A Bivariate Approach to the Determination of Effective Pollution by Farm-households

F. Perali[1], P. Polinori[2], C. Salvioni[3] and M. Veronesi[1]

[1] Department of Economics, University of Verona, Italy
[2] Department of Economics, University of Perugia, Italy
[3] Department MQTE, University of Chieti-Pescara, Italy

Contribution appeared in Arfini, F. (Ed.) (2005) "Modelling Agricultural Policies: State of the Art and New Challenges", proceedings of the 89th EAAE Seminar, pp. 384 - 399

February 2-5, 2005

Parma, Italy

A Bivariate Approach to the Determination of Effective Pollution by Farm-households

F. Perali[*], P. Polinori[**], C. Salvioni[***] and M. Veronesi[*]

Abstract

Following a bivariate probit approach and using the 1996 survey conducted by the Italian Institute for the Studies of Agricultural Markets (ISMEA) in Italy, this study shows that the effective pollution of Italian farm-households depends on both the actual level of pollution, calculated using the OECD "Environmental Indicators for Agriculture" (1997, 1999, 2001a, b), as well as the level of environmental concern, measured using the survey information about the adoption of environmentally sensitivity technology and production techniques. Omission of the level of environmental concern from the estimation of effective pollution probability yields evidence of policy unfairness in the allocations of subsidies between farms. This evidence, however, disappears when the level of actual pollution and environmental concern are estimated jointly in a bivariate framework. In addition, the study shows that geographic location of the farm, the common market organizations and the farm types play a significant roles in the determination of the level of effective pollution.

Keywords: Actual Pollution, Aggregation, Bivariate Probit, Environmental Concern, Effective Pollution.

JEL Classification: C35, Q20, Q28.

Acknowledgments

The author thanks the Italian Institute for the Studies of Agricultural Markets (ISMEA) for providing the data for this study.

[*] Department of Economics, University of Verona, Italy.
[**] Department of Economics, University of Perugia, Italy.
[***] Department MQTE, University of Chieti-Pescara, Italy.

Introduction

To insure the environmental sustainability of the development of rural areas the Community has recently included cross-compliance requirements in the CAP. Farmers who do not respect the obligations arising from about 40 legislative acts, which apply directly at farm level[1], will see their payments reduced or cancelled.

The objective of the paper is to analyse the problem of defining a procedure for monitoring and evaluating whether farmers are meeting cross-compliance requirements. We present a methodology that accounts for the effective pollution of farms which is the actual level of pollution corrected by the level of environmental sensitivity of the farm. Different farms may have comparable levels of actual pollution, but may differ in their concerns about the environment in terms of the adoption of pollution abatement and control technologies.

The proposed methodology begins by calculating several indexes measuring the impact on the environment and on the human health of the production process at the farm level, and then aggregates them in an aggregate index of actual pollution. An aggregate index of environmental concern is computed, taking into account the adoption of low impact or abatement technologies. The effective level of pollution is measured by an aggregate index that is the product of the previous two indexes. This latter index is the one we propose to use to monitor and evaluate the farm environmental impact.

In order to make the proposed methodology operative even in those cases in which the data on the farm production process and behaviour, needed to calculate the simple and aggregate indexes, are not available, the paper also shows how the farm propensity to pollute can be inferred by making use of commonly available information on the farm. In fact, as it is shown, the effective pollution probability of a farm follows a bivariate decision process: first, the farmer considers the impact exerted by the farm activities on the natural resources (soil, water, air) and on human health, and then he considers what he can do to control that environmental impact.

The application to the Italian case shows that the omission of the level of environmental concern from the estimation of effective pollution probability yields misleading evidence of policy unfairness in the allocations of subsidies between farms. This evidence, however, disappears when the level of actual pollution and environmental concern are estimated jointly in a bivariate framework.

The data set comes from a nationwide survey about the socio-economic characteristics of Italian agriculture conducted by the Italian Institute for the Studies of Agricultural Markets (ISMEA) in Italy in 1996. This data set contains several relevant variables for the definition of the pollution of the farm and its environmental effort. The results also show that the geographic location of the farm, the product specialization and the types of farm-household play a significant role in the determination of the level of effective pollution. This latter results can be useful in defining the environmental payments for farm with specific product and location characteristics. The presentation of the results accounting for farm type is of interest because it

[1] Minimal list of statutory legal standards – environment, food safety and animal welfare – with others at the request of Member State (Reg. (CE) n. 1782/2003, Reg. (CE) n.1257/2003).

permits synchronizing the environmental action with the rural development policy. We suggest the adoption of a policy that considers the level of effective pollution and that uses it to discriminate between farms in the attribution of subsidies.

