*, **2**, 367-374.

Scholefield, D., Lockyer, D.R., Whitehead, D.C. and Tyson, K.C. (1991) A model to predict transformations and losses of nitrogen in UK pastures grazed by beef cattle, *Plant and Soil*, **132**, 165-177.

Shephard, R.W. (1970) *Theory of Cost and Production Functions*, Princeton University Press, New Jersey.

Williams, J.R., and Kissel, D.E. (1991) Chapter 4, Water percolation: An indicator of nitrogen-leaching potential in, *Managing Nitrogen for Groundwater Quality and Farm Profitability*, (Eds. R.F. Follett, D.R. Keeney and R.M. Cruse), Soil Science Association of America, Madison, Wisconsin.