# Efficiency, subsidies and environmental adaptation of animal farming under CAP

Carlos San Juan        Stefan Sperlich
Departamento de Economía
Universidad Carlos III de Madrid,
c/ Madrid 126, E - 28903 Getafe, Spain

Werner Kleinhanß
Federal Agricultural Research Centre, Braunschweig

Carmen Murillo
Departamento de Economía, Universidad de Cantabria, Santander

October 28, 2005

## Abstract

The purpose of this paper is to model the interaction between the targets of the current CAP: environmental adaptation, subsidies and efficiency the of animal farming. To this end we first have to identify the production frontier and relative efficiency level for each animal oriented type of farm in the sample. The production frontier and efficiency index for each type of farm (assuming no specific production functions) are identified using DEA techniques. We then address the relationship between relative efficiency, farm size and environmentally friendly behavior realizing a non parametric 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 where 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 absolute 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 ones are more efficient.[1]

**Keywords**: Efficiency, subsidies, DEA, non-parametric regression, ecological 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, bird influence) 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 towards the decoupling of income from prices. Additionally, CAP MTR introduces a modulation of the direct payment (e.g., limiting direct payments by size).[2] 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 are, in a nutshell:

- 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, 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 money,

- a reduction in direct payments ("modulation") for lager farms to increase the budget for the new rural development policy.

In the WTO negotiations millennium round the exporting countries of the Cairns group asked for the total abolition of the agricultural subsidies. The Commission has successfully argued for multi-functionality 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.[3]

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

---

[2]Commissioner Fischler, the European Commission staff 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".

[3]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, well illustrate the complementary relationship between animal farming and landscape preservation.

CAP reform increases in the level of direct payments were followed by a rise in efficiency, especially for the environmentally friendly farms. This could actually politically justify the reinforcement of the status quo with minor changes. Fischler, on the other hand, strongly warns EU farmers to 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 paper is the analysis of this triangular relation between efficiency, environmental friendliness and subsidies in the EU. More specific, we try to find some answers on the following chain of question.

First, it is still costly to be environmental 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 environmental friendly farms? More specifically we can study how much they improve compared to 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 status quo farmer's behavior.

So, to answer these questions we propose basically three 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 having in mind that, within the farm possibilities of 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).

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

Note that the calculation of $E_W$, $E_{DP}$ is not of direct interest in our paper but serves as an auxiliary step.

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

When recalling the second question it is clear that 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 $E_{DP}$, i.e. for the statistical part of this analysis we need the same steps as we need to answer the first above mentioned 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, e.g. haging become less environmental 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 paper.

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 semi-log-linear regression (i.e. correlation and semi-log-correlation) of $E_W$ on $DP$. The results are alarming: even though efficiency increases weakly with absolute $DP$ increase (not surprising), it decreases for a relative increase in $DP$ (perhaps surprising?). One conclusion is that it is counterproductive when $DP$ is positively related to size.

This produces our fourth and final question: How are the $DP$ related to size $ESU$? Here we have simply looked at different linear and log-linear relationships between $DP$ and $ESU$. We belief 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. That correlation is not surprising as "decoupled subsidies" in terms of 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 paper 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 cows) crisis. We repeat efficiency calculations for two 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 polemic about the influence of subjective modelling.

Efficiency is measured using an index calculated with DEA (Data Envelope Analysis), and again 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 "crop subsidies" on the top of agricultural products. This allows us to compare relative efficiency with direct subsidies included in the efficiency index to the case when they are eliminated in the so called counterfactual efficiency calculations. In the following we always refer to $E_{DP}$ when speaking of efficiency, and to $E_W$ when speaking of conventional efficiency.

The implications are important for the future application of the recently approved CAP Reform 2003 on historical basis. Results can potentially be translated into promoting the wrong type of farming, as in past years, e.g., 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

Recall that we are interested in the trade-off between efficiency, environmental conservation and the CAP's use of the direct payment (DP).

The DPs were originally introduced in the McSharry CAP reform to decouple income from guarantee prices, and to control the intervention stocks. The theoretical idea behind the introduction of direct payments is the anti-production premium which assumes that the farm income is the result of the difference between income and cost plus the anti-production premium. This premium should be related with the income loss due to the output reduction. In practice direct payments were fixed by area, by head of certain animal 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 reenforced through the reduction of the guarantee 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 guarantee 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 (NUTS2 in Eurostat jargon). Notice that the output reduction is not directly observed. Therefore, the policy measure (i.e. the anti production premium) 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/AUA) 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 often required in parametric production functions for inputs and outputs (i.e. under perfect competition), even in the most flexible specifications. Further, as the topic is known to be rather polemic by nature, we prefer not to use specific parametrization for our models, i.e. neither for the production function when calculating efficiency nor for the regression when estimating the effect of size and environmental friendliness on efficiency.[4]

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

The producer level of direct support is measured by the total amount of direct payments in farm accounts; size by European Size Units (ESU), and livestock unit equivalents per agricultural utilized area (LU/AUA) 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[5]:

---

[4]As a referee noticed, the effect of size on efficiency could also be studied by other methods, e.g. though the measurement of scale efficiency, see Førsund (1996) or Banker, Cooper, Seiford, and Zhu (2004).

[5]We always have used nominal prices and real quantities accounted by the farmers as given in FADN. It could be discussed to use world market prices instead, but 1. 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 on world market prices. 2. It is beyond the scope of this paper to study the efficiency under different trade liberalization scenarios.

| OUTPUT | $Q_v$ | plant 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, agro-chemicals, seeds, water and other crop specific inputs, fuels and lubricants) |
| | W | wages |
| | AUA = $A_I + A_{NI}$ | Agricultural Utilized Area of farm aggregate adjusted by quality (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 |

The livestock is considered in the farmer's output because output is defined here by stock variation and sales, for details see MCKay, Lawrence, and Vlastuin (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 by departing from the acquisition price and inputting 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.

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 (AUA) irrigated (Ha.); $A_{NI}$: AUA non-irrigated (Ha.); $P_I$: price AUA irrigated by region (Euros/Ha); $P_{NI}$: price AUA non-irrigated by region (Euros/Ha). The average price by region is a weighted average of the irrigated and non irrigated different type of lands in the region (e.g. grass land divided in two quality levels, plus arable land, greenhouses, permanent crops (by type of tree: fruits other than citruses, citruses, vinegars, olive) see for details Decimavilla and San Juan (2002). For Germany we took the total of AUA as the differentiation of irrigated and non irrigated land does not make much sense.

There are three things that might cause 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.

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 in to 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 non-parametric 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.