The rest of the paper is organized as follows. We first describe the methodology followed to define the indexes used to measure the impact of farms on the natural resources (soil, water, air) and on human health, the farmers' attitude towards the environment and the effective environmental impact. We then present how the effective pollution probabilities can be modelled following a bivariate probit approach. Finally, the results of the application of the proposed methodology to a sample of Italian farms are presented.

Measures of the environmental impact at the farm level

Following the indication provided by the OECD (1997, 1999, 2001a, b), the impact of farms on the natural resources (soil, water, air) and on human health can be measured by using four classes of indicators[2] (Table A1):

1. Environmental pressure indicators;
2. Pesticide risk indicator;
3. Efficiency indicators;
4. Use indicators.

The first class of indicators includes the chemical loads associated with the use of fertilizers and pesticides per hectare of land, the water intensity indicator, the gas emissions indicators for machinery and livestock per hectare of land. The pesticide risk indicator combines information on pesticide toxicity and exposure with that on pesticide use. The third class of indicators includes the technical and economic efficiency indicators of water and fertilizers use. The last class contains the water and pesticides use indicators. These indicators can be normalized and aggregated into an Aggregate Index (*AI*) measuring the actual impact of the farms on the environment and the human health[3].

Then, the effective environmental impact can be obtained by taking into account the level of environmental concern of the farm. An Environmental Concern Index (*ET*) can be calculated accounting for, for example, the expenditures on abatement techniques, such as water depurators, or the adoption of low impact farming practises, such as integrated pest management, and the level of information of the farmer about abatement and control techniques.

Omission of the level of environmental concern from the estimation of the effective pollution can be the cause of policy unfairness. Take for instance the case of two farms with the same level of actual pollution but different sensitivity for the environment. It would be unfair to apply the same sanction to the farm with a high index of environmental sensitivity.

[2] See Table A1 for a list of all the indexes and the variables used to calculate them.
[3] The normalization, aggregation and rating procedures used are described in the Appendix and in Table A2.

The adoption of abatement and control technologies can lower the polluting load of this farm. We suggest linking the environmental sanction or payment to the value of the Effective Aggregate Index (*EAI*), obtained by the product of the Aggregate Index of actual pollution (*AI*) and the Environmental Concern Index (*ET*).

Finally, an Aggregate Production Intensity Index (*IIA*) can be calculated on the basis of the following information: the number of animals on the agriculture forage land; net returns on the labour unities; gross sales on net returns; set aside land on total agricultural land. The correlation analysis between the Aggregate Production Intensity Index and the Environmental Concern Index allows the identification of those farms adopting intensive techniques of production without caring for the environmental impact, that is those farms that are expected to produce the highest results, in terms of pollution abatement, whenever they adopted a more environmentally sensitive attitude. In this respect, these farms are the best candidate for environmental abatement payments.

Bivariate Probit Approach to the Determination of Effective Pollution Probability

This section outlines the econometric procedure for the determination of the farm's effective pollution probability. A farm can meet the cross-compliance requirements if it has a low/very low level of actual pollution or if it presents a high level of actual pollution (*AI_high=1*) associated with a high/very high level of environmental concern (*ET_high=1*), that is if it has a low level of effective pollution.

We can rescale the previously defined aggregate indexes to obtain two dichotomous variables:

$$AI\_high_i = 1 \text{ } if \text{ } AI = 4 \text{ } or \text{ } AI = 5; \quad = 0, \text{ } otherwise$$

$$EI\_high_i = 1 \text{ } if \text{ } ET = 1 \text{ } or \text{ ET} = 2; \quad = 0, \text{ } otherwise$$

where:

AI_high corresponds to a high level of the actual pollution Aggregate Index; it takes value 1 if the farm has a very high or high environmental impact (*AI = 4* or *AI = 5*) and 0 otherwise.

ET_high corresponds to a high level of the Environmental Concern Index; it takes value 1 if the farm has a very high or high environmental concern (*ET = 1* or *ET = 2*) and 0 otherwise.

Let define the farm's probability to meet the cross-compliance requirements, hence not to see its payments reduced, *P(SUB = 1)*. This probability is defined as follows:

$$P(SUB_i = 1) = P(ET\_high_i = 1 \cap AI\_high_i = 1) \tag{0.1}$$

As we already argued in the previous paragraph, the probability of effective level of pollution depends both on the farm level of actual pollution *(AI)* and on its level of environmental concern *(ET)*. The two variables follow a joint distribution, in this latter case the two variables are correlated and the value/probability of one depends on the value/probability of the other. The jointness in the distribution of the two variables can be tested and the correlation parameter can be estimated by making use of a bivariate probit model.

Estimation of the probability (0.1) to meet the cross-compliance requirements requires specification of actual pollution and environmental concern decisions. The farm's propensity to be environmentally sensitive can be defined as y^*_{1i}, which depends essentially on a vector of farm or farmer specific variables (x_{1i}). These include such characteristics as whether the farmer is informed about developments in environmental techniques, education, age, geographic location or productivity sector of the farm. Let y^*_{2i} denote the farmer's propensity to polluting. It depends on a vector of farm or farmer specific variables (x_{2i}) such as education, risk attitude and geographic location, productivity sector and the farm type.

We assume that other factors may exist that influence the attitude towards pollution of the farm that may be known to the farmer but not to the researcher. For example, the farmer's decisions may also be influenced by the fact that the farmer can consider the difference between the expected subsidies and the actual cost of the adoption of pollution abatement technologies. These unobserved factors may be included in the variable R^*. Thus, the polluting decision rules are as follows:

$$\begin{aligned} ET\_high_i &= 1 \; if \;\; y^*_{1i} = x_{1i}\beta_1 + \gamma_1 R_i^* + \varepsilon_{1i} > 0 \\ &= 0 \; if \;\; y^*_{1i} = x_{1i}\beta_1 + \gamma_1 R_i^* + \varepsilon_{1i} \le 0 \\ AI\_high_i &= 1 \; if \;\; y^*_{1i} = x_{1i}\beta_1 + \gamma_1 R_i^* + \varepsilon_{1i} > 0 \\ &= 0 \; if \;\; y^*_{1i} = x_{1i}\beta_1 + \gamma_1 R_i^* + \varepsilon_{1i} \le 0 \end{aligned} \tag{0.2}$$

$$y^*_{1i} = x_{1i}\beta_1 + \gamma_1 R_i^* + \varepsilon_{1i} \tag{0.3}$$

$$y^*_{2i} = x_{2i}\beta_2 + \gamma_2 R_i^* + \varepsilon_{2i} \tag{0.4}$$

where R^* represents the unobserved factors affecting the operator's attitude towards the environment. The implicit assumption is that, other things equal, an increase of R^* is expected to increase the probability of observing a higher environmental concern and a higher level of actual pollution, hence of receiving the payment. Thus we assumed the coefficients γ_1 and γ_2 positive. Given R^* is not known to the researcher, it cannot be treated as a regressor. It will be absorbed into the error terms $v_{1i} = \gamma_{1i} R_i * + \varepsilon_{1i}$ and $v_{2i} = \gamma_{2i} R_i * + \varepsilon_{2i}$ (Alberini et al., 1996). Note that farmer's environmental concern decision may not be independent of the level of pollution decision. As a consequence of no independence of the probabilities of environmental concern and pollution, we estimate (0.3) and (0.4) jointly as a bivariate probit model under the assumption that v_1 and v_2 are jointly normally distributed, the error terms ε_1 and ε_2 are independent of each other and $Cov(v_{1i} v_{2i}) = \gamma_1 \gamma_2 R_i * = \rho$. We assume also no endogeneity, that is $E(x_{1i} v_{1i}) = 0;\ E(x_{2i} v_{2i}) = 0$.

The effective pollution probability in Equation (0.1) can be written as $P(SUB_i = 1) = F(x_{1i}\beta_1, x_{2i}\beta_2; \rho)$, where F is the bivariate normal cumulative distribution function. The coefficients vectors β_1 and β_2 can be estimated by different variants of probit depending on the degree of observability in *AI_high* and *ET_high* (Meng and Schmidt, 1985). Since in our study both variables are observed for all *i*, the effective pollution probability can be estimated from a bivariate probit model with full observability (Zellner and Lee, 1965; Ashford and Sowden, 1970).

Data

The data set comes from a nationwide survey of the socio-economic characteristics of Italian agriculture conducted by the Italian Institute for the Studies of Agricultural Markets (ISMEA) in 1996. It collected information from 1881 Italian farms, where 95 percent (1777) are family farms. The survey combines information about household and farm characteristics, time use, farm profits, off-farm money income, governmental and inter-household transfers, consumption, technology, non-farm assets and information about the degree of autonomy in both farm and household decision making of the household members. The holistic design of the ISMEA survey minimizes the need to "crosswalk" surveys to produce estimates of total farm household income, extended and full incomes, and guarantees a high level of quality and data consistency (Smeeding and Weinberg, 1998).

The design of the ISMEA survey includes a section about the inputs and the outputs of the farm useful for the definition of the environmental impact indicators. For example, it considers the quantity of fertilizers and pesticides used for type of crop, the total quantity of water and combustible used in the production. Another section provides information about the level of concern of the farmer for the environment: if the farm received a premium for the environment, the total environmental expenditure, if the farmer reads environmental magazines, if he applies soil fertility analysis, integrated pest or weed or cryptogram management, the presence of hyper fertility program support or fertilizer use reduction. The survey provides also information about the number of labor units, the net returns, the gross sale, the forage cultivation land and the total agricultural land, which are useful to derive the level of production intensity of the farm.

The farm households are classified by macro region in North, Center and South-Island, by type of production (cereals, industrial crops, vegetables, fruit, olives, grapes, floriculture, livestock) and by farm types (limited resource, retirement, residential, commercial, small, medium and large farm-households, and non-family farms).

The variables used to estimate the effective pollution probability are described in Table 1 and Table 2.

The farm characteristics include the geographic location, the farm's altitude, the farm-type, the hectares of land, the type of production (crops and livestock), the farm to global income ratio, the intensity level of production.

The farmer characteristics include the farmer's age, the level of education, a variable taking value 1 if the farmer is born in the farm in order to capture the unobserved affection of the farmer to the land and so his concern about the environment.

The household characteristics include the number of children under 12 years old, the number of hours spent in pure leisure per week, the farmer risk propensity, who takes the decisions on the farm (wife, husband or alone).

Results

The results of the bivariate probit process are reported in Table 3. As expected the coefficient of correlation (rho) between the farmer's environmental concern decision and the level of pollution decision is found to be positive and significantly different from zero. The positive correlation indicates that the decision of the farmer to pollute increases the likelihood that the operator also takes into account the adoption of some pollution abatement technology, and vice versa. This shows the importance of the bivariate model[4].

The sign and significance of variable coefficients in both environmental concern (*ET_high*) and actual pollution (*AI_high*) equations suggest interesting conclusions. Among the farm types only the medium and the large farms have significant positive influence on the high level of environmental concern but at the same time they significantly affect the likelihood of high actual pollution. Limited resource farms have lower high actual pollution probability than large farms but they do not have significant influence on the likelihood of environmental concern. This is a first example confirming that the omission of the level of environmental concern from the estimation of effective pollution probability would yield unfair allocations of subsidies between farms. Farms specialized in crop production have both higher probability to be

[4] Several dummies variable have been omitted for the purpose of identification and we did not find evidence of multicollinearity.

highly environmental concern and higher likelihood to be actually polluting than livestock farms.

Moreover, the coefficients of geographic location variables, *south* and *north*, are negative but not significant in the environmental concern equation, while they are positive and significant at the 5% level in the equation of actual pollution. The altitude variable *plain* suggests that the propensity to be highly environmentally friendly is not affected by the farm's location but it increases the probability of being actually polluting. Note that, as expected, the coefficient of the high intensity productivity variable (*IIA_alto*) is significant and negative in the environmental concern equation. In general, it is not true that the farm-households adopting more intensive techniques are also those caring more for the environment.

It is interesting to note that all the family variables considered in this study have significant influence on the level of actual pollution, while they are not significant to explain the level of environmental concern. The wife decision in the farm affects the actual pollution probability negatively, whereas the number of children under 12 positively. But they do not influence significantly the environmental sensitivity for two different reasons: first, the female mean age of our sample is about 50 years so women might not have received in the past an environmental education, but at the same time, they express a native propensity to care about the family and so to care about the environment polluting less. We support this affirmation analyzing the coefficient of the variable age that is significant but negatively correlated with the likelihood to be highly environmentally sensitive. The older the farmer, the lower the probability that he is concerned about the environment. The second reason explains why the number of children under 12 does not affect the probability to be highly environmental concerned. The farm-households that have children may care primarily about them and may not spend time and money on concern for the environment.

The sign and the significance level of the coefficients of the farmer's personal characteristics variables have interesting implications. Lack of education does not affect the likelihood of being environmentally concerned and actually polluting. As anticipated, the age has a negative impact on the environmental sensitivity since the sample mean age is about 50 years, so they did not receive in the past an environmental education. Note also that the coefficient of the risk propensity variable is negative in both equations. Farmers more adverse to risk are less likely to be environmentally friendly and actually polluting. Finally, the positive and significant coefficient of the variable number of hours of pure leisure per week in the environmental concern equation suggests that larger the percentage of hours in pure leisure, larger is the chance for the farm to be highly environmentally concern.

Conclusions

Following a bivariate probit approach and using the 1996 survey conducted by the Italian Institute for the Studies of Agricultural Markets (ISMEA) in Italy, this study shows that the probability of effective pollution of Italian farm-households depends jointly on the actual level of pollution and the level of environmental concern of farms.

We found that large farms have significant positive influence on the high level of environmental concern but at the same time they significantly and positively affect the likelihood of high actual pollution, while limited resource farms have lower high actual pollution probability than large farms but they do not significantly affect the likelihood of environmental concern. Further, crop production farms have both higher probability to be highly environmental concern and higher likelihood to be actually polluting than livestock farms. These results show that the omission of the level of environmental concern from the estimation of effective pollution probability would yields misleading of fairness in the allocations of subsidies between farms.

It is also important to note that, in general, it is not true that the farm-households adopting more intensive techniques are also those caring more for the environment. The coefficient of the high intensity productivity variable is significant and negative in the environmental concern equation.

Then, we considered how the farmer and family's characteristics affect the actual pollution and environment concern likelihood. We found that wife's decision making on the farm affects negatively the probability of high actual pollution, whereas the number of children under 12 positively, but they do not influence significantly the high environmental sensitivity. We think that farm-households that have children may care about them above all and they may not spend time and money on concern about the environment, but at the same time, women express a native propensity to care about the family and so to care about the environment polluting less. Interesting is also that older the farmer, the lower the probability that he is concerned about the environment, this could be related to the fact that they might not have received in the past an environmental education. Further, farmers more averse to risk are less likely to be environmentally friendly and actually polluting, and larger the percentage of hours spent in pure leisure, the larger is the chance for the farm to be highly environmentally sensitive.

In conclusion, we suggest the adoption of a policy that considers the level of effective pollution and that uses it to discriminate between farms in the attribution of subsidies. Furthermore this study shows the necessity of providing more environmental education to farm-households in order to increase their environmental sensitivity.

References

Alberini A., Eskeland G.S., Krupnick A. and McGranahan G. (1996): "Determinants of Diarrheal Disease in Jakarta", *Water Resources Research*, 32 (7): 2259-69.

Ashford J.R. and Sowden R.R. (1970): "Multivariate Probit Analysis", *Biometrika*, 26: 535-46.

Caiumi A. and Perali F. (1997): "Female Labor Force Participation: A Comparison between Urban and Rural Families", *American Journal of Agricultural Economics*, 79: 595-601.

Falconer K. (2002): "Pesticide Environmental Indicators and Environmental Policy", *Journal of Environmental Management*, 65: 285-300.

Hadi A.S. (1993): "Procedures for the Identifications of Multiple Outliers in Linear Models", *Journal of the American Statistical Association*, 88 (424): 1265-72.

Johnson J. L. (1983): "Sex Differentials in Unemployment Rates: A case for No Concern", *Journal of Political Economy*, 91: 293-303.

Kovach J., Petzoldt C., Degni J. and Tette J. (1992): "A Method to Measure the Environmental Impact of Pesticides", *New York's Food & Life Sciences Bulletin*, 139, NYS Agricul. Exp. Sta. Cornell University, Geneva, NY.

Meng C. and Schmidt P. (1985): "On the Cost of Partial Observability in the Bivariate Probit Model", *International Economic Review*, 26: 71-86.

Metcalf R.L. (1975): *Introduction to Insect Pest Management*, New York, Wiley & Son, New York: 235-73.

Mohanty M.S. (1995): "Asymptotic Properties of the Two-Stage Bivariate Probit Estimator in the Presence of Partial Observability", *Journal of Quantitative Economics*, 11: 101-25.

Id. (2002): "A Bivariate Probit Approach to the Determination of Employment: A Study of Teen Employment Differentials in Los Angeles County", *Applied Economics*, 34: 143-56.

Niemeijer D. (2002): "Developing Indicators for Environmental Policy: Data-Driven and Theory-Driven Approaches Examined by Examples", *Environmental Science & Policy*, 5: 91-103.

Oaxaca R.L. (1973): "Male-female Wage Differentials in Urban Labor Markets", *International Economic Review*, 14: 693-709.

OECD (1997): *Environmental Indicators for Agriculture*, Vol. 1, Concepts and Frameworks, Paris.

OECD (1999): *Environmental Indicators for Agriculture*, Vol. 2, Issues and Designs, Paris.

OECD (2001a): *Environmental Indicators for Agriculture*, Vol. 3, Methods and Results, Paris.

OECD (2001b): *Environmental Indicators for Agriculture. Towards a Sustainable Development*, Paris.

Penrose L.J., Thwaite W.G. and Bower C.C. (1994): "Rating Index as a Basis for Decision-Making on Pesticide Use Reduction and for Accreditation of Fruit Produced Under Integrated Pest Management", *Crop Protection*, 13: 146-52.

Smeeding T.M and Weinberg D. (1998): "Toward a Uniform Household Income Definition", Paper prepared for the Canberra Group on Household Income Measurement Meeting in The Hague, The Netherlands.

Zellner A. and Lee T.H. (1965): "Joint Estimation of Relationships Involving Discrete Random Variables", *Econometrica*, 33: 382-94.

Appendix

The procedure followed to calculate the Aggregate Index of actual pollution (*AI*), the Environmental Concern Index (*ET*) and the Effective Aggregate Index of effective pollution (*EAI*) is a two step procedure: normalization and, aggregation and rating (Table A2).

a) Normalization

The normalization of the environmental indicators (x_i) is obtained by applying the following expression:

$$x_{ni} = (x_i - x_{i,\min}) / (x_{i,\max} - x_{i,\min});\ 0 < x_{ni} < 1$$

where:

x_{ni} = i-th normalized environmental indicator;
x_i = i-th environmental indicator;
$x_{i,min}$ = i-th environmental indicator's minimum value;
$x_{i,max}$ = i-th environmental indicator's maximum value.

The presence of outliers has been detected by making use of the method proposed by Hadi (1993), The identified outliers have been then included in the analysis after assigning them an unitary value. This procedure, on the contrary of that one based on re-weighting the indicators, has been preferred as it does not change the original ranking of the indicator.

b) Aggregation and rating

The aggregation procedure consists of two steps: first, we sum the normalized indicators in order to obtain an aggregate index $X_{ai} = \sum_{i=1}^{I} x_{ni}$. Second, we divide the distribution of X_{ai} into quintiles and assign to each of them a subjective weight 1-5 (Penrose et al., 1994; Kovach et al., 1992; Metcalf, 1975). Unitary value of the aggregate normalized indicator corresponds to the best behaviour of the farm towards the environment, while a value of five to the worst; note that instead in the case of the Environmental Concern Index, a value of 1 indicates very low environmental concern and value 5 very high environmental concern.

Table 1. Definition of Variables

Variable	Definition
Farm characteristics	
north	dummy==1 if the farm is in the North of Italy, 0 otherwise
center	dummy==1 if the farm is in the Center of Italy, 0 otherwise
south	dummy==1 if the farm is in the South of Italy, 0 otherwise
hill	dummy==1 if the farm is in hill, 0 otherwise
plain	dummy==1 if the farm is in plain, 0 otherwise
mountain	dummy==1 if the farm is in mountain, 0 otherwise
farm1	dummy==1 if Limited Resources farm, 0 otherwise
farm2	dummy==1 if Retirement farm, 0 otherwise
farm3	dummy==1 if Residential farm, 0 otherwise
farm4	dummy==1 if Small farm, 0 otherwise
farm5	dummy==1 if Medium farm, 0 otherwise
farm6	dummy==1 if Large farm, 0 otherwise
ocm_veg	dummy==1 if Crop farm, 0 otherwise
ocm_anim	dummy==1 if Livestock farm, 0 otherwise
IIA_low	dummy==1 if Low intensity farm, 0 otherwise
IIA_medium	dummy==1 if Medium intensity farm, 0 otherwise
IIA_high	dummy==1 if High intensity farm, 0 otherwise
re_d	dummy==1 if Global Income is > 50% than Net Returns, 0 otherwise
ha_tot	total agricultural land
Farmer characteristics	
education	dummy==1 if the farmer does not have education, 0 otherwise
age	age of the farmer
cap_YES	dummy==1 if the farmer is born in the farm, 0 otherwise
cap_NO	dummy==1 if the farmer is not born in the farm, 0 otherwise
risk_low	dummy==1 if the farmer has low risk propensity, 0 otherwise
risk_medium	dummy==1 if the farmer has medium risk propensity, 0 otherwise
risk_high	dummy==1 if the farmer has high risk propensity, 0 otherwise
leisure	number of hours of pure leisure per week
Family characteristics	
nchild	number of children <12 years old
dec_wife	dummy==1 if the wife takes decision in the farm, 0 otherwise
dec_housband	dummy==1 if the husband takes decision in the farm, 0 otherwise
dec_alone	dummy==1 if the farmer takes decisions by himself, 0 otherwise

Table 2. Descriptive Statistics for Selected Variables

Variable	Mean	Std. Dev.	Min	Max
Farm Characteristics				
farm1	0.075	0.263	0	1
farm2	0.023	0.151	0	1
farm4	0.122	0.328	0	1
farm5	0.665	0.472	0	1
farm6	0.037	0.188	0	1
ocm_veg	0.671	0.470	0	1
north	0.400	0.490	0	1
south	0.388	0.487	0	1
hill	0.776	0.417	0	1
plain	0.157	0.364	0	1
IIA_basso	0.078	0.268	0	1
IIA_alto	0.146	0.353	0	1
re_d	0.926	0.262	0	1
ha_tot	24.222	42.509	0.026	714.400
Personal Characteristics				
education	0.845	0.362	0	1
age	51.122	13.011	18	89
cap_YES	0.805	0.396	0	1
risk_basso	0.399	0.490	0	1
risk_alto	0.298	0.457	0	1
leisure	44.518	17.505	3	102
Family Characteristics				
nchild	1.166	1.096	0	7
dec_wife	0.405	0.491	0	1
dec_housband	0.026	0.158	0	1
1881 Observations				

Table 3. Bivariate Probit Estimate of Effective Pollution

Variable	Coeff.	t-Stat.	Coeff.	t-Stat.
	ET_high		AI_high	
Constant	-0.957**	-2.150	-2.839**	-2.839
Farm Characteristics				
farm1	0.186	0.520	0.557**	2.120
farm2	0.454	1.020	-0.704*	-1.750
farm4	0.409	1.200	-0.402	-1.540
farm5	0.678**	2.050	0.352	1.460
farm6	0.778*	1.920	1.182**	3.960
ocm_veg	0.742**	6.430	0.225**	2.770
north	-0.085	-0.740	0.885**	8.790
south	-0.180	-1.590	0.351**	3.480
hill	-0.466**	-1.620	0.580**	2.770
plain	-0.287	-2.300	1.348**	7.120
IIA_basso	-0.189	-1.100	0.147	1.110
IIA_alto	-0.299**	-2.150	0.079	0.780
re_d	-0.531**	-3.360	0.154	1.020
ha_tot	-0.331*	-1.930	-0.465**	-2.990
Personal Characteristics				
education	-0.113	-1.000	0.025	0.250
age	-1.152**	-3.320	0.022	0.080
cap_YES	0.163	1.270	0.163	1.610
risk_basso	-0.389**	-3.780	-0.215**	-2.500
risk_alto	-0.083	-0.800	0.030	0.340
leisure	0.560**	2.220	-0.251	-1.190
Family Characteristics				
nchild	0.450	1.150	1.670**	5.140
dec_wife	0.021	0.240	-0.111*	-1.670
dec_housband	-0.129	-0.470	-0.231	-1.010
rho	0.086	3.353		
Log Likelihood	-1862.12			
Sample Size	1881			

** Significant at 5% level; * Significant at 10% level.

Table A1. Environmental Indicators and Indexes

INDICATORS AND INDEXES	ID	Unit	Formula
1 - ENVIRONMENTAL IMPACT INDICATORS			
1.1 - Environmental Pressure Indicators			
1.1.1 - Chemical Load Indicators			
- fertilizers	ccch	kg/ha	$y_1 = x(\Sigma_i \alpha_i) / H$
- pesticides	ccfh	kg/ha	$y_2 = x\beta / H$
- total chemical load/ha	ccth	kg/ha	$Y = \Sigma_i y_i;\ i = 1, 2;$
1.1.2 - Use Intensity Indicator			
- water	iua	mc/ha	$Y = x / H$
1.1.3 - Gas Emissions Indicators			
- machinery/ha	inq_mach	mg/ha	$y_1 = z(\Sigma_i \phi_i \gamma) / H$
- livestock/ha	inq_livh	mg/ha	$y_2 = t(\Sigma_I \phi_I \gamma) / H$
- total gas emissions/ha	inq_gash	mg/ha	$Y = \Sigma_i y_i;\ i = 1, 2;$
1.2 - Pesticide Risk Indicator			
- pesticide risk indicator based on LD50	ind	kg/ha	$Y = [\Sigma_i (x_i / tox_i) \times 1000] / H$
1.3 - Efficiency Indicators			
1.3.1 - Technical Efficiency			
- water use	ietac	kg/kg	$y_1 = P/x$
- fertilizers use	ietc	kg/kg	$y_2 = P/x$
- total technical efficiency	iet	kg/kg	$Y = \Sigma_i y_i;\ i = 1, 2$
1.3.2 - Economic Efficiency			
- water use	ieea	€/kg	$y_1 = Q/x$
- fertilizers use	ieec	€/kg	$y_2 = Q/x$
- total economic efficiency	iee	€/kg	$Y = \Sigma_i y_i;\ i = 1, 2;$
1.4 - Use Indicators			
- water	ua	kg/kg	$Y = x / \bar{A}$
- pesticides	iuf	kg/kg	$Y = x / \bar{A}$
2 - AGGREGATE INDEX	AI	1 – 5	
3 - ENVIRONMENTAL CONCERN INDEX	ET	1 – 5	
4 - EFFECTIVE AGGREGATE INDEX	EAI	1 – 5	EAI = AI*ET
5 - Intensity production indicators			
- number of animals (n) / forage cultivation land (F)	int1	num/ha	$y_1 = n / F$
- Net Returns (NR) / Labour Units (LU)	int2	€/num	$y_2 = NR / LU$
- Gross Sales/ Net Returns (NR)	int3	€/€	$y_3 = GS / NR$
- Set Aside / total agricultural land	int4	ha /ha	$y_4 = SA / H$
- AGGREGATE INTENSITY INDEX	IIA	1 - 5	$Y = \Sigma_i y_i;\ i = 1, \dots, 4;$

LEGEND
x = quantity of input used; H = total agriculture land; α = average percentage of Nitrogen (i = 1), Phosphor (i = 2), Potassium (i = 3); β = average active principal of pesticides; z = quantity of combustible used; ϕ = gas emission of CO_2 (i = 1), CH_4 (i = 2), N_2O (i = 3); γ = GWPs over 100 years for CO_2 (γ = 1), CH_4 (γ = 21), N_2O (γ = 310); t = quantity of manure used; tox = toxicity index = ($\beta \times 1000000$) / LD50; P = output; Q = monetary output; $\bar{A}$ = average quantity of input used in the sample.

Table A1. Normalization, Aggregation and Rating

ENVIRONMENTAL INDICATORS AND INDEXES	ID	Range
1. Normalized total chemical load indicator/ha	ccth_st	0-1
2. Normalized water use intensity indicator	iua_st	0-1
3. Normalized gas emission indicator/ha	inq_gash_st	0-1
4. Normalized pesticide risk indicator	ind_st	0-1
5. Normalized water use indicator	ua_st	0-1
6. Normalized pesticide use indicator	iuf_st	0-1
7. Normalized technical efficiency indicator	iet_st	0-1
8. Normalized economic efficiency indicator	iee_st	0-1
A. AGGREGATE INDEX (AI =1+2+3+4+5+6+7+8)	AI	1-5
1 = Very low environmental impact; 2 = Low environmental impact; 3 = Medium environmental impact; 4 = High environmental impact; 5 = Very high environmental impact.		
9. Normalized intensity indicator 1 (animals/ha	int1_st	0-1
10. Normalized intensity indicator 2 (NR/LU)	int2_st	0-1
11. Normalized intensity indicator 3 (GS/NR)	int3_st	0-1
12. Normalized intensity indicator 4 (set aside/total land)	int4_st	0-1
B. AGGREGATE INTENSITY INDEX	IIA	1-5
1 = Very low intensive farm; 2 = Low intensive farm; 3 = Medium intensive farm; 4 = High intensive farm; 5 = Very high intensive farm.		
13. Normalized premium for environmental concern	amb_pre_st	0-1
14. Environmental expenditure / Total expenditure	amb	0-1
15. Soil fertility analysis	ana_fert	0-1
16. Integrated pest management	log_inse	0-1
17. Integrated weed management	log_crit	0-1
18. Integrated cryptogam management	log_male	0-1
19. Hyper-fertility program support	iper_fert	0-1
20. Normalized magazine readings	riv_spec	0-1
21. Fertilizer use reduction	inizia_h	0-1
22. Normalized (Set aside/total agricultural land)	int4_st	0-1
C. ENVIRONMENTAL CONCERN INDEX (ET = 13+14+15+16+17+18+19+20+21+22)	ET	1-5
1 = Very low environmental concern; 2 = Low environmental concern; 3 = Medium environmental concern; 5 = Very high environmental concern.		
D. EFFECTIVE AGGREGATE INDEX (EAI = AI × ET)	EAI	1-5
1 = Very low polluting farm; 2 = Low polluting farm; 3 = Medium polluting farm; 4 = High polluting farm; 5 = Very high polluting farm		