AGRICULTURAL
ECONOMICS

Agricultural Economics 36 (2007) 49–64

Efficiency, subsidies and environmental adaptation of animal farming under CAP

Werner Kleinhanß[a], Carmen Murillo[b], Carlos San Juan[c], Stefan Sperlich[d,*]

[a]*Federal Agricultural Research Centre, Braunschweig, Germany*
[b]*Departamento de Economía, Universidad de Cantabria, Santander, Spain*
[c]*Departamento de Economía, Universidad Carlos III de Madrid, c/Madrid 126, 28903 Getafe, Spain*
[d]*Georg-August Universität, Institut für Statistik und ökonometrie, Platz der Göttinger Sieben 5, 37073 Göttingen, Germany*

Received 7 March 2005; received in revised form 15 November 2005; accepted 24 April 2006

Abstract

Q1 The purpose of this article is to model the interaction between the targets of the current CAP: environmental adaptation, subsidies, and efficiency of animal farming. To this end we first have to identify the production frontier and relative efficiency level for each animal-oriented farm in the sample. The production frontier and efficiency index for each type of farm (assuming no specific production functions) are identified using Data Envelope Analysis (DEA) techniques. We then address the relationship between relative efficiency, farm size, and environmentally friendly behavior by carrying out a nonparametric regression of efficiency, on economic size, a proxy for the degree of environmental appropriateness, and regional dummies. Calculations of the efficiency of the farms including direct subsidies are compared with the counterfactual exercise in the case in which direct subsidies are not considered. Finally, we look for relations between subsidies and factors such as farm size, efficiency and environmentally friendly behavior. One key result shows that, on average direct payments generally tend to increase efficiency. However, in most of the cases the mean efficiency decreases as the percentage of direct payments rises. Direct payments are found to be positively related to environmentally friendly production, at least in Germany. However, in general, the direct payment system is not sufficient to offset the fact that the less environmentally friendly farms as well as the larger farms are more efficient.

Q2 *Keywords:* Efficiency; Subsidies; DEA; Nonparametric regression; Environmental friendly farming; Natural resources

1. Motivation and organization

Environmental adaptation and efficiency have become key issues in new European agricultural policy. The recent animal epidemics (e.g., mad cow disease, foot-and-mouth disease, avian influenza) and consumer reactions have drawn attention to the need for environmental adaptation of animal husbandry. The agreement of the Council of Ministers in June 2003 (Mid-Term Review of the Common Agricultural Politic [CAP MTR]), means a step toward the decoupling of income from prices. Additionally, CAP MTR introduces a modulation of the direct payment (e.g., limiting direct payments by size).[1] Decoupled aid means that in the future, the vast majority of subsidies will be paid independently of the volume of production. According to the Commission, the key elements of the new reformed CAP, in a nutshell, are:

- a single farm payment to be made to EU farmers, independent of production; limited coupled elements may be maintained to avoid abandonment of production;
- this payment will be linked to respect for environmental friendly behavior, food safety, animal and plant health and animal welfare standards, as well as the requirement to keep all farmland in good agricultural and environmental condition ("cross-compliance");
- a strengthened rural development policy using more EU budget;
- a reduction in direct payments ("modulation") for lager farms to increase the budget for the new rural development policy.

Q3 * Corresponding author. Tel.: ??? ??? ???; fax: ??? ??? ???.
E-mail address: stefan.sperlich@wiwi.uni-goettingen.de (S. Sperlich).

[1] Commissioner Fischler, the European Commission member responsible for Agriculture, Rural, Development and Fisheries, remarked (SPEECH/03/326): "The agriculture reform has been agreed. . . . Farmers will enjoy more income stability, more freedom to produce what the market wants, and a system of support which is much easier to justify from a social point of view. Consumers and taxpayers will receive more for their money: more transparency, more quality, more environmental protection and animal welfare".

In the millennium round WTO negotiations, the exporting countries of the Cairns group asked for the total abolition of the agricultural subsidies. The Commission has successfully argued for multifunctionality to maintain or even increase direct payments when reducing intervention prices in the WTO negotiations. This approach should provide a tool for promoting environmentally friendly practices in agriculture.[2]

Another perspective is that supporters of direct payments argue that this tool is used to increase the efficiency of farms with structural problems. Critics argue that subsidies allow farmers to continue producing below the efficiency frontier. Some critics, especially from outside the EU, also argue that by allowing inefficient European farmers to survive, the CAP harms efficient agriculture overseas. One of the main justifications for the direct payments to farmers is their positive societal impact through nature and environmental conservation and increased efficiency. Direct payments are defined as subsidies decoupled (not linked to) from the output level. Surprisingly few studies are available, which calculate whether the 1992 CAP reform increases in the level of direct payments were followed by a rise in efficiency, especially for the environmentally friendly farms. This could in fact politically justify the reinforcement of the status quo with minor changes. Commissioner Fischler, on the other hand, argued that EU farmers will become more competitive by increasing their efficiency: "Severing the link between subsidies and production will make EU farmers more competitive and market orientated". Certainly, see Álvarez (2001), efficiency does not automatically imply that farms are competitive. However, at least empirically, strengthened market orientation does not always correlate clearly with environmental conservation.

The aim of this article is the analysis of this triangular relationship between efficiency, environmental friendliness, and subsidies in the EU. More specifically, we seek answers to the following questions.

First, is it still costly to be environmentally friendly in the EU? Note that answering this question is a rather complex task due to endogenity problems and because the notion of efficiency is ambiguous; it could mean economic efficiency from the farmer's point of view (e.g. including direct payments) or conventional technical efficiency (looking at the production process without subsidies).

The second question is: How much do subsidies compensate small and environmentally friendly farms? More specifically we can study how much they improve compared with the large environmentally unfriendly farms after, rather than before subsidies. This is equivalent to revealing the impact of the direct payment system on farm efficiency under the farmer's status quo behavior.

So, to answer these questions we propose three basic steps:

In the first step, efficiency is calculated taking into account the direct payment (*DP*) received, calling this coefficient E_{DP}. Direct payments are defined as subsidies not directly linked to the output level. In other words, the farmer has to take decisions bearing in mind that, within the possibilities of farm production, certain outputs include byproducts (positive environmental externalities) with a direct payment as a monetary compensation, whereas others have no compensation or even a cost (environmental tax for a negative externalities).

Afterward, efficiency is computed excluding direct payments, calling this coefficient E_W. The results rank farms according to efficiency in a world without direct subsidies, so they are neither rewards for positive externalities nor penalties for negative externalities. The E_W calculation assumes the existing input/output prices (i.e., coupled support levels remain unchanged). For example, the efficiency impact of a hypothetical variation of the intervention prices is (by intention) not considered in this exercise. It may be emphasized that we indeed want to compare the efficiency with versus without *DP* under the farmer's "status quo" behavior.

In other words we want to measure the loss/gain of economic (i.e., monetary) efficiency for being environmentally friendly. We always refer to E_{DP} when speaking of economic efficiency, and to E_W when speaking of conventional efficiency. However, the calculation of E_W and E_{DP} is not of direct interest in our article but serves instead as an auxiliary step.

After having calculated E_{DP} and E_W, we regress efficiency on environmentally friendly (EF), and on other factors such as economic size (ESU) to get rid of possible endogenity. We find that in general, being environmental friendly and/or small is costly. Also we find strong positive interaction between size and environmental unfriendliness. This might not be surprising, but now we can look at what the European Union was really doing to counteract (and give credibility to their "cross-compliance" argument).

To answer the second question all we have to do is to take the regression curves from the above and compare the regression of E_W with the regression of E_{DP}, that is, for the statistical part of this analysis we need the same steps as we need to answer the first question.

Note that in this counterfactual exercise we study how efficient the farms would be under the same allocation policy but without receiving *DP*. This point is important to understand because we are not interested in measuring the efficiency (without *DP*) after farms have adapted to the new situation, for example, having become less environmentally friendly or increased farm size etc. Such a study would be interesting when looking at competitiveness but is clearly beyond of the scope of this article.

Next, the third question deals with the issue of promoting conventional efficiency, E_W, using direct payments. The question to answer is: Do the subsidies (at least) promote conventional efficiency? Investigating this is fairly simple because due to the way we have calculated E_W, there is no problem of endogenity when looking at the linear and semilog-linear regression (i.e., correlation and semilog correlation) of E_W on *DP*.

This produces our fourth and final question: How are the *DP* related to size? Here we have simply looked at different linear

[2] E.g. in Spain, animal farms with extensive practices (i.e. low livestock units per land area), frequently located in mountainous areas or Mediterranean forest grazing lands, illustrate the complementary relationship between animal farming and landscape preservation.

and log-linear relationships between *DP* and size measured in European Size Units (*ESU*). We believe that this is fair enough in our context as we are not interested in the intentions of the European Community (as then one would have to take into account the possible endogenity of *ESU*). Moreover we are interested in the absolute (direct, indirect, or spurious) relation between *DP* and *ESU*. We find that this relation is up to 97% which leads us to conclude that this subsidy policy is counterproductive. The positive correlation is not surprising as direct payments are mainly based on "per head" and "per hectare," but not exclusively, and therefore the strength of positive correlation is surprising.

To study these questions at the European level, we chose Spain and Germany as representatives of Continental and Mediterranean livestock raising. The article aims to replicate efficiency levels under different policy scenarios. We use the German and the Spanish sample data from the Farm Accounting Data Network (FADN) from 1999 and 2000 by type of animal farming with positive plant production. Note that more recent data, though available, are strongly disturbed by the BSE (also called mad cow) crisis. We repeat efficiency calculations for 2 years to test for the influence of random weather variability (e.g., pasture availability).

For our analysis we always use nonparametric methods when parametric model (mis)specification could provoke serious disturbances in our conclusions. When we speak of model (mis)specification we do not refer to variable selection but to functional form specification. This greatly increases the econometric effort as well as the variance of our results but avoids any problems of model specification.

Efficiency is measured using an index calculated with DEA (Data Envelope Analysis), and with a counterfactual index ignoring the direct payments for each individual farm (for details see Section 3.1). The counterfactual index measures the level of efficiency distortion on the economic behavior of the holding because under CAP it is possible to "farm subsidies" on the
Q4 top of agricultural products. This allows us to compare relative efficiency with direct subsidies included in the efficiency index with the case in which they are eliminated in the so called counterfactual efficiency calculations. From the above discussion one might already see that in both cases, E_W and E_{DP} will have to be calculated as what in the classical DEA literature is called "technical efficiency" but with all variables measured in monetary units. We are not speaking of efficiency as the product of allocative times technical efficiency (in physical units) as is often done in DEA literature. For more details see Section 3.

The implications are important for the future application of the recently approved CAP Reform 2003 on a historical basis. Results can potentially be translated into promoting the wrong type of farming, as in past years, for example, the conversion of price support into direct payments based on the past year's level of protection (the "historical rights" argument).

2. Model and data

DPs were originally introduced in the McSharry CAP reform to decouple income from guaranteed prices, and to control the intervention stocks. The theoretical idea behind the introduction of direct payments is the antiproduction premium that assumes that the farm income is the result of the difference between income and cost plus the antiproduction premium (decoupling of price and income policy). This premium should be related to the income loss due to the output reduction. In practice direct payments were fixed by area, by head of certain LIVESTOCK and others. This means that in a common market organization with overproduction, the Commission typically calculates the *DP* to compensate farmers for an income reduction.

That policy can be reinforced through the reduction of the guaranteed prices (e.g., reducing the intervention price to approach world market prices) with the target being to eliminate intervention stocks and to control the CAP budget by reducing intervention acquisitions and/or export restitutions (subsidies). In fact, the total elimination of the export restitutions is under negotiation in the current *millennium round* of the WTO (World Trade Organization). The Commission is also seeking to transform every direct payment into single farm payments.

In practice, the real impact of guaranteed price reductions on farm income depends on the level of efficiency of the farm with respect to the standard level imputed by the policy makers in the regions. Therefore, the policy measure (i.e., direct payment) is calculated mainly based on the set-aside area and/or the number of heads per holding. On top of that, the *cross compliance* criterion requires a minimum area per head (of cattle) to qualify for a direct payments program. For this reason our environmentally friendly proxy is livestock unit equivalents per agricultural utilized area (LU/UAA) as we want to replicate the policy makers' criteria.

The econometric task is to assess the impact of being environmentally friendly on being efficient, and to determine the relations between direct payments and other factors such as farm economic size, environmentally friendly, behavior and economic efficiency. Under price intervention it is difficult to assume that marginal costs equal prices as is often required in parametric production functions for inputs and outputs (i.e., under perfect competition), even in the most flexible specifications. Furthermore, we prefer not to use specific parameterization for our models, neither for the production function when calculating efficiency, nor for the regression when estimating the effect of size and environmental friendliness on efficiency.[3]

Therefore we have decided to apply rather sophisticated nonparametric methods as otherwise the chosen parametric specifications would clearly have a direct impact on the results. Certainly, it is well known that this, at least for moderate sample

[3] As a referee noticed, the effect of size on efficiency could also be studied by other methods, e.g. through the measurement of scale efficiency, see Førsund (1996) or Banker et al. (2004).

sizes, greatly complicates precise inference (e.g., significance testing). On the other hand, the results we see are uncorrupted in the sense that they do not vary with the subjective model chosen.

The producer level of direct support is measured by the total amount of direct payments in farm accounts; size (*ESU*), and livestock unit equivalents per agricultural utilized area (LU/UAA) as a proxy for measuring how environmentally friendly the farm is, see discussion above. In contrast, the output efficiency cannot be directly observed and will have to be calculated in a first step by DEA, see also Section 3.1.

Our variables to calculate efficiency of production are the following[4]:

OUTPUT	Q_v	crop output
	Q_a	animal output
INPUT	K_f	farm capital, mainly buildings and machinery at present value;
	C_a	fodder and other animal linked inputs;
	C_v	crop linked inputs (fertilizer, agrochemicals, seeds, water and other crop specific inputs, fuels and lubricants);
	W	wages;
	UAA =	Utilized Agricultural Area of farm aggregate adjusted by quality;
	$A_I + A_{NI}$	(including pasture, arable land and permanent crops adjusted by quality, i.e., geographical situation and if irrigated or not);
	SP	"shadow price"/opportunity costs for producing in a non-subsidized way.

Livestock is considered in the farmer's output because output is defined here by stock variation and sales, for details see McKay et al. (1983). In total, five inputs can be used plus SP. The so called "shadow price" (SP) is calculated from direct payments (including premiums) with a negative sign, therefore we set SP $= -DP$.

The present value of K_f is calculated based on the acquisition price and imputing a technical amortization by average life of every type of asset (building or machinery and transport vehicles, for details see Ball et al., 2004). This excludes the livestock units and the land area.

[4] We have always used nominal prices and real quantities accounted by the farmers as given in FADN. It could be argued that world market prices should be used instead, but 1(a). it is very difficult to calculate new output and input levels for each individual farm under world market conditions. Therefore efficiency measures would be biased by assumptions of world market prices. 2(b). It is beyond the scope of this article to study the efficiency under different trade liberalization scenarios.

When we say "adjusted by quality," this means that we have calculated the value of input "land" for Spain as follows:

$$\text{land value } = A_I * P_I + A_{NI} * P_{NI}, \tag{1}$$

where A_I: Agricultural Utilized Area (UAA) irrigated (ha.); A_{NI}: UAA nonirrigated (Ha.); P_I: price UAA irrigated by region (Euros/ha); P_{NI}: price UAA nonirrigated by region (Euros/ha). The average price by region is a weighted average of the different types of irrigated and nonirrigated lands in the region (e.g. grass land divided into two quality levels, plus arable land, green houses, permanent crops (by type of tree: fruits other than citruses, citruses, vinegars, olive) see for details Decimavilla and San Juan (2002). For Germany we use total UAA since differentiating between irrigated and non irrigated land does not make much sense.[5]

Having the efficiency at hand, we use a regression model to study the level of compatibility between different targets of the new CAP, including environmental conservation and competitiveness at the farm level.

We want to measure the impact of being environmentally friendly by filtering out the regional and size effects. As indicated before, the increase in consumer demand for more environmentally friendly products on the one hand, and the justification of the incentives paid under CAP on the other hand, have made "being environmentally friendly" an important issue of output efficiency.

Beyond the objective of more market-oriented agriculture, the new CAP uses efficiency as a key factor. Therefore, we estimated the following model [in EUR]:

$$E = g[\ln(EF), \ln(ESU)] + \beta^T R + e, \tag{2}$$

where E is efficiency, EF indicates the environmentally friendly degree proxy measure, ESU is the European Size Unit, and R is a vector of dummy variables for agricultural regions divided into North, Center, Northeast, South, and East for Spain, and North, Center, South and East for Germany, respectively. Farm location is reported at a general agricultural regional level, a geographical unit that includes several provinces (in Spain) or Länder (in Germany). Therefore, in the nonparametric model we use location parameters to control for possible different regional endowments. The aggregation by large agrarian regions of Spain is based on the geographical specialization reported by Mora and San Juan (2003). The aggregation of the LU is made with the standard procedure used by FADN and EUROSTAT. The variables come from the individual accounting collected under FADN normalization. The detailed input and output information of each farm account is fully utilized to calculate the aggregate variables that include all production costs.

[5] There are three things that might arouse curiosity; why we do not aggregate the two outputs, why we include SP as input instead of considering *DP* as a third output; and why we have not aggregated more (or less) on the input side. The reasoning for this comes immediately from the DEA method so we have postponed this discussion until the end of Section 3.1.

The function $g: \Re^2 \rightarrow \Re$ is not specified further because the impact of $\ln(EF)$ and $\ln(ESU)$ turned out to be nonlinear and to have (strong) interaction. Finally, the "error" term e stands for the not further specified heterogeneity. As $g(\cdot, \cdot)$ is nonparametric, we could have directly used the covariates *ESU* and *EF* in model (2). The logarithm does, therefore, not impose any model specification here. This variable transformation is only due to *smoothing necessities*, see Section 3.2.

As we analyze here the subsidy policy, for a fair evaluation we have to choose the same measure that the European Community applies, that is, Livestock Units per Agricultural Utilized Area, i.e., LU/UAA (see discussion in the Introduction). The LU/UAA indicator is highly related to other environmental amenities of the farm such as untouched landscape, traditional buildings, wild animal habitats, biodiversity, preservation of the regional nonintensive productive (endemic) livestock types, and shows the potential of grazing feed of the livestock, usually negative related to feedstuff consumption. Under the current CAP regulation (since year 1993) the farms must present a minimum LU/UAA to qualify for *DP*. For that reason some farms in our sample show no *DP*, but it could also happen that the farm does not have subsidies for other reasons (i.e., certain products do not have *DP*-specific programs but farms can be engaged in other environmental programs that provide subsidies to improve environmental behavior, in which case *DP* are included). Table 1 gives the *DP* distribution by type of animal and country.

Note that the smaller EF is, the more "environmentally friendly" the farm is. Note further that defining EF by LU/(land value) does not change the overall final results.

All variables are taken from FADN, except land prices which come from the Agricultural Land Prices Survey for Spain (Encuesta de precios de la Tierra (Base 1997), Boletin Mensual de Estadistica Agraria, November 2002). For Germany we use land rents derived from the rental prices of rented land; in cases of missing values we have used the regional average derived from the underlying sample.

We selected farms oriented to livestock production, that is, farms with larger animal than crop output in both years. In order to include only farms with similar production functions and with the possibility of cropping vegetal products for reuse on the farm or for sale, we selected only farms with both positive animal and crop production. We believe that farms without land have a noncomparable production function and should therefore be excluded, for example fattening farms.

For our analysis, the sample is split by country (Germany and Spain), and by type of livestock farming (cattle farming, pig farming, and sheep and goats) as these different farm types are uniform neither in their treatment by the CAP nor in their production processes. So we did all calculations separately for each year, country, and farm type. As mentioned earlier, for comparison reasons we will determine an efficiency index with $[E_{DP}]$ and without $[E_W]$ direct payments.

We use data from the sample in Spain and Germany (1999–2000) of the FADN. Every year, the survey gathers information on the characteristics of the farm (UAA, LU, type of livestock and crops, economic size of the farm, etc.) and nominal production (animal and crop output) for a representative sample of holdings at the regional level in Germany and Spain. Unfortunately, for Germany information on sheep and goats is only available for very few farms, thus a statistical analysis is not sensible.

The same FADN survey also provides detailed information on input expenditures by farm. As mentioned above, for the selected farms, livestock output is always greater than crop output (fodder, field crop, grain cereals, vineyards, potatoes, industrial crops, plants, fruits, dried pulses, olive groves and others). Table 2 summarizes the number of farms that are used for all the calculations, separated by year, animal type and country.

3. Methodologies

Even though these methodologies are not completely new, some readers might not be familiar with DEA or with nonparametric regression. Therefore we offer a brief overview for a better understanding and interpretation of the results presented later.

3.1. Estimation of efficiency

DEA is a nonparametric approach using linear programming methods to determine the envelopment of the (decision-making units) DMUs thus identifying "the best practice" for each productive unit.

Table 1
Numbers and percentages of farms with *DP* per country and year

Type/year	Numbers		Percentages	
	1999	2000	1999	2000
		Spain		
Cattle	996	1313	69.408	85.094
Sheep & goats	553	679	100	100
Pig farming	233	232	91.373	93.173
Poultry	24	26	80.000	81.250
Total	1847	2288	79.750	89.937
		Germany		
Cattle	353	358	97.245	98.623
Pig farming	245	245	100	100
Total	598	603	98.355	99.178

Table 2
Number of farms used in our calculations

Year		1999	2000
	Sheep and goats	553	679
Spain	Cattle	1,435	1,543
	Pig	255	249
Germany	Cattle	604	604
	Pig	355	355

Then, measures are calculated relative to this frontier for each individual Debreu-Farrell efficiency (see Cooper et al. (2000), for a comprehensive description of the methodology). The main advantage of DEA is that there is no need to specify a particular functional form for the production frontier, though the assumption that there is no random error might be seen as a drawback. However, excluding measurement errors, this question depends only on the definition of "efficiency."

Let $x \in \Re^p$ and $y \in \Re^q_+$ denote input and output vectors, respectively, with which we may define the following set of the feasible input–output combination,

$$\Psi = \{(x, y) \in \Re^{p+q} : \quad \text{x can produce y}\}. \tag{3}$$

For any $y \in \Re^q_+$ we may define the previous set by the input requirement set defined as

$$X(y) = \{x \in \Re^p : (x, y) \in \Psi\}, \tag{4}$$

where the input efficient frontier may be defined by the following isoquant:

$$\delta X(y) = \{x \in X(y) : \theta x \notin X(y) \quad \forall \theta < 1\}, \tag{5}$$

and therefore, the corresponding Farrell input-oriented measure of efficiency (Farrell, 1957) is specified as the following distance function:[6]

$$\theta(x, y) = \inf\{\theta : \theta x \in X(y)\}. \tag{7}$$

So $\theta(x, y)$ defines the input efficiency (the maximum contraction) along a fixed ray away from the efficient input. For example a value of $\theta(x, y) = 1$ means that the producer is input efficient while a value of $\theta(x, y) < 1$ indicates that the producer is input inefficient and he may reduce inputs in that proportion while maintaining the output level.

Alternatively, one could formulate (5)–(7) as an output-oriented problem. In practice, the input orientated is more popular due to its easier interpretation. However, in particular if we include *DP* as output instead of SP $= -DP$ as input, an output-orientated DEA version would be more intuitive as a farmer is interested in maximizing the direct subsidies received. But maximizing *DP* and minimizing SP is the same, and similar problems could be discussed for many of the other inputs when choosing an output orientated DEA. Nevertheless, we repeated some of the calculations with the output-orientated DEA, which certainly suffers to some extent from similar criticism as the input-orientated DEA. In the Appendix we give for the Spanish data histogram plots of the E_{DP} and E_W for 1999 and 2000 when the indices are calculated with an input-orientated DEA, and histograms of the differences of the input minus the output-orientated E_{DP}. As can be seen, the differences are minor (but for pig farms an output orientated DEA discriminates the farms somewhat more). Not surprisingly, the final conclusions turn out to be the same.

Further alternatives are for example DEA methodologies that allow us to combine both minimizing input and maximizing output, see Gonzälez Fidalgo (2001) or Banker et al. (2004). On the other hand these methods also need several assumptions that do not necessarily fit in our context. For example, in Gonzälez Fidalgo (2001) minimizing input and maximizing output is restricted to always occur in the same proportion. Finally, DEA also allows for fixing some of the input, and respectively output factors, see Banker and Morey (1986), but we found this does not fit to our case because all input factors can be potentially changed by the farmer's decision, for example additional land could be rented.

Finally, some readers may be puzzled by the counterfactual exercise, that is, including, respectively excluding, SP in DEA. First, note that as we just calculate two different convex hulls under the status quo allocation, DEA is not a regression problem. Second, compare our treatment of SP with Seiford and Zhu (2002)'s treatment of undesirable inputs/outputs under the assumption that classification of DEA efficiencies and inefficiencies are invariant to data transformations. Third, our key argument is that in our counterfactual exercise we really want to compare how farms do economically under the status quo allocation (and thus their present environmental friendliness) with versus without *DP*.

For the rest of the article we therefore concentrate on the presentation of the numerical results based on the input-orientated DEA. The estimation of this above-defined concept requires some assumptions (see Färe et al. 1994) for both the production possibility set (mainly convexity and free disposability of inputs and outputs) and the distance function. The first model proposed under the methodology called DEA (Charnes et al. 1978) was defined under constant returns to scale, but later papers have considered alternative assumptions such as the case of variable returns to scale by Banker et al. (1984). In any case, and under some regularity assumptions on the data-generating process specified in Kneip et al. (1998), DEA allows consistent estimation of the above concepts (see Simar and Wilson (2000) for a review of the DEA statistical properties).

For a sample of n producers, the DEA estimate of the production set ($\widehat{\Psi}$) under the least restrictive returns to scale assumption (i.e., variable returns) is:

$$\widehat{\Psi} = \{(x, y) \in \Re^{p+q} : x \geq \sum_{i=1}^{n} \gamma_i x_i, \quad y \leq \sum_{i=1}^{n} \gamma_i y_i, \quad \sum_{i=1}^{n} \gamma_i = 1, \quad \forall \gamma_i \geq 0\}, \tag{8}$$

where γ_i is the intensity vector of firm i and defines its *best practice or benchmark firm* by a linear combination of all the firms observed in the sample. Constraint $\sum_{i=1}^{n} \gamma_i = 1$ imposes variable returns to scale into the benchmark technology while

[6] Equivalently, Farrell's input efficiency may be described by the Shephard input distance function

$$\delta(x, y) = (\theta(x, y))^{-1} = \sup\{\delta \mid \frac{x}{\delta} \in X(y)\}. \tag{6}$$

the first two constraints in the Equation (8) imply that an excess of outputs or inputs can be disposed off freely.

The DEA estimates of equations (4) and (5) are then

$$\hat{X}(y) = \{x \in \Re^p \mid (x, y) \in \hat{\Psi}\}, \tag{9}$$

$$\delta\hat{X}(y) = \{x \in \hat{X}(y) \mid \theta x \notin \hat{X}(y), \quad \forall\theta < 1\}, \tag{10}$$

while the estimate of the Farrell (technical) efficiency measure, see Farrell (1957), is computed by linear programming techniques as follows:

$$\hat{\theta}(x_j, y_j) = \min\left\{\theta : \sum_{i=1}^{n} \gamma_i x_i \leq \theta x_j, \quad y_j \leq \sum_{i=1}^{n} \gamma_i y_i, \right.$$
$$\left. \sum_{i=1}^{n} \gamma_i = 1, \quad \forall\gamma_i \geq 0 \right\}. \tag{11}$$

Since by construction $\hat{\psi} \subseteq \psi$, the estimator $\hat{\theta}(x_j, y_j)$ constitutes a downward-biased estimator of $\theta(x_j, y_j)$. The analyzed firm j is technically efficient if and only if $\hat{\theta}(x_j, y_j) = 1$ and it is placed on the estimated frontier, while a value such that $\hat{\theta}(x_j, y_j) < 1$ means that the firm is inefficient.

We conclude with three remarks that can now be better understood.

As animal and crop outputs cannot be easily substituted, we must not aggregate them, instead, we consider the calculation of efficiency as a two-dimensional output problem.

As indicated above, including SP representing the costs paid for not producing in a subsidized manner can also be understood as including direct payments as negative inputs. There are two reasons why we prefer not to consider them as a third output: on the one hand, the production factors considered here do not (directly) produce *DP*, so there is no reasonable argument for allowing them to appear on the left-hand side of the production function; on the other hand, many farms in Spain have zero subsidies and would thus form a noninterpretable hyper-plane in the DEA. Nevertheless we admit that handling *DP* in the model input subsidies or output-related subsidies is a crucial point; both approaches could be applied. The "Global Trade Analysis Project" (known as GTAP), for example, handles *DP* as subsidies on inputs. The main part of *DP* in the EU is linked to land (even headage premia on beef cattle are linked via livestock density restrictions). Therefore it seems worthwhile to handle *DP* as input subsidies. For the other option (handling *DP* as output), *DP* level could increase for a constant level on inputs (e.g., land), which is contradictory to CAP regulations.

Finally, there is the issue whether more input variables should be aggregated for a nonparametric analysis such as DEA for example to obtain stronger results with respect to larger differences in the resulting efficiency index. However, this question is nothing more than a discussion of the bias—variance trade off dilemma: aggregation leads to more bias but less variance and vice versa. We have decided here to opt for high resolution, that is high variance, small bias. So, since we always conduct nonparametric analysis with high-resolution level, none of our results will suffer errors due to possible misspecification.

3.2. *Regression analysis*

Next, we are interested in estimating model (2). As mentioned above we do not want to assume any particular functional form for $g(\cdot, \cdot)$ except that it is a smooth function, that is, it has continuous second derivatives.

We will now present a brief overview of the algorithms of nonparametric flexible function regression. In particular we explain the estimation of the parameters β and their asymptotic covariance, as well as the estimation of the nonparametric function of $g(\cdot, \cdot)$ in a semiparametric model of the form as described in Eq. (2). We assume $E[e|EF, ESU, R] = 0$, Var $[e] < \infty$. The estimation of $g(\cdot)$ and β involves two steps: first the estimation of β and its covariance using the method of Robinson (1988), and afterward the estimation of $g(\cdot, \cdot)$ using local linear smoothing by Ruppert and Wand (1994). For a more detailed introduction to non and semiparametric modeling see also Härdle et al. (2004).

The basic idea is to construct an estimator that creates a smooth surface (or hyperplane), that is, in the one-dimensional case a smooth line, into the point cloud that presents its functional form. The smoothness of that surface can be (pre) determined by choosing a respectively large *smoothing parameter* h, called bandwidth. h can often also be data driven.

First, it is important to understand that this estimator works locally, for example we estimate the desired function, the hyperplane, separately at each point we are interested in. Therefore we need to introduce some additional notation. Consider a regression problem of the form $E[Y|X = x_0[= G(x_0)$, $Y \in \Re$, $X, x \in \Re^d$ with $G(\cdot)$: $\Re^d \to \Re$ being an unknown smoothing function. We aim to estimate $G(x_0)$ for some point $x_0 \in \Re^d$. Having $\{X_i, Y_i\}_{i=1}^n$ observed, this can be done by local least squares Q5

$$\begin{pmatrix} \hat{G}(x_0) \\ \widehat{\nabla G}(x_0) \end{pmatrix} = \operatorname{argmin}_{a_0, a_1} \sum_{i=1}^{n} \left\{Y_i - a_0 - a_1^T(X_i - x_0)\right\}^2$$
$$\times K_h(X_i - x_0), \tag{12}$$

$a_0 \in \Re$, $a_1 \in \Re^d$ and $\nabla G(\cdot)$ being the gradient of $G(\cdot)$. Further, $K_h(v) = \prod_{j=1}^{d} \frac{1}{h} K(\frac{v_j}{h})$ is a $\Re^d \to \Re$ weight function. In our calculations we chose $K(v) = \frac{15}{16}(1 - v^2)^2 \mathbb{1}\{|v| \leq 1\}$. So we used a weighted least squares estimator for linear regression, which becomes a local (linear) estimator due to the weights K_h giving a lot of weight to points (X_i, Y_i) where X_i is close to x_0 but zero weights to points far from x_0. Consistency, asymptotic theory and properties are well known and studied for the multivariate case in Ruppert and Wand (1994), for a general introduction see Fan and Gijbels (1996).

If we eliminate the vector a_1 in Eq. (12) and thus maximize only over a_0, the minimizing argument is a local constant estimator of $G(x_0)$. In this case it is easy to derive the explicit formula

$$\tilde{G}(x_0) = \frac{\sum_{i=1}^{n} K_h(X_i - x_0)Y_i}{\sum_{i=1}^{n} K_h(X_i - x_0)}. \tag{13}$$

As one can see, in the weighting function, the smoothing parameter h comes in: the larger the h, and consequently the environment with positive weighting, the smoother the resulting hyperplane, that is $h \to \infty$ gives a linear function for G whereas $h = 0$ yields a G being the interpolation of the Y_i's. In a context like ours, the choice of the smoothing parameter should be considered in the same way as choice of degrees of freedom, tha is the empirical researcher can allow for more flexibility or impose more smoothness on the function. To allow for high flexibility without increasing the variance to unreasonable levels, we chose smoothing parameters that do not restrict the functional forms unless the plotted surface becomes wiggly.

Returning to our model (2), we will apply the local linear estimation method, that is Eq. (12), on $\{W_i{:=}(\ln(EF_i), \ln(ESU_i)), (E_i - \hat{\beta}R_i)\}_{i=1}^{n}$. The remaining question is how to obtain $\hat{\beta}$. The estimator of β is defined as

$$\hat{\beta} = S^{-1}_{R-\tilde{R},R-\tilde{R}} S_{R-\tilde{R},E-\tilde{E}} \tag{14}$$

where for any matrix or vector sequences R_i, B_i we set $S_{R,B} = \frac{1}{n}\sum_{i=1}^{n} R_i B_i^T$ and $\tilde{R}_i = \widehat{E}[R_i|W_i]$, $\tilde{B}_i = \widehat{E}[B_i|W_i]$ with B_i being either R_i or E_i. We estimate the conditional expectations ($\widehat{E}$) via a local constant smoother as defined in the Eq. (13). It is easy to see that the variance of $\hat{\beta}$ can be estimated by $\hat{\sigma}^2 S^{-1}_{R-\tilde{R},R-\tilde{R}}$ where $\hat{\sigma}^2$ is a consistent estimator of the conditional variance of E: $\sigma^2 = \text{var}\,[E|W_i, R_i]$. For more details see Robinson (1988). Note that all these models have been proven to work perfectly for dependent data as well. It is worthwhile to mention this as the indices calculated by DEA are not independent.

Certainly, as $g(\cdot,\cdot)$ is nonparametric, we could have directly used the covariates *ESU* and *EF* in the model (2). As mentioned above, the logarithm therefore does not impose any model specification here. The problem is that both variables have rather skewed distributions with many data-sparse areas. In contrast, ln (*ESU*) as well as ln (*EF*) look quite normal around the mode with rather short tails at the end. It is thus only for the sake of reasonable behavior of our smoothing techniques that we prefer to apply our smoothing methods to the log-transformed data, see also Biedermann and Dette (2003) for more details.

4. Empirical results

All above-mentioned calculations have been realized separately for Spain for 1999–2000 for cattle, pig, sheep and goat farms; and Germany for 1999–2000 for cattle and pig farms. The presentation and discussion of results follows the reasoning and list of questions given in the first section. This includes: estimation of our models in Eq. (2) to analyze the impact of environmentally friendly behavior, and farm size on efficiency for 1999 and 2000; comparison of regression results based on E_W with those based on E_{DP}; calculating the correlations between subsidies and other factors such as farm economic size, environmentally friendly behavior and economic efficiency; where appropriate, we use and compare results of both calculations for further conclusions.

As the calculation of efficiency alone is not of interest in this article, but only as an auxiliary step, we neither explicitly present nor discuss the results of the DEA calculations here but have deferred them to the Appendix. Instead, we start directly with the analysis of the impact of environmentally friendly behavior and farm size on efficiency. That is, we focus on the regression problem of equation

$$E = g[\ln(EF), \ln(ESU)] + \beta^T R + e, \tag{14}$$

for E being E_{DP} (*DP* included in model) as well as for E being E_W (*DP* not included in model). By comparing efficiency calculated with direct payments included (E_{DP}) and the resulting efficiency when subsidies are ignored (E_W), we check if and how the CAP distorts efficiency. When we speak of significance in the following, we always refer to the 10% significance level.

In non and semiparametric regression, the choice of smoothness controlled via the bandwidth (h), is often either not discussed or quite polemic. Therefore, we tried out several bandwidths and present here the results for those where the estimated surface starts to become smooth. In practice, for two dimensions and smooth densities as we have in this application, this provides a reasonable tradeoff between bias and variance of the estimates. For the parametric part β of model (2) it should be emphasized that the results for the (semi) parametric estimation of the regional dummies turned out to be quite robust with respect to the bandwidth choice for the nonparametric part. This is expected if, for example, the regional dummies are almost uncorrelated with the other covariates ln (*ESU*) and ln (*EF*).

First let us make some remarks on the results concerning the regional dummies, that is, on $\hat{\beta}$, summarized in Table 3 for Spain in 1999 and 2000. We divided Spain into five regions: North, Center, Andalusia, Ebro (along the Ebro river), and Levante. The last one has been used as a normalizing region. Note that Andalusia could be replaced by "South," and Levante by "East". Ebro represents the northeastern Spanish region including the northeastern Mediterranean coast and the Ebro river valley with a mainly Mediterranean climate that traditionally has been considered as a homogenous agricultural region.

Surprisingly, the North and Ebro seem to be less efficient. However, these results are only significant for cattle farming, whereas the Ebro is only significantly less efficient than the other regions when considering sheep and goats. For cattle farming, Levante seems to be best though not significantly better than the center and the south. These findings hardly change as a function of the dependent variable (E_{DP} or E_W).

In Table 4 the corresponding results for Germany are given, also for 1999 and 2000, but without sheep and goat farms. We

Table 3
Estimates (upper lines) with standard error (lower lines) for the regional dummies "β" in model (2) for Spain

Year	Dep. var.	North	Center	Andalusia	Ebro
		Cattle			
1999	E_{DP}	−0.4617	−0.2189	−0.3761	−0.4907
		0.10397	0.10400	0.11497	0.10549
	E_W	−0.4672	−0.2537	−0.3298	−0.5062
		0.10214	0.10217	0.11295	0.10364
2000	E_{DP}	−0.4192	−0.1790	−0.3592	−0.4065
		0.08852	0.08943	0.10014	0.09058
	E_W	−0.4233	−0.2001	−0.3419	−0.4030
		0.08752	0.08842	0.09901	0.08955
		Pig			
1999	E_{DP}	−0.3459	−0.0607	0.08320	−0.0926
		0.15810	0.04511	0.08340	0.03744
	E_W	−0.2148	−0.0158	0.12308	−0.0410
		0.15722	0.04486	0.08294	0.03723
2000	E_{DP}	−0.0329	−0.0160	0.20771	−0.0460
		0.11033	0.05789	0.09105	0.04707
	E_W	−0.0948	−0.0195	0.05673	−0.0431
		0.10395	0.05454	0.08578	0.04434
		Sheep and Goats			
1999	E_{DP}	−0.1201	−0.0228	−0.0901	−0.1380
		0.06367	0.04137	0.05517	0.04526
	E_W	−0.1459	−0.0326	−0.0989	−0.1851
		0.05864	0.03811	0.05081	0.04168
2000	E_{DP}	−0.1049	−0.0726	−0.2321	−0.2156
		0.05082	0.03033	0.03679	0.03437
	E_W	−0.0969	−0.0667	−0.2181	−0.2632
		0.04845	0.02891	0.03506	0.03276

Table 4
Estimates (upper lines) with standard error (lower lines) for the regional dummies "β" in model (2) for Germany

Year	Dep. var.	North	Center	South
		Cattle		
1999	E_{DP}	−0.0279	0.02466	0.07394
		0.02255	0.02271	0.02350
	E_W	−0.0261	0.03302	0.08543
		0.02310	0.02327	0.02407
2000	E_{DP}	0.01820	0.05785	0.08961
		0.02208	0.02235	0.02322
	E_W	0.02153	0.06496	0.09295
		0.02199	0.02226	0.02313
		Pig		
1999	E_{DP}	−0.0787	−0.0657	−0.0272
		0.05942	0.06424	0.06067
	E_W	−0.1328	−0.1073	−0.0860
		0.05821	0.06294	0.05944
2000	E_{DP}	−0.1331	−0.0989	−0.0765
		0.06012	0.06545	0.06171
	E_W	−0.2057	−0.1580	−0.1574
		0.05955	0.06483	0.06113

divided Germany in to four regions: North, Center, South, and East. As in Spain, the latter has been used as a normalizing region.

In Germany, the South is the most efficient in cattle farming, but it is hard to say anything about differences in the rest of the country. It may be that central Germany is more efficient than the East and North, but this is not significant for 2000. In pig farming, the most efficient farms are the (quite large) farms in Eastern Germany. However, there are significant changes between years and dependent variables, E_{DP} or E_W.

Interesting is the outcome of the impact estimates for environmentally friendly behavior (EF) and farm size (ESU) on (conventional) efficiency. As the functional form of $g(\cdot, \cdot)$ in model (2) is nonparametric, the results are presented graphically, see Fig. 1–10. Here, only the graphs for 2000 are shown; the graphs for 1999 are available on www.uc3m.es/uc3m/dpto/CJM/webmonnet.html or upon request.

In all the graphs shown, the outer 2% boundaries are cut off (i.e. not plotted) to avoid interpreting the so called boundary effects typical in nonparametric estimation. Since the $g(\cdot, \cdot)$ function is an unknown function from $\Re^2$ to $\Re$, it is presented via two graph: a three-dimensional and a two-dimensional graphs. The two-dimensional graph shows three functions representing the three slices of the full (i.e., three) dimensional plot that describe the impact of the environmentally friendly behavior proxy ($\ln EF$) on efficiency when farm size ($\ln ESU$) is fixed at: the median (solid line), the upper 95% quantile, i.e., large farms (dotted line), and at the lower 5% quantile, i.e., small farms (dashed line).

Our results show that:

In general one can say that the distortion of efficiency caused by direct subsidies is visible in our plots for Spain and Germany, especially for pig farms (see below).

The less environmentally friendly farms are generally more efficient under the actual price structure. This holds regardless of whether direct subsidies are included in the calculation of efficiency. Generally, efficiency (for both E_{DP} and E_W) increases when livestock units per agricultural utilized area increase (under the actual input and output price structure). Environmentally friendly pig farms are the exception. When we include direct payments (i.e., consider E_{DP}) they indeed show a high level of efficiency. These findings are especially strong in the medium and small holdings, particularly in Germany, but can also be found in Spain.

Positive externalities justify the introduction of subsidies to reward nature conservation. The target of helping to preserve the environment and increase efficiency was introduced by the MacSharry reform of the CAP by linking subsidies to environmentally friendly farms. This reform of the CAP (in force since 1993), yields some visible effects, especially in pig farming.

The efficiency level of the more environmentally friendly pig farms is much higher than those of the conventional hog holdings, that is intensive ones, when focusing on E_{DP}, again, especially in Germany. It is rather interesting to note that they lose their efficiency advantage compared with the less environmentally adapted farms when we recalculate efficiency under the counterfactual hypothesis of no direct payments. In other

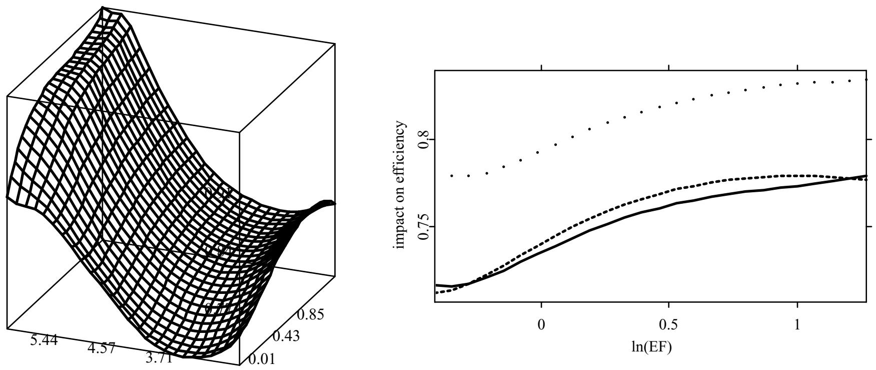

Fig. 1. German cattle farms 2000, E_W. Left: $g_W(\ln(EF), \ln(ESU))$ from Eq. (2), where the lower left (right) axis is the impact of $\ln(EF)$ ($\ln(ESU)$). Right: g_W with ESU fixed at median (solid), 95% quantile that is, large farms (dotted line), and 5% quantile that is, small farms (dashed line). The outer 2% are cut off in both figures.

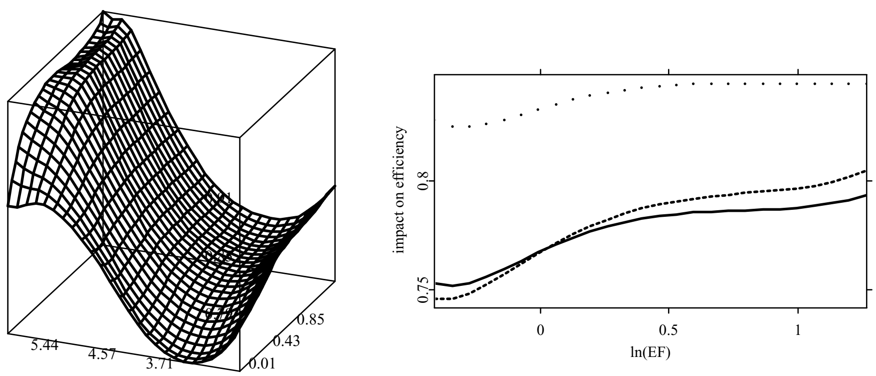

Fig. 2. German cattle farms 2000, E_{DP}. Left: $g_{DP}(\ln(EF), \ln(ESU))$ from Eq. (2), where the lower left (right) axis is impact of $\ln(EF)$ ($\ln(ESU)$). Right: g_{DP} with ESU fixed at median (solid), 95% quantile that is, large farms (dotted line), and 5% quantile that is, small farms (dashed line). The outer 2% are cut off in both figures.

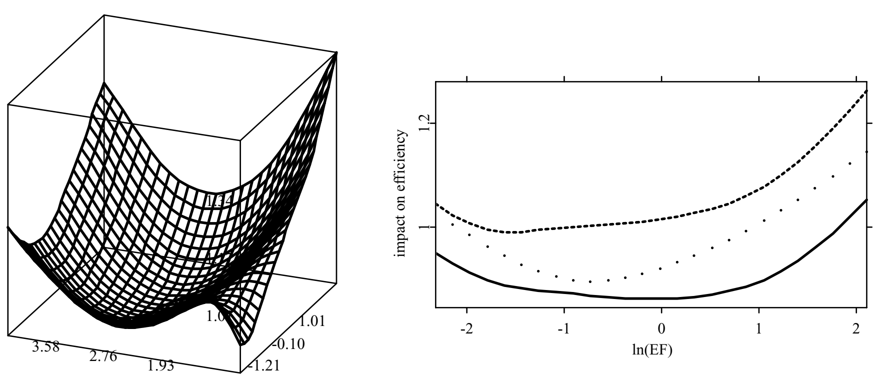

Fig. 3. Spanish cattle farms 2000, E_W. Left: $g_W(\ln(EF), \ln(ESU))$ from Eq. (2), where the lower left (right) axis is impact of $\ln(EF)$ ($\ln(ESU)$). Right: g_W with ESU fixed at median (solid), 95% quantile that is, large farms (dotted line), and 5% quantile that is, small farms (dashed line). The outer 2% are cut off in both figures.

words, it is fairly clear that the most environment-friendly pig farms are efficient thanks to direct payments under the CAP.

Conversely, in cattle farming, the conventional holdings are more efficient than those most environmentally friendly. In this sector, the results hold true with and without direct payments, i.e. independently of whether we consider E_{DP} or E_W. This also holds for both countries, Germany and Spain.

However, in Germany, the discrepancy in efficiency between intensive and extensive animal farming becomes less important when subsidies are taken into account, whereas in Spain we cannot find any effect of the subsidies with respect to environmental friendliness in the cattle sector. Our results seem to indicate that to some extent (and certainly in Germany) there is a positive impact of subsidies on environmental friendliness. Moreover, we found empirical evidence that the current direct payments system is "environmentally oriented" in all types of farms studied in Germany but only for pig farming in Spain, (see Tables 5–9). Taken together, this information offers the conclusion that the

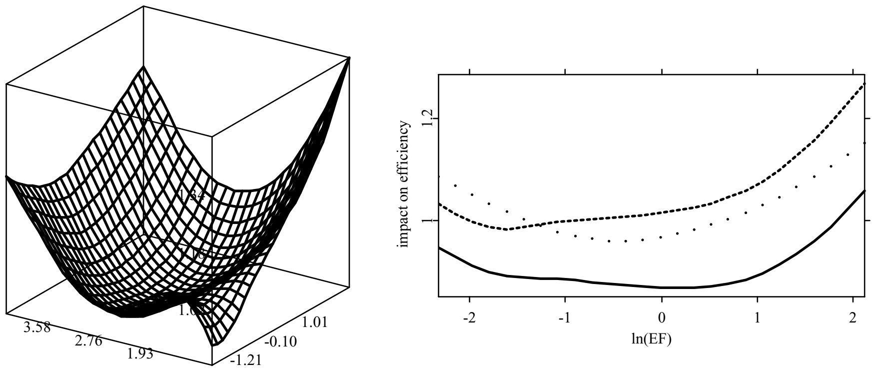

Fig. 4. Spanish cattle farms 2000, E_{DP}. Left: $g_{DP}(\ln(EF), \ln(ESU))$ from Eq. (2), where the lower left (right) axis is impact of $\ln(EF)$ ($\ln(ESU)$). Right: g_{DP} with ESU fixed at median (solid), 95% quantile that is, large farms (dotted line), and 5% quantile that is, small farms (dashed line). The outer 2% are cut off in both figures.

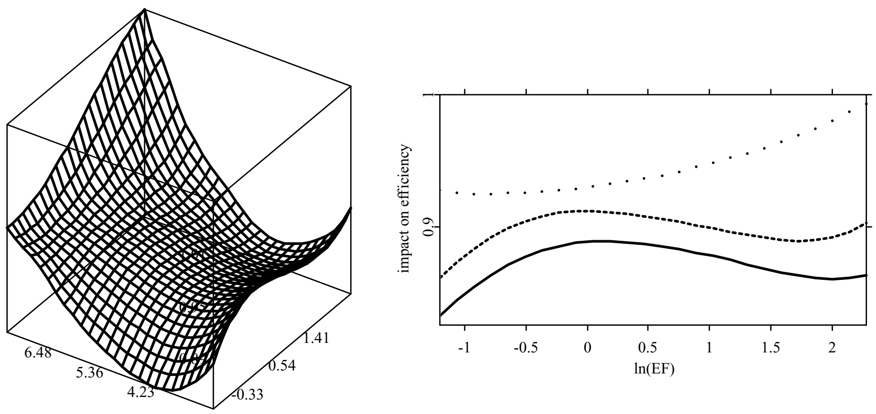

Fig. 5. German pig farms 2000, E_W. Left: $g_W(\ln(EF), \ln(ESU))$ from Eq. (2), where the lower left (right) axis is impact of $\ln(EF)$ ($\ln(ESU)$). Right: g_W with ESU fixed at median (solid), 95% quantile that is, large farms (dotted line), and 5% quantile that is, small farms (dashed line). The outer 2% are cut off in both figures.

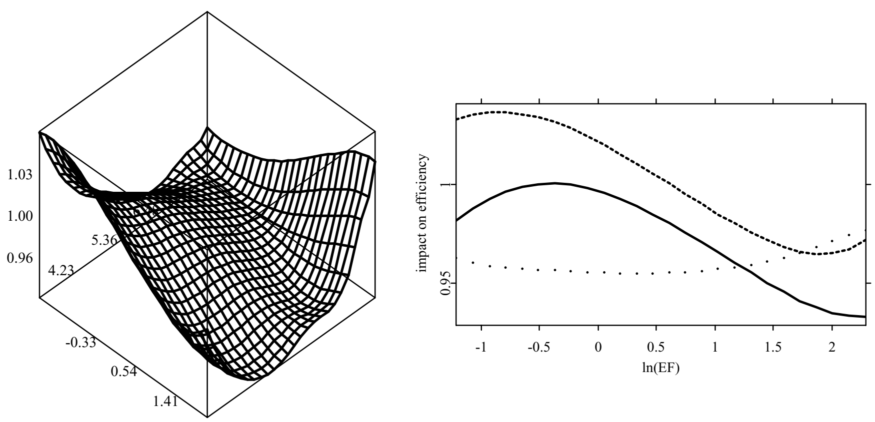

Fig. 6. German pig farms 2000, E_{DP}. Left: $g_{DP}(\ln(EF), \ln(ESU))$ from Eq. (2), where the lower left (right) axis is impact of $\ln(EF)$ ($\ln(ESU)$). Right: g_{DP} with ESU fixed at median (solid), 95% quantile that is, large farms (dotted line), and 5% quantile that is, small farms (dashed line). The outer 2% are cut off in both figures.

present direct subsidy structure plays a significant role in helping farms to preserve the natural environment under competitive conditions. Additionally, the new CAP MTR could improve the efficiency of environmentally friendly farms by increasing the incentives for agrienvironmental measures.

To further investigate this point we directly calculated different correlations between subsidies and *EF* and tested them for significance, see Tables 5 to 9. These results show a significant positive correlation between the direct subsidies and the proxy of adaptation to the natural environment, that is direct payment

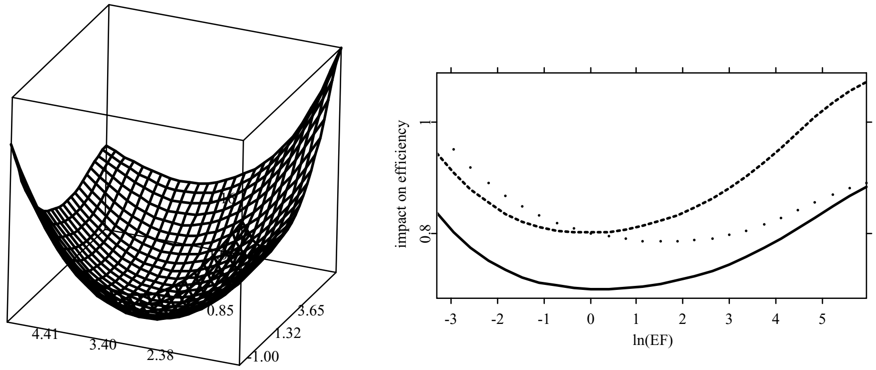

Fig. 7. Spanish pig farms 2000, E_W. Left: $g_W(\ln(EF), \ln(ESU))$ from Eq. (2), where the lower left (right) axis is impact of $\ln(EF)$ ($\ln(ESU)$). Right: g_W with ESU fixed at median (solid), 95% quantile that is, large farms (dotted line), and 5% quantile that is, small farms (dashed line). The outer 2% are cut off in both figures.

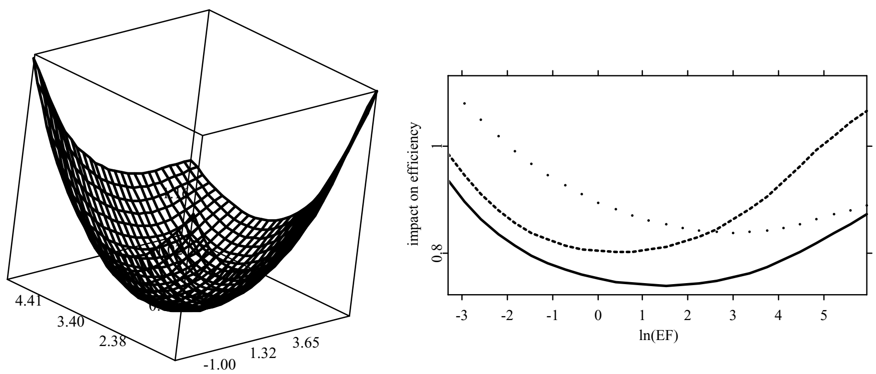

Fig. 8. Spanish pig farms 2000, E_{DP}. Left: $g_{DP}(\ln(EF), \ln(ESU))$ from Eq. (2), where the lower left (right) axis is impact of $\ln(EF)$ ($\ln(ESU)$). Right: g_{DP} with ESU fixed at median (solid), 95% quantile that is, large farms (dotted line), and 5% quantile that is, small farms (dashed line). The outer 2% are cut off in both figures.

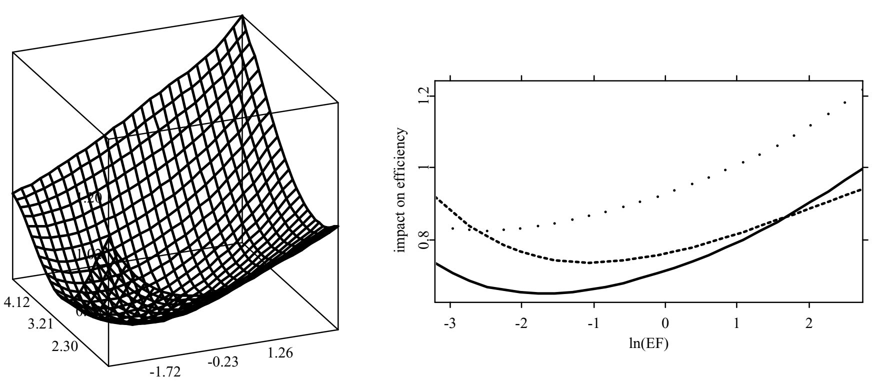

Fig. 9. Spanish sheep and goat farms 2000, E_W. Left: $g_W(\ln(EF), \ln(ESU))$ from Eq. (2), where the lower left (right) axis is impact of $\ln(EF)$ ($\ln(ESU)$). Right: g_W with ESU fixed at median (solid), 95% quantile that is, large farms (dotted line), and 5% quantile that is, small farms (dashed line). The outer 2% are cut off in both figures.

negatively correlated with LU/UAA, for both countries and years over all models considered.

All figures indicate clearly that economic size matters for efficiency performance. Usually, one would expect the larger farms to outperform the smaller farms, but we found several exceptions. Particularly, the smaller pig farms perform rather well and are also competitive.

The counterfactual exercise (looking at E_W) without subsides supports these results both in Germany and Spain. We cannot make such a clear statement for cattle farming. Nevertheless, in Spain the small cattle farms are above the mean efficiency index. The counterfactual study without direct subsidies upholds these results. This finding is interesting since it fits perfectly with the conclusions of Älvarez and Arias (2003), who point

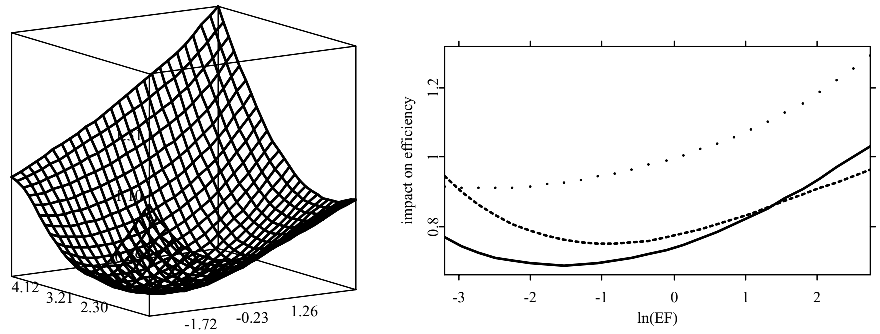

Fig. 10. Spanish sheep and goat farms 2000, E_{DP}. Left: $g_{DP}(\ln(EF), \ln(ESU))$ from Eq. (2), where the lower left (right) axis is impact of $\ln(EF)$ ($\ln(ESU)$). Right: g_{DP} with ESU fixed at median (solid), 95% quantile that is, large farms (dotted line), and 5% quantile that is, small farms (dashed line). The outer 2% are cut off in both figures.

out that increasing cattle farm size while holding managerial ability constant can be an important source of diseconomies of size.

In general the biggest farms reach the highest levels of efficiency for all animal types. This result also holds true in the counterfactual exercise, except for German pig farming. It is important to add here that the size impact on efficiency often interacts with the environmental behavior, and that this impact (i.e., the one of $\ln(EF)$) on efficiency is often much stronger than the size effect. The counterfactual does not seems to have an impact on the efficiency of small farms versus the medium-sized farms.

German pig farming is an interesting case because the small, environmentally friendly farms are more efficient than the biggest farm when we account for direct payments. But this is not the case in either German cattle farming or for any type of animal farm in Spain. There, direct payments do not seem to positively discriminate (helping to raise relative efficiency) by size and adaptation to natural environment. Again, especially for pig farming in Spain, small intensive farms seem to be above the mean efficiency index and above the biggest farms. The counterfactual, ignoring the subsidies when estimating efficiency, upholds these results as well. This was a remarkable difference from the results for German pig farming, where the direct payments made the environmentally friendly farms more efficient in comparison with the more conventional farms (intensive). In Germany, the direct payment unambiguously increases the efficiency of environmentally friendly farms when compared with conventional (intensive) farms, which is not the case in Spain.

We finally address the question of which factors the subsidies are correlated with and how strongly. As we are interested in both absolute and relative mean increases, we calculated all the following correlations and their $p-$values

$$corr(DP, EF), corr(DP, \ln(EF)), corr\ (\ln(1+DP),$$
$$\ln(EF)), corr(\ln(1+DP), EF)corr(DP, ESU),$$

Table 5
Correlations (upper lines) with P-values for significance (lower lines) for Spain

Year	Cattle				
	Variables	$(\cdot),(\cdot)$	$(\cdot),\ln(\cdot)$	$\ln(1+\cdot),(\cdot)$	$\ln(1+\cdot),\ln(\cdot)$
1999	DP, EF	−0.1267	−0.4134	−0.3864	−0.1854
		0.0000	0.0000	0.0000	0.0000
	DP, ESU	0.6567	0.4884	0.1369	0.2255
		0.0000	0.0000	0.0000	0.0000
	DP, E_W	0.0108	–	0.1099	–
		0.6827	–	0.0000	–
	DP, E_{DP}	0.2007	–	0.1592	–
		0.0000	–	0.0000	–
2000	DP, EF	−0.1808	−0.4311	−0.3264	−0.1956
		0.0000	0.0000	0.0000	0.0000
	DP, ESU	0.6163	0.4573	0.1061	0.1644
		0.0000	0.0000	0.0000	0.0000
	DP, E_W	0.1351	–	0.0897	–
		0.0000	–	0.0004	–
	DP, E_{DP}	0.2755	–	0.1330	–
		0.0000	–	0.0000	–

Table 6
Correlations (upper lines) with p-values for significance (lower lines) for Spain

Year	Pig farming				
	Variables	$(\cdot),(\cdot)$	$(\cdot),\ln(\cdot)$	$\ln(1+\cdot),(\cdot)$	$\ln(1+\cdot),\ln(\cdot)$
1999	DP, EF	−0.0893	−0.2574	−0.4323	−0.1306
		0.1549	0.0000	0.0000	0.0372
	DP, ESU	0.4885	0.3425	0.1834	0.1188
		0.0000	0.0000	0.0033	0.0581
	DP, E_W	−0529	–	−3088	–
		0.4005	–	0.0000	–
	DP, E_{DP}	0.1252	–	−1852	–
		0.0458	–	0.0030	–
2000	DP, EF	−0.1427	−0.4211	−0.4718	−0.3836
		0.0243	0.0000	0.0000	0.0000
	DP, ESU	0.1682	0.2106	0.0464	0.0284
		0.0078	0.0008	0.4661	0.6553
	DP, E_W	−0.1462	–	−0.2370	–
		0.0210	–	0.0002	–
	DP, E_{DP}	0.0580	–	−1355	–
		0.3622	–	0.0326	–

Table 7
Correlations (upper lines) with p-values for significance (lower lines) for Spain

Year	Sheep and goats				
	Variables	$(\cdot),(\cdot)$	$(\cdot), \ln(\cdot)$	$\ln(1+\cdot),(\cdot)$	$\ln(1+\cdot), \ln(\cdot)$
1999	DP, EF	−0.0817	−0.2722	−0.2705	−0.1269
		0.0548	0.0000	0.0000	0.0028
	DP, ESU	0.9683	0.6143	0.8555	0.5959
		0.0000	0.0000	0.0000	0.0000
	DP, E_W	0.0480	–	−0.2046	–
		0.2594	–	0.0000	–
	DP, E_{DP}	0.1136	–	−0.0638	–
		0.0075	–	0.1341	–
2000	DP, EF	−0.0862	−0.2845	−0.2750	−0.1298
		0.0247	0.0000	0.0000	0.0007
	DP, ESU	0.9731	0.6248	0.8297	0.5706
		0.0000	0.0000	0.0000	0.0000
	DP, E_W	0.0917	–	−0.0956	–
		0.0168	–	0.0127	–
	DP, E_{DP}	0.1516	–	0.0241	–
		0.0000	–	0.5302	–

$$corr(DP, \ln(ESU)), corr(\ln(1+DP), \ln(ESU)),$$
$$corr(\ln(1+DP), ESU) corr(DP, E_W),$$
$$corr(\ln(1+DP), E_W), corr(DP, E_{DP}),$$
$$corr(\ln(1+DP), E_{DP})$$

It is clear that when efficiency is calculated including DP, E_{DP} will be (positively) correlated with it; this result can change when looking at $\ln(1+DP)$. Furthermore, we expect that the

Table 8
Correlations (upper lines) with p-values for significance (lower lines) for Germany

Year	Cattle				
	Variables	$(\cdot),(\cdot)$	$(\cdot), \ln(\cdot)$	$\ln(1+\cdot),(\cdot)$	$\ln(1+\cdot), \ln(\cdot)$
1999	DP, EF	−0.0914	−0.1102	−0.1686	−0.1432
		0.0248	0.0067	0.0000	0.0004
	DP, ESU	0.9705	0.6109	0.6085	0.4422
		0.0000	0.0000	0.0000	0.0000
	DP, E_W	0.0762	–	−0.1909	–
		0.0612	–	0.0000	–
	DP, E_{DP}	0.1168	–	0.0142	–
		0.0040	–	0.7275	–
2000	DP, EF	−0.0633	−0.0746	−0.1398	−0.1102
		0.1201	0.0671	0.0006	0.0067
	DP, ESU	0.9695	0.5994	0.6128	0.4480
		0.0000	0.0000	0.0000	0.0000
	DP, E_W	0.1002	–	−1075	–
		0.0138	–	0.0082	–
	DP, E_{DP}	0.1200	–	0.0085	–
		0.0031	–	0.8356	–

Table 9
Correlations (upper lines) with p-values for significance (lower lines) for Germany

Year	Pig farming				
	Variables	$(\cdot),(\cdot)$	$(\cdot), \ln(\cdot)$	$\ln(1+\cdot),(\cdot)$	$\ln(1+\cdot), \ln(\cdot)$
1999	DP, EF	−0.2678	−0.3708	−0.4654	−0.4130
		0.0000	0.0000	0.0000	0.0000
	DP, ESU	0.9699	0.8309	0.7181	0.6435
		0.0000	0.0000	0.0000	0.0000
	DP, E_W	0.2451	–	0.0324	–
		0.0000	–	0.5426	–
	DP, E_{DP}	0.2823	–	0.1816	–
		0.0000	–	0.0006	–
2000	DP, EF	−0.2692	−0.3541	−0.4525	−0.4199
		0.0000	0.0000	0.0000	0.0000
	DP, ESU	0.9595	0.8268	0.7232	0.6547
		0.0000	0.0000	0.0000	0.0000
	DP, E_W	0.2796	–	0.0473	–
		0.0000	–	0.3746	–
	DP, E_{DP}	0.2923	–	0.1153	–
		0.0000	–	0.0298	–

subsidies for each year are highly correlated with those of the last year. This means that, when calculating E_W, one has certainly not eliminated the effect of all subsidies paid to this farm, but only ignored the cash received this particular year. So, in E_W the long-term effect of (formerly received or not received) direct payments is still reflected. For this reason it is clear that we are more interested in $corr(DP, E_W)$ and $corr(\ln(1+DP), E_W)$ than in $corr(DP, E_{DP})$ and $corr(\ln(1+DP), E_{DP})$.

The results for Spain for 1999 and 2000 are presented in Tables 5–7.

First, let us briefly summarize the signs in the tables: In both years, 1999 and 2000, signs are always negative for any correlation considered between DP and EF for Spain as well as for Germany. This means that, as mentioned above, environmentally friendly behavior is indeed supported financially by the CAP (recall that the smaller the EF the more environmentally friendly the farm is). For both countries and years all correlations considered between DP and ESU are positive. This means that large farms generally get more financial support than small farms. Looking at the pair $corr(DP, E_W)$, $corr(\ln(1+DP), E_W)$ we get the following pattern for both years

	cattle		pig		Sheep and goats	
Spain	+	+	–	–	+	–
Germany	+	–	+	0		

Now let us come to a more detailed analysis of the results.

Looking only at the absolute values of the calculated correlation coefficients, it seems evident (compare the p-values of the significance tests) that the distortion of efficiency caused by direct subsidies is significant in Spain and Germany. Recall, however, that here we are not correcting for endogeneity, which is why we had to do the regression analysis with the counterfactual exercise.

Direct payment correlation with farm size shows the level of real *inverse modulation* of the actual CAP subsidies. In Germany and Spain we found a clear positive correlation between subsidies and farm size. Our results are not surprising in the sense that some *DP* are directly related to "size," For example payments per animal. However, they are definitely surprising when we compare the official political intention with the high level of correlation. Often, the correlation between *DP* and size is higher than 90%, that is, subsidies can mainly be explained by farm size. Moreover, results are independent of farm type, country, or the year in which they are tested. These results are congruent with the generally accepted hypothesis that all direct subsidies are (indirectly) linked mostly to output level and size.

Looking at the correlations between E_{DP} and DP, the results are, as expected, all positive or zero except $corr(\ln(1 + DP), E_{DP})$ for pig farming in Spain. There is no doubt that when direct payment enters as positive output or negative input of a farm, then farms obtaining those payments seem to be efficient. Looking at the p-values this is significant for most cases in both countries and years. The hypothesis that subsidies increase economic efficiency has to be examined by comparing the real efficiency DEA index with counterfactual efficiencies under the hypothesis of not having received subsidies in a given year.

Then, in this more interesting counterfactual exercise, our results show that the level of efficiency on average increases with the units of direct payments as well (again with the exception of Spanish pig farms). However, looking at a percentage increase of direct payment, that is at $corr(\ln(1 + DP), E_W)$, efficiency decreases or stagnates for all years and countries except for Spanish cattle farms. In other words, focusing on efficiency, a policy that grants subsidies per farm (decoupled from size) seems much more reasonable than the current, counterproductive policy of giving subsidies mainly based on size. This is also one of the key points in the recommendation list of Bertola et al. (2002), reinforced by our empirical results. In fact a subsidy per agricultural worker is less discriminatory than the current system (see Mora and San Juan, 2004).

As indicated above, when we interpreted the graph's outcomes, the results show significant positive correlation between the direct subsidies and the proxy of adaptation to the natural environment, for example direct payments correlated negatively with LU/UAA. That is, we found an overall significant decreasing level of direct payments as the livestock units per agricultural utilized area increased in all farm types. So the subsidy policy does take environmental friendliness into account. However, it is evident from the tables that this correlation is much weaker, almost negligible in fact, compared to the overwhelmingly strong correlation between subsidies and farm size.

5. Conclusions

The main empirical conclusions are the following:

Looking at conventional efficiency (E_W), large farms with intensive holdings are generally more efficient. Direct payments are a potential source of efficiency distortions. In fact, the amount of direct payment growth after the CAP reform of 1992 significantly affects the relative level of efficiency E_{DP}. The results show a positive correlation between subsidies and efficiency (both E_W and E_{DP}) when looking at the absolute amounts. However, the mean efficiency decreases or stagnates as the percentage of direct payments rises. This holds for all type of farms, years, and countries analyzed except for Spanish cattle farms. This means that a combination of direct subsidies and size is counterproductive. Unfortunately, our results show clearly that this is precisely what the present subsidy policy is doing: farm size explains around 90% (or more) of the direct payments.

The strong subsidization of large farms cannot be justified by their presumably high efficiency or environmental friendliness. Neither our graphical nor our numerical results confirm such a hypothesis.

The direct subsidies have been justified as a reward for the positive externalities that the agricultural activities generate. We have found that the current direct payment system is not sufficient to correct the fact that the less environmentally friendly farms are the most efficient farms. The only exception we found was German pig farming, in which the efficiency of the most environmentally friendly farms normally rises strongly with direct payments versus the counterfactual exercise. This effect is more important in Germany, where conventional farms become less efficient than the environmentally friendly farms after receiving direct payments. In Spain, the most conventional (intensive) Spanish hog holdings reach efficiency levels "similar" to the environmentally friendly farms when accounting for direct payments (or vice versa). But the smallest and less environmentally adapted farms are most efficient. Given that the CAP regulations are common for the whole EU, we suspect that the differences in the environmental standards and their enforcements between Member States have lead finally to the observed differences between Germany and Spain.

Nevertheless, together with the estimated correlations, we conclude that there is some empirical evidence that the actual direct payments system is "environmentally oriented" for all types of farms studied (levels of direct payments decrease as the livestock units per agricultural utilized area increases). Thus, our calculations show that the current subsidy schedule plays a significant role in helping farmers to conserve the natural environment, even though we have seen throughout our data that while this policy succeeds especially in Germany, it is still not sufficient to motivate farmers to change their production toward a more environmental friendly one.

All this should encourage a look at what the EU is doing to change the situation (the cross-compliance provisions of the 2003 CAP reform) in the future.

Acknowledgments

We would like to thank Alois Kneip, Antonio Álvarez Pinilla, and two anonymous referees for their helpful discussion. This research was supported by the Spanish "Dirección General de Investigación del Ministerio de Ciencia y Tecnología", number SEJ2005-08269/ECON, SEJ2004-04583/ECON, and by FUNCAS.

Appendix

All histograms shown here refer to the Spanish data. We first give the histogram plots for 1999, E_{DP} and E_W, followed by those for 2000. Note that these indices are calculated based on an input-orientated DEA. Due to the high level of disaggregation, in most cases we have the mode at 1 but little density close to 1. A further aggregation of inputs would make the histograms flatter on the right tail. Obviously, there are only marginal differences between 1999 and 2000. Finally, we give histograms for 1999 and 2000 of the differences in E_{DP} when calculated by input-orientated DEA minus those calculated by output orientated DEA ("orientation differences for E_{DP}"). As expected, the differences have by far the highest density close to zero. Together with the former plots (E_{DP} in 1999 and 2000) we can conclude that these small changes cannot affect the overall conclusions of our analysis.

References

Älvarez Pinilla, A., 2001. Concepto y mediciön de la eficiencia productiva. In: Älvarez Pinilla, A. (Ed.) La mediciön de la eficiencia y la productividad, Pirämide, Madrid, pp. 19–40.

Älvarez Pinilla, A., Arias, C., 2003. Diseconomies of size with fixed managerial ability, Am. J. Agric. Econ. 85(1), 134–142.

Ball, V. E., Butault, J. P., San Juan, C., 2004. Measuring real capital input in
Q6 OECD Agriculture, Can. J. Agri. Economics, forthcoming.

Banker, R. D., Morey, R. C., 1986. Efficiency analysis for exogenously fixed inputs and outputs, Oper. Res. 34 (4), 513–521.

Banker, R. D., Charnes, A., Cooper, W. W., 1984. Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis, Manage. Sci. 30 (9), 1078–1092.

Banker, R. D., Cooper, W. W., Seiford, L. M., Zhu, J., 2004. Returns to Scale in DEA. In: Cooper, W.W., Seiford, L.M., and Zhu, J. (Eds.) *Handbook on Data Envelopment Analysis*, Kluwer, London, pp. 41–73.

Bertola, G., Boeri, T., Brücker, H., Coricelli, F., de la Fuente, A., Dolado, J. J., Fitzgerald, J., Garibaldi, P., Hanson, G., Jimeno-Serrano, J. F., Portes, R., Saint-Paul, G., Spilimbergo, A., 2002. Who's afraid of the big enlargement? CEPR Policy Paper 7

Biedermann, S., Dette, H., 2001. Optimal designs for testing the functional form of a regression via non-parametric estimation techniques, Stat. & Prob. Lett. 52, 215–224.

Charnes, A., Cooper, W. W., Rhodes, E., 1978. Measuring the inefficiency of decision making units, Eur. J. Operational Research 2, 429–444.

Cooper, W. W., Seiford, L. M., Tone, K., 2000. Data Envelopment Analysis. A Comprehensive Text with Models, Applications, References and DEA-Solver Software, Kluwer Academic Publishers, London.

Decimavilla, E., San Juan, C., 2002. Encuesta de precios de la tierra 2001. Base 1997. Metadología, Boletin Mensual de Estadística Agraria 11, 7–76.

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

Fan, J., Gijbels, I., 1996. Local polynomial regression, Chapman and Hall, London.

Farrell, M. J., 1957. The Measurement of Productive Efficiency, J. R. Stat. Soc. Ser. A 120, 253–281.

Fischler, F., 2003. The new, reformed agricultural policy. Final press conference after the decision at the Council on agriculture, Luxembourg 26 Jun 2003. SPEECH/03/326.

Førsund, F. R., 1996. On the calculation of the scale elasticity in DEA models, J. Product. Anal. 7, 283–302.

Gonzälez Fidalgo, E., 2001. La estimaciön de la eficiencia con métodos no paramétricos. In: Älvarez Pinilla, A. (Ed.) La mediciön de la eficiencia y la productividad, Pirämide, Madrid, pp. 139–166.

Härdle, W., Müller, M., Sperlich, S., Werwatz, A., 2004. Nonparametric and Semiparametric Models, Springer Verlag, Berlin and Heidelberg.

Kneip, A., Park, B. U., Simar, L., 1998. A note on the convergence of nonparametric DEA estimators for production efficiency scores, Econo. Theory 14, 783–793.

Mckay, L., Lawrence, D., Vlastuin, C., 1983. Profit, output supply, and input demand functions for multiproduct firms: the case of Australian Agriculture, Inter. Economic Re. 24, 323–339.

Mora, R., San Juan, C., 2003. Geographical Specialisation in Spanish Agriculture Before and After Integration in the European Union, Regional Science and Urban Economics, forthcoming Q7

Mora, R., San Juan, C., 2004. Product concentration and farm specialization in Spain after implementation of the CAP and its reform. In: Poppe, K. (Ed.) New Roads for Farm Accounting and FADN, LEI, The Hague, pp. 105–121.

Robinson, P., 1988. Root-N-consistent semiparametric regression. Econometrica 56, 931–954.

Ruppert, D., Wand, M. P., 1994. Multivariate locally weighted least squares regression, Ann. Stat. 22, 1346–1370.

Seiford, L. M., Zhu, J., 2002. Modeling undesirable factors in efficiency evaluation, Eur. J. Oper. Res. 142, 16–20.

Simar, L., Wilson, P., 2000. Statistical Inference in Non–parametric frontier models: the state of the Art, J. Product. Anal. 13, 49–78.

Queries

Q1 Author: Please spell out CAP.
Q2 Author: Please provide JEL classification.
Q3 Author: Please provide the telephone and fax number.
Q4 Author: Please check whether edit in sentence “The counterfractual . . . .” retain the intended meaning.
Q5 Author: The sentence “The basic idea . . . .” not clear. Please check.
Q6 Author: Please update Ball et al. (2004)).
Q7 Author: Please update More and San Juan (2003).