The function $g : \Re^2 \to \Re$ 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 non-parametric, we could have directly used the co-variates $ESU$ and $EF$ in the 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 subvention policy, for a fair evaluation we have to choose the same measure which the European Community applies, i.e. Livestock Units per Agricultural Utilized Area, i.e. LU/UAA (see discussion in the introduction). The LU/UAA indicator is an indicator highly related to other environmental amenities of the farm such us untouched landscape, traditional buildings, wild animal habitats, bio-diversity, preservation of the regional non intensive 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/AUA 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 (or some products do not have DP specific programs but some farms can be engaged in other environmental programs that provide subsidies to improve environmental behavior in that 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

|  | Spain | | | |
| --- | --- | --- | --- | --- |
|  | numbers | | percentages | |
| Type \ Year | 1999 | 2000 | 1999 | 2000 |
| 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 | | | |
|  | numbers | | percentages | |
| Type \ Year | 1999 | 2000 | 1999 | 2000 |
| Cattle | 353 | 358 | 97.245 | 98.623 |
| Pig Farming | 245 | 245 | 100 | 100 |
| Total | 598 | 603 | 98.355 | 99.178 |

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

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 animal production, i.e. farms with larger animal than vegetal output in both years. To include only farms with similar production functions and with the possibility of cropping plant products for re-use on the farm or for sale, we selected only farms with both positive animal and plant production. We believe that farms without land have a non-comparable production function and will therefore be excluded, e.g. think of fattering farms.

For our analysis, the sample is split by country (Germany and Spain), and by type of animal farming (cattle farming, pig farming and sheep and goats) as these different farm types are neither uniform in the treatment by CAP nor in the production process. 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 (years 1999–2000) of the Farm Accounting Data Network (FADN). Every year, the survey gathers information on the characteristics of the farm (AUA, LU, type of livestock and crops, economic size of the farm, ...) 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, the livestock production is always greater than plant 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.

| Year | | | ′99 | ′00 |
|---|---|---|---|---|
| | | sheep and goats | 553 | 679 |
| Spain | | cattle | 1435 | 1543 |
| | | pig | 255 | 249 |
| Germany | | cattle | 604 | 604 |
| | | pig | 355 | 355 |

Table 2: Number of farms that are used for all our calculations.

# 3    Methodologies

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

## 3.1    Estimation of Efficiency

Since we do not want our results and conclusions to be dependent on a particular specification of the production function, we have chosen the DEA method. Data envelopment analysis (DEA) is a non-parametric approach using linear programming methods to determine the envelopment of the DMUs (decision making units) thus identifying "the best practice" for each productive unit.

Then, measures are calculated relative to this frontier for each individual Debreu-Farrell efficiency (see Cooper, Seiford and Tone (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 us define some basic concepts of non-parametric efficiency measurement. 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:

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

10

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. E.g. 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) to (7) as an output oriented problem. In practice, the input orientated is more popular due to its easier interpretation, e.g. it corresponds to the idea of cost efficiency. However, as a referee pointed out, in particular the inclusion of SP= $-$DP would make the output orientated DEA version more intuitive as a farmer is not interested in minimizing / reducing the direct subsidies received. However, the contrary argument would hold for many of the other inputs when choosing an output orientated DEA. The interpretation in an input orientated DEA would not be that the farmer with an index $E < 1$ is interested in reducing DP (by factor $E$), but that he could reach the same output level with less DP and is therefore inefficient at the present state. Nevertheless, we repeated some of the calculations with the output orientated DEA, that certainly suffers to some extend similar criticism as the input orientated one. 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 even discriminates the farms somewhat more). Not surprisingly, the final conclusions turn out to be the same.

Further alternatives are e.g. DEA-methodologies that allow combining both minimizing input and maximizing output, see González Fidalgo (2001) or Banker, Cooper, Seiford, and Zhu (2004). On the other hand these method also needs several assumptions that not necessarily meet in our context as for example in González Fidalgo (2001) minimizing input and maximizing output is restricted to occur always in the same proportion. Finally, DEA also allows for fixing some of the input, 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, e.g. additional land could be rented.

For the rest of the paper 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, Grosskopf and Lovell, 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, Cooper and Rhodes, 1978) was defined under constant returns to scale, but later papers have considered alternative sets of assumptions such the case of variable returns to scale by

---

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

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

Banker, Charnes and Cooper (1984). In any case, and under some regularity assumptions on the data generating process specified in Kneip, Park and Simar (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:

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

where $\gamma_i$ 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 the two first constraints in the equation (8) imply that an excess of outputs or inputs can be disposed of freely.

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

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

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

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

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

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 as $\hat{\theta}(x_j, y_j) < 1$ means that the firm is inefficient.

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

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

As already 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 preferred 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 non-interpretable hyper-plane in the DEA. Nevertheless we admit that handling DP in the model as input subsidy or output related subsidy 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 subsidy. 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, it could be argued as to whether more input variables should be aggregated for a non-parametric analysis such as DEA, e.g. 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: more aggregating leads to more bias but less variance and vice versa. We have decided here to opt for high resolution, i.e. high variance, small bias. So, since we always conduct non-parametric 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 on $g(\cdot,\cdot)$ except that it is a smooth function, i.e. it has continuous second derivatives.

We will now present a brief insight into the algorithms of non-parametric flexible function regression. In particular we explain the estimation of the parameters $\beta$ and its asymptotic covariance, as well as the estimation of the non-parametric function of $g(\cdot,\cdot)$ in a semipara-metric model of the form as described in equation (2). We assume $E[e|EF, ESU, R] = 0$, $Var[e] < \infty$. The estimation of $g(\cdot)$ and $\beta$ will be made in two steps: first the estimation of $\beta$ and its covariance using the method of Robinson (1988), and afterwards the estimation of $g(\cdot,\cdot)$ using local linear smoothing by Ruppert and Wand (1994). For a more detailed introduction to non- and semi-parametric modeling see also Härdle, Müller, Sperlich, and Werwatz (2004).

The basic idea is to construct an estimator that creates a smooth surface (or hyper-plane), e.g. 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. Actually, this parameter can often also be data driven.

First, it is important to understand that this estimator works locally, e.g. we estimate the desired function, the hyper-plane, separately at each point we are interested in. Therefore we need to introduce some additional notations. Consider for a moment 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. Imagine 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:

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

$a_0 \in \Re$, $a_1 \in \Re^d$ 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 that 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 in equation (12) the vector $a_1$ 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 give 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 hyper-plane, i.e. $h \to \infty$ 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, i.e. the empirical researcher would allow for more flexibility or impose more smoothness on its functions. To allow for high flexibility without increasing the variance to unreasonable levels, we chose smoothing parameters that did not restrict the functional forms unless the plotted surface became wiggly.

Returning back to our model (2), we will apply the local linear estimation method, i.e. equation (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 $\beta$ is defined as

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

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 local constant smoother as defined in the equation (13). It is easy to see that the variance of $\hat{\beta}$ can be estimated by $\hat{\sigma}^2 S_{R-\tilde{R}, R-\tilde{R}}^{-1}$ with $\hat{\sigma}^2$ being a consistent estimator of the conditional variance of $E$: $\sigma^2 = Var[E|W_i, R_i]$. For more details see Robinson (1988). Note that all these models have been proved to work perfectly also for dependent data. It is worthwhile to mention this as the indices calculated by DEA are not independent.

Certainly, as $g(\cdot, \cdot)$ is non-parametric, we could have directly used the co-variates $ESU$ and $EF$ in the model (2). As mentioned above, the logarithm does therefore not impose any model specification here. The problem is that both variables have a rather skewed distribution 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 on 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; for 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 equation (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 an auxiliary step (see first section), we will neither explicitly give 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

$$E \;=\; g[\ln(EF), \ln(ESU)] + \beta^T R + e \;,$$

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 checked if and how the CAP policy distorts efficiency. The estimation procedures applied here have been explained in detail in Section 3.2. When we speak of significance in the following, we always refer to the 10% significance level.

In non- and semi-parametric regression, the choice of smoothness controlled via the bandwidth (named $h$ in Section 3.2) and chosen by the empirical researcher, 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 trade-off between bias and variance of the estimates. For the parametric part $\beta$ 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 non-parametric part. This is expected if e.g. 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, i.e. on $\hat{\beta}$, summarized in Table 3 for Spain in 1999 and 2000. We divided Spain into 5 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 even though not significantly better than the center and the south. These findings hardly change with the dependent variable, being $E_{DP}$ and its counterfactual opposite $E_W$.

| 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 3: Estimates (upper lines) with standard error (lower lines) for the regional dummies "$\beta$" in model (2) for **Spain**.

In Table 4 the corresponding results for Germany are given, also for 1999 and 2000, but without sheep and goat farms. We divided Germany in to only four regions: North, Center, South, and East. Again, the latter has been used as a normalizing region.

In Germany, the south is the most efficient in cattle farming whereas it is hard to say something 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) ones in Eastern Germany. However, significance changes with year and the dependent variable, $E_{DP}$ or $E_W$.

Much more interesting for us is the outcome of the impact estimates for environmentally

| 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 |

Table 4: Estimates (upper lines) with standard error (lower lines) for the regional dummies "$\beta$" in model (2) for **Germany**.

friendly behavior ($EF$) and farm size ($ESU$) on (conventional) efficiency. As the functional form of $g(\cdot, \cdot)$ in model (2) is non-parametric, the results are given graphically, see Figures 1 to 10. Here, only the graphs for 2000 are shown, the graphs for 1999 are available on page `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 graphs: a three dimensional and a two dimensional one. The two dimensional one 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 in terms of the actual price structure. This holds regardless of whether direct subsidies are included in the calculation of efficiency or not. Generally, efficiency (for both, $E_{DP}$ and $E_W$) increases when
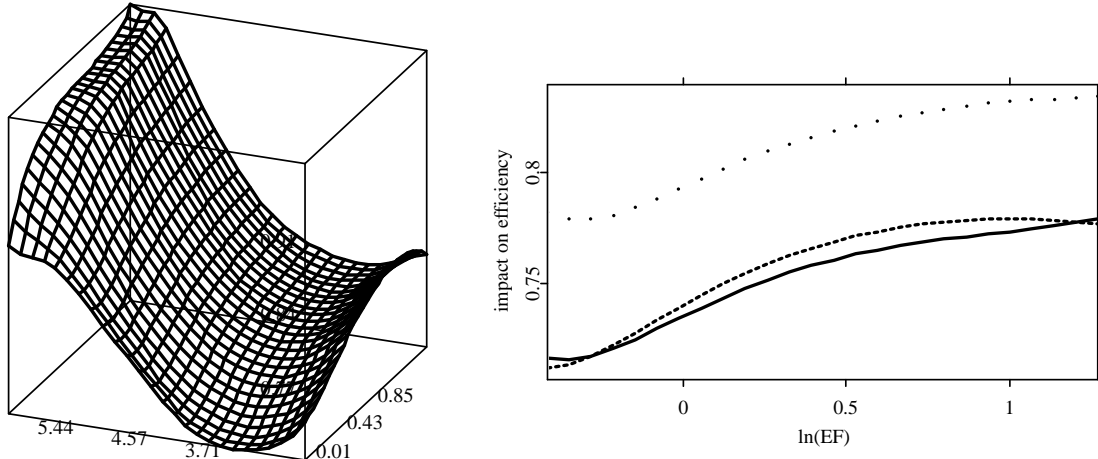
Figure 1: German cattle farms 2000, $\mathbf{E_W}$. Left: $g_W(\ln(EF), \ln(ESU))$ from equation (2) where lower left axis is the impact of $\ln(EF)$, lower right of $\ln(ESU)$. Right: $g_W$ with $ESU$ fixed at median (solid), 95% quantile i.e. large farms (dotted line), and 5% quantile i.e. small farms (dashed line). Outer 2% are cut off in both figures.
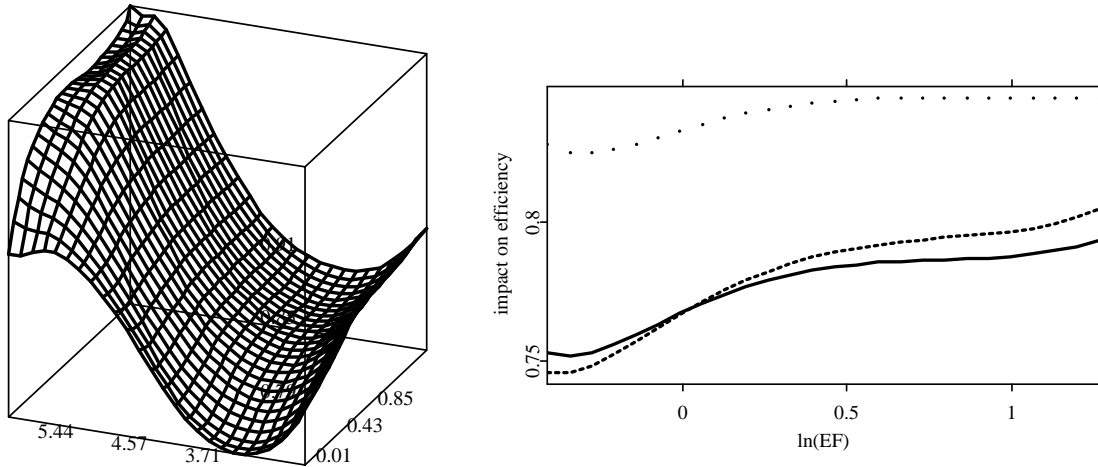


Figure 2: German cattle farms 2000, $\mathbf{E_{DP}}$. Left: $g_{DP}(\ln(EF), \ln(ESU))$ from equation (2) where lower left axis is impact of $\ln(EF)$, lower right of $\ln(ESU)$. Right: $g_{DP}$ with $ESU$ fixed at median (solid), 95% quantile i.e. large farms (dotted line), and 5% quantile i.e. small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

livestock units per agricultural utilized area increase (under the actual price structure of inputs and outputs). 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, in particular 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 1992, (in force since 1993), yields some visible effects, especially
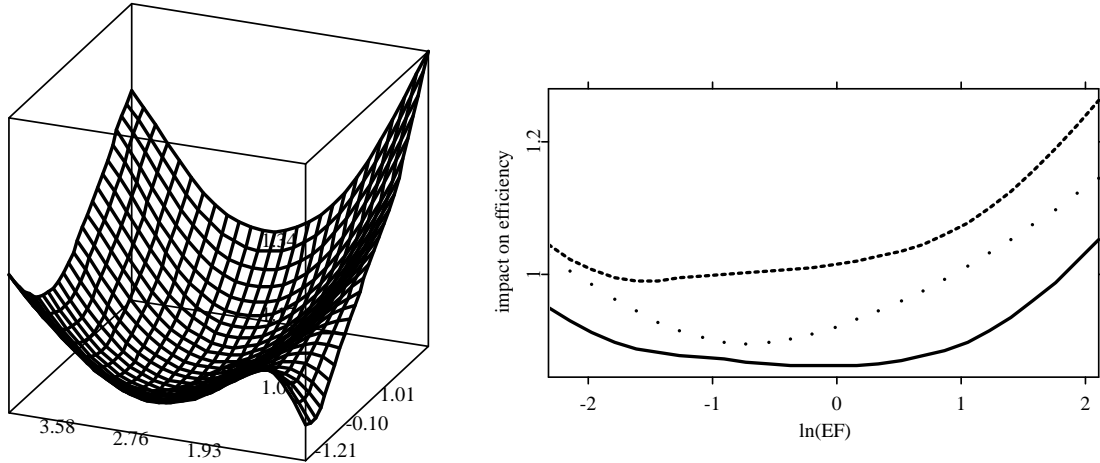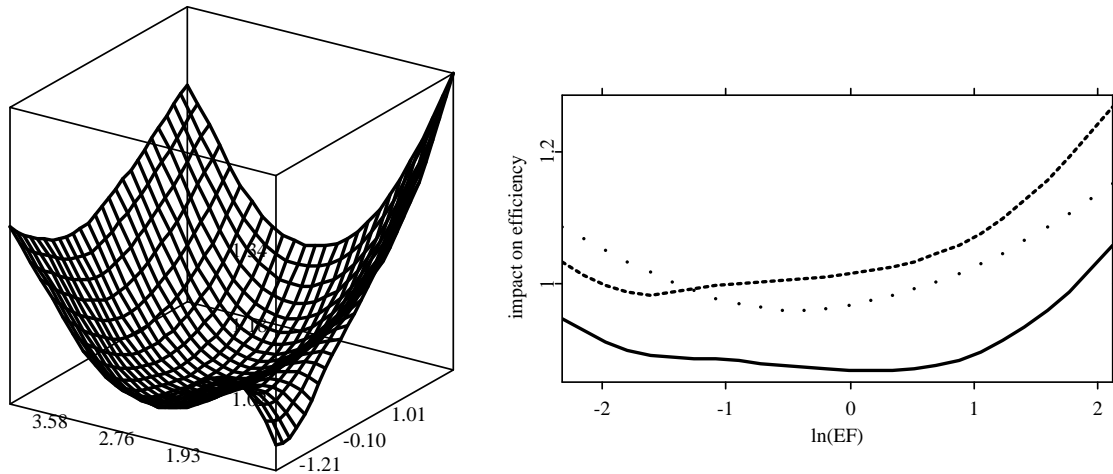
Figure 3: Spanish cattle farms 2000, $\mathbf{E_W}$. Left: $g_W(\ln(EF), \ln(ESU))$ from equation (2) where lower left axis is impact of $\ln(EF)$, lower right of $\ln(ESU)$. Right: $g_W$ with $ESU$ fixed at median (solid), 95% quantile i.e. large farms (dotted line), and 5% quantile i.e. small farms (dashed line). Outer 2% are cut off in both figures.



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

in the sector of pig farming.

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

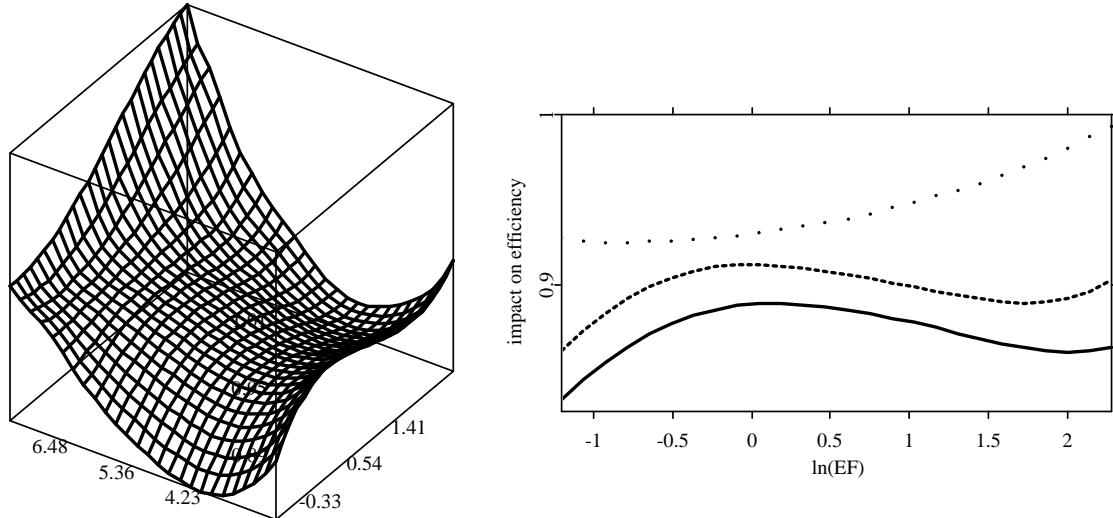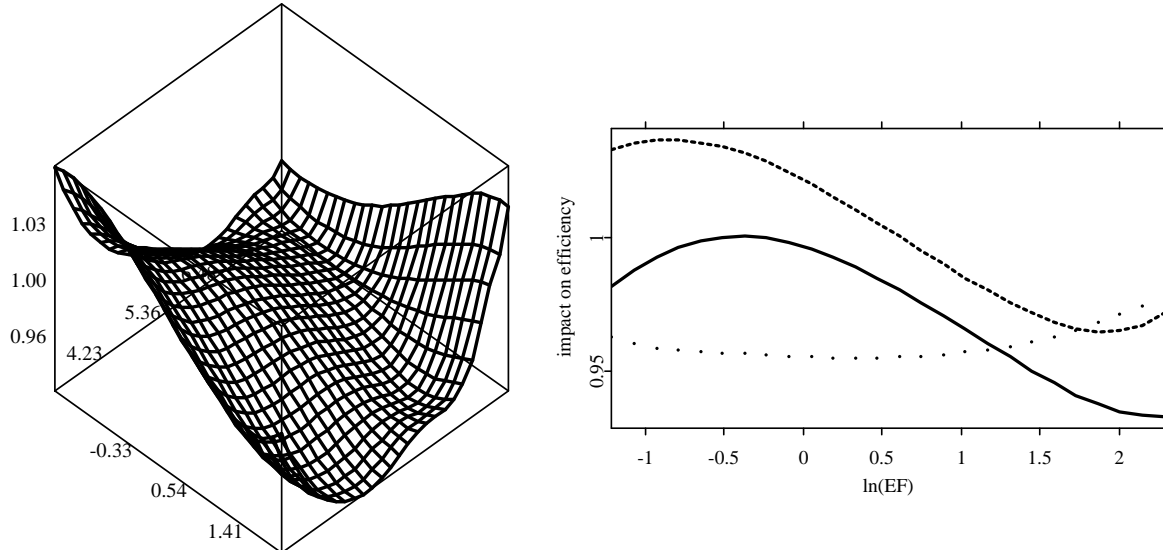Conversely, in cattle farming, the conventional holdings are more efficient than those most

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



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

adapted to the environment. In this sector, the results hold true with and without direct payments, i.e. independently of whether looking at $E_{DP}$ 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 certain in Germany) there is a positive distortion caused by subsidies for environmental friendliness. Moreover, we found empirical evidence that the current direct payments system is "environmentally oriented"
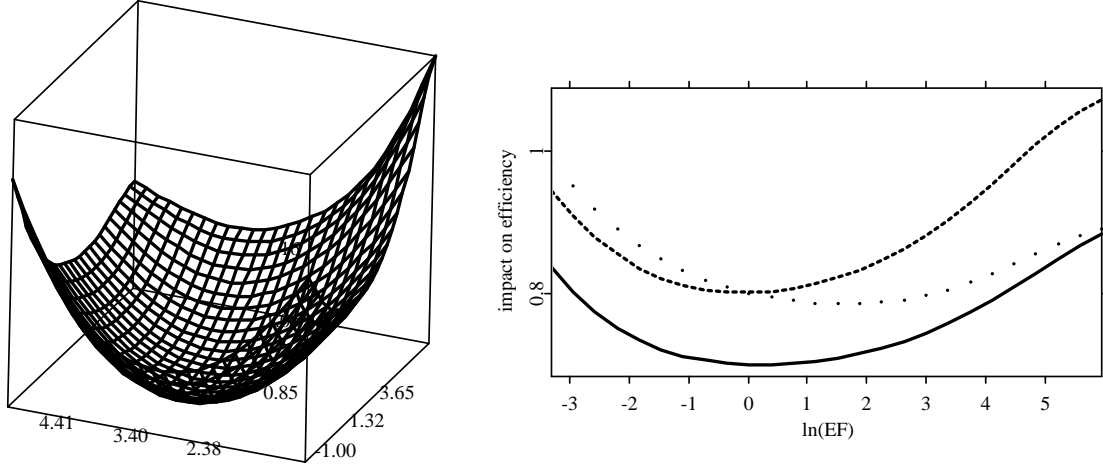
Figure 7: Spanish pig farms 2000, $\mathbf{E_W}$. Left: $g_W(\ln(EF), \ln(ESU))$ from equation (2) where lower left axis is impact of $\ln(EF)$, lower right of $\ln(ESU)$. Right: $g_W$ with $ESU$ fixed at median (solid), 95% quantile i.e. large farms (dotted line), and 5% quantile i.e. small farms (dashed line). Outer 2% are cut off in both figures.
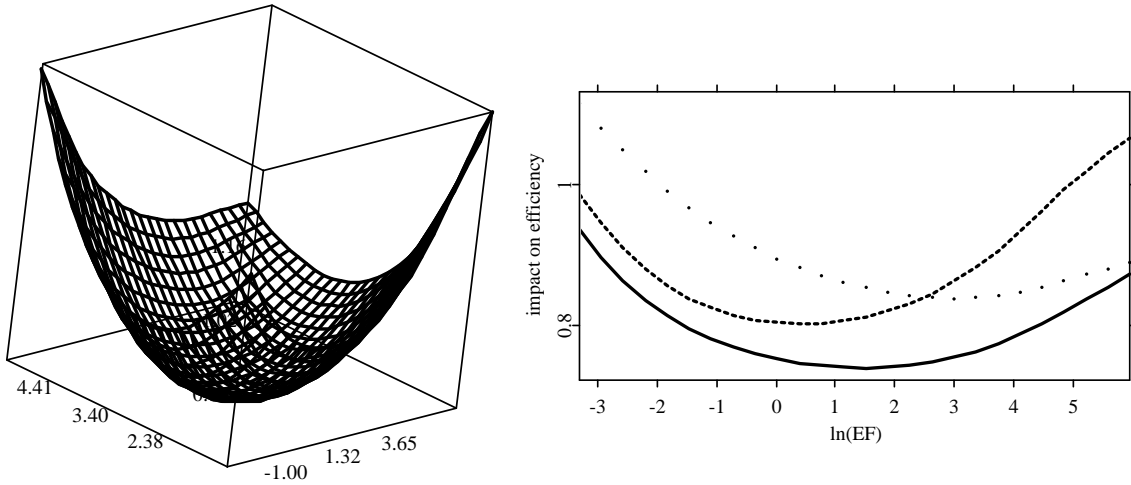


Figure 8: Spanish pig farms 2000, $\mathbf{E_{DP}}$. Left: $g_{DP}(\ln(EF), \ln(ESU))$ from equation (2) where lower left axis is impact of $\ln(EF)$, lower right of $\ln(ESU)$. Right: $g_{DP}$ with $ESU$ fixed at median (solid), 95% quantile i.e. large farms (dotted line), and 5% quantile i.e. small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

in all types of farms studied in Germany but only for pig farming in Spain, (see Tables 5 to 9). Taken together, this information offers the conclusion that the 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 the environmentally friendly farm by increasing the incentives for ecological farming.

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 of the direct subsidies and the proxy of adaptation to the natural environment, i.e. direct payment negatively correlated with LU/AUA, for both
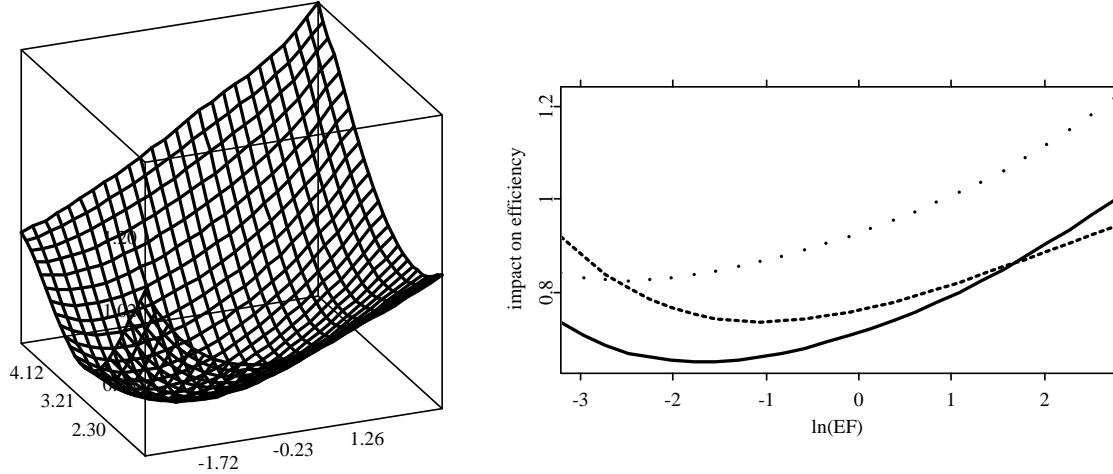
21

Figure 9: Spanish sheep and goat farms 2000, $\mathbf{E_W}$. Left: $g_W(\ln(EF), \ln(ESU))$ from equation (2) where lower left axis is impact of $\ln(EF)$, lower right of $\ln(ESU)$. Right: $g_W$ with $ESU$ fixed at median (solid), 95% quantile i.e. large farms (dotted line), and 5% quantile i.e. small farms (dashed line). Outer 2% are cut off in both figures.
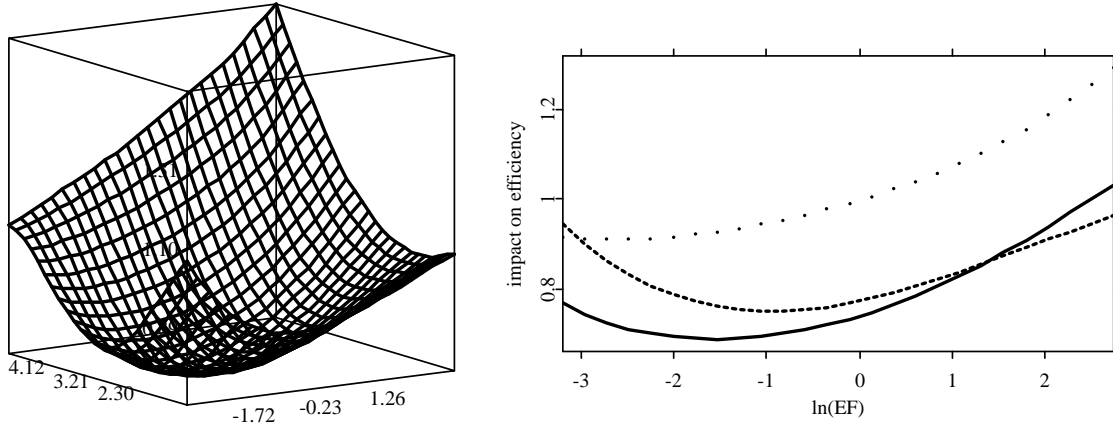


Figure 10: Spanish sheep and goat farms 2000, $\mathbf{E_{DP}}$. Left: $g_{DP}(\ln(EF), \ln(ESU))$ from equation (2) where lower left axis impact of $\ln(EF)$, lower right of $\ln(ESU)$. Right: $g_{DP}$ with $ESU$ fixed at median (solid), 95% quantile i.e. large farms (dotted line), and 5% quantile i.e.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

countries and years over all models considered.

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

Also the counterfactual exercise (looking on $E_W$) 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 result. This finding is interesting

22

since this result fits perfectly with the conclusions of Álvarez and Arias (2003), who pointed out that increasing cattle farm size while holding managerial ability constant can be an important source for 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 the small farms versus the medium farm.

German pig farming is an interesting case because the small environmentally friendly farms become 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 calculating the efficiency upholds these results as well. This was a remarkable difference in the results of 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) ones, which is not the case in Spain.

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

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

It is clear that when efficiency is calculated including $DP$, $E_{DP}$ will be (positively) correlated with it; the result can change when looking on $\ln(1+DP)$. Further, we expect that the 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 very year. So, in $E_W$ the long-term effect of (former received or not received) direct payments is still reflected. For this reason it is clear that we are more interested in $corr(DP, E_W), \ corr(\ln(1+DP), E_W)$ than in $corr(DP, E_{DP}), \ corr(\ln(1+DP), E_{DP})$.

Again, first we give the results for Spain for 1999 and 2000, drawn in Table 5 to Table 7.

First, let us briefly summarize the signs we see 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 European Community (recall that

**Cattle**

| year | variables | $(\cdot),(\cdot)$ | $(\cdot),\ln(\cdot)$ | $\ln(1+\cdot),(\cdot)$ | $\ln(1+\cdot),\ln(\cdot)$ |
|------|-----------|---------|---------|---------|---------|
| 1999 | $DP, EF$ | -.1267 | -.4134 | -.3864 | -.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 | $DP, EF$ | -.1808 | -.4311 | -.3264 | -.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 5: Correlations (upper lines) with p-values for significance (lower lines) for **Spain**.

**Pig Farming**

| year | variables | $(\cdot),(\cdot)$ | $(\cdot),\ln(\cdot)$ | $\ln(1+\cdot),(\cdot)$ | $\ln(1+\cdot),\ln(\cdot)$ |
|------|-----------|---------|---------|---------|---------|
| 1999 | $DP, EF$ | -.0893 | -.2574 | -.4323 | -.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 | $DP, EF$ | -.1427 | -.4211 | -.4718 | -.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 |
|      | $DP, E_W$ | -.1462 | – | -.2370 | – |
|      |          | 0.0210 | – | 0.0002 | – |
|      | $DP, E_{DP}$ | 0.0580 | – | -.1355 | – |
|      |          | 0.3622 | – | 0.0326 | – |

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

**Sheep and Goats**

| year | variables | $(\cdot),(\cdot)$ | $(\cdot),\ln(\cdot)$ | $\ln(1+\cdot),(\cdot)$ | $\ln(1+\cdot),\ln(\cdot)$ |
|------|-----------|-------------------|----------------------|------------------------|----------------------------|
| 1999 | $DP, EF$ | -.0817 | -.2722 | -.2705 | -.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 | – | -.2046 | – |
|      |          | 0.2594 | – | 0.0000 | – |
|      | $DP, E_{DP}$ | 0.1136 | – | -.0638 | – |
|      |          | 0.0075 | – | 0.1341 | – |
| 2000 | $DP, EF$ | -.0862 | -.2845 | -.2750 | -.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 | – | -.0956 | – |
|      |          | 0.0168 | – | 0.0127 | – |
|      | $DP, E_{DP}$ | 0.1516 | – | 0.0241 | – |
|      |          | 0.0000 | – | 0.5302 | – |

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

**Cattle**

| year | variables | $(\cdot),(\cdot)$ | $(\cdot),\ln(\cdot)$ | $\ln(1+\cdot),(\cdot)$ | $\ln(1+\cdot),\ln(\cdot)$ |
|------|-----------|-------------------|----------------------|------------------------|----------------------------|
| 1999 | $DP, EF$ | -.0914 | -.1102 | -.1686 | -.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 | – | -.1909 | – |
|      |          | 0.0612 | – | 0.0000 | – |
|      | $DP, E_{DP}$ | 0.1168 | – | 0.0142 | – |
|      |          | 0.0040 | – | 0.7275 | – |
| 2000 | $DP, EF$ | -.0633 | -.0746 | -.1398 | -.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 8: Correlations (upper lines) with p-values for significance (lower lines) for **Germany**.

**Pig Farming**

| year | variables | $(\cdot),(\cdot)$ | $(\cdot),\ln(\cdot)$ | $\ln(1+\cdot),(\cdot)$ | $\ln(1+\cdot),\ln(\cdot)$ |
|------|-----------|--------|--------|--------|--------|
| 1999 | $DP, EF$ | -.2678 | -.3708 | -.4654 | -.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 | $DP, EF$ | -.2692 | -.3541 | -.4525 | -.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 | – |

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

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 ones. Looking at the pair $corr(DP, E_W)$, $corr(\ln(1 + DP), E_W)$ we get the following schema 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. But do not forget that here we are not correcting for endogeneity. That 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", e.g. 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%, i.e. subsidies can mainly be explained

26

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 believed 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 certainly 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 really 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 the particular 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. This again with the exception of Spanish pig farms. However, looking at a percentage increase of direct payment, i.e. on $corr(\ln(1 + DP), E_W)$, efficiency decreases or stagnates for all years and countries except for Spanish cattle farms. In other words, focusing on the efficiency, a policy which grants subsidies per farm seems much more reasonable than the current (and thus 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 actual 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, e.g. 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 takes environmental friendliness into account. However, it is evident from the tables that this correlation is much weaker, almost negligible compared to the overwhelmingly strong correlation between subsidies and farm size.

## 5   Conclusions

The main empirical conclusions are the following.

Maybe not surprisingly, 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 what the present subvention

policy is doing: farm size explains about 90% (or more) of the direct payments.

Note further that the strong subsidization of large farms cannot be justified with their presumably high efficiency or environmental friendliness. Neither our graphical nor our numerical results confirm, or even contradict, 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 ones. 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 guess 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 numerical results on correlations, we conclude that there is some empirical evidence that the actual direct payments system is "environmentally oriented" for all types of farms studied (decreasing level of direct payments as the livestock units per agricultural utilized area increases). Thus, our calculations show that the actual 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 in particular in Germany, it is still not sufficient to motivate farmers to change their production towards a more environmental friendly one.

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

## 6    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, see discussion in the previous sections, we have in most cases 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 most density close to zero. Together with the former plots ($E_{DP}$ in 1999 and 2000) we can already conclude that these small changes cannot affect the all over conclusions of our analysis.
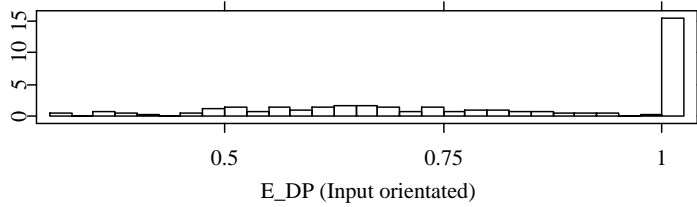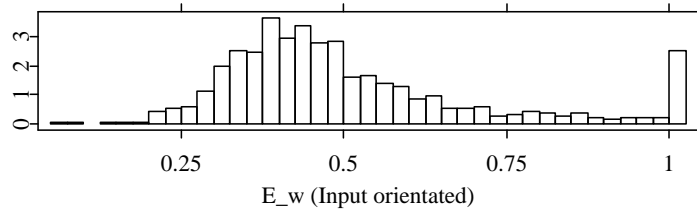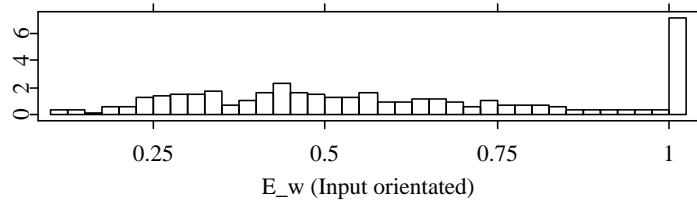
## Cattle (1999)
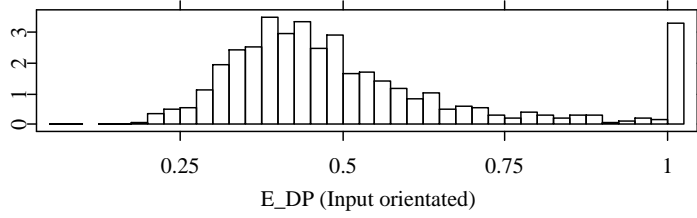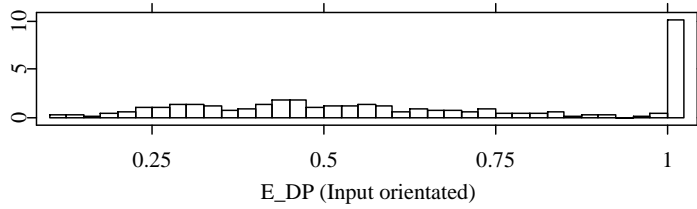
E_DP (Input orientated)

## Pigs (1999)

E_DP (Input orientated)

## Sheeps and Goats (1999)

E_DP (Input orientated)

## Cattle (1999)

E_w (Input orientated)

## Pigs (1999)

E_w (Input orientated)

## Sheeps and Goats (1999)

E_w (Input orientated)

## Cattle (2000)



E_DP (Input orientated)

## Pigs (2000)



E_DP (Input orientated)

## Sheeps and Goats (2000)



E_DP (Input orientated)

## Cattle (2000)



E_w (Input orientated)

## Pigs (2000)



E_w (Input orientated)

## Sheeps and Goats (2000)



E_w (Input orientated)

## Cattle (1999)



Orientation differences in E_DP

## Pigs (1999)



Orientation differences in E_DP

## Sheeps and Goats (1999)



Orientation differences in E_DP

## Cattle (2000)



Orientation differences in E_DP

## Pigs (2000)



Orientation differences in E_DP

## Sheeps and Goats (2000)



Orientation differences in E_DP
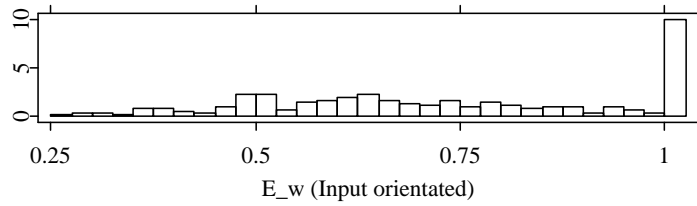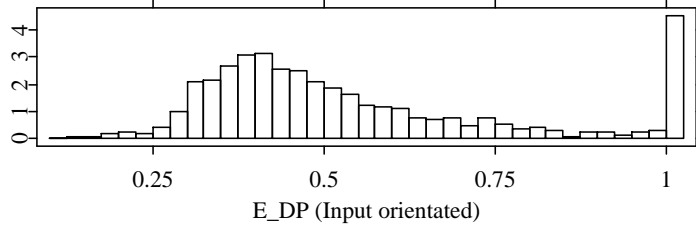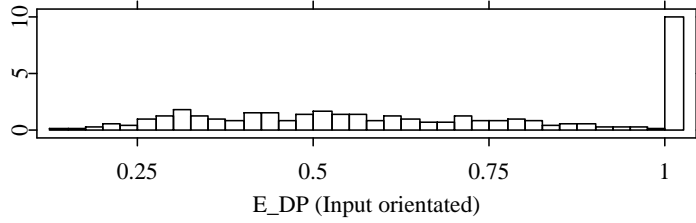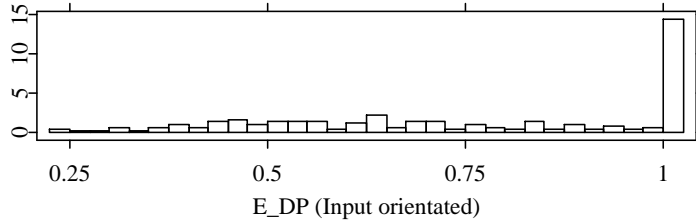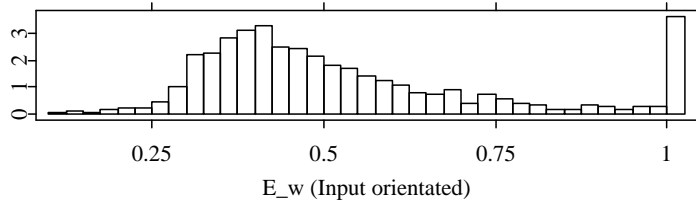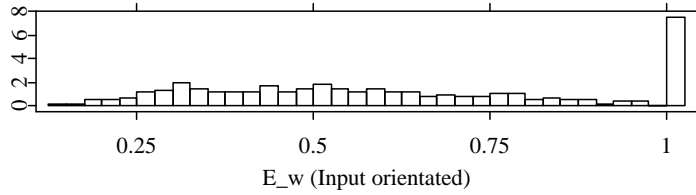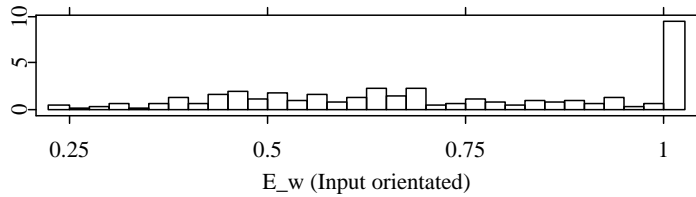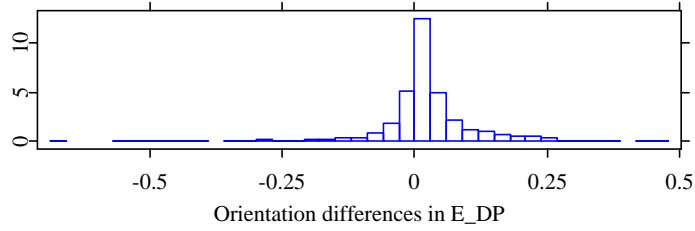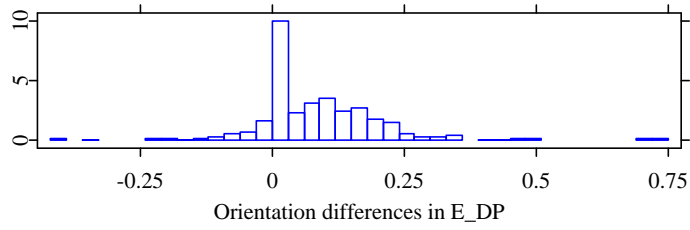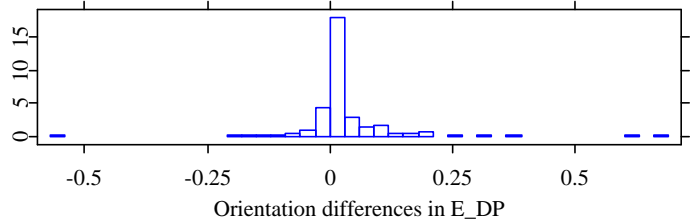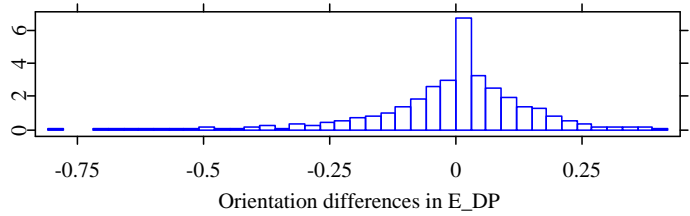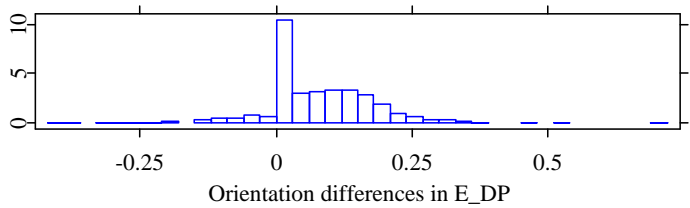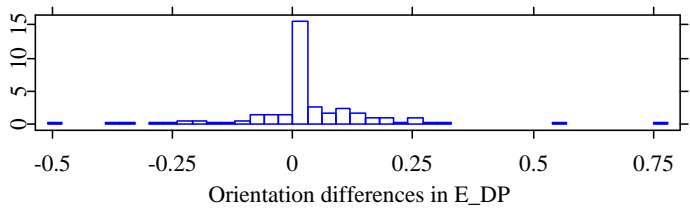
# References

Álvarez Pinilla, A. (2001) Concepto y medición de la eficiencia productiva, pp. 19-40 in A. Álvarez (ed.) *La medición de la eficiencia y la productividad*, Pirámide, Madrid.

Álvarez Pinilla, A. and Arias, C. (2003) Diseconomies of Size with Fixed Managerial Ability, *Amer. J. Agr. Econ.*, **85**, (1), 134–142.

Ball, V. E., J. P. Butault, and C. San Juan (2004) Measuring Real Capital Input in OECD Agriculture Canadian Journal of Agricultural Economics, November (forthcoming)

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

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

Banker, R.D., W.W. Cooper, L.M. Seiford, and J. Zhu (2004) Returns to Scale in DEA. in W.W. Cooper, L.M. Seiford and J. Zhu (eds.) *Handbook on Data Envelopment Analysis*, Kluwer, London.

Bertola, G., T. Boeri, H. Brücker, F. Coricelli, A. de la Fuente, J.J. Dolado, J. Fitzgerald, P. Garibaldi, G. Hanson, J.F. Jimeno-Serrano, R. Portes, G. Saint-Paul, and A. Spilimbergo (2002) Who's Afraid of the Big Enlargement? *CEPR Policy Paper 7*, `www.cepr.org/pubs/books/cepr/booklist.asp`

Biedermann, S., and H. Dette (2001) Optimal designs for testing the functional form of a regression via non-parametric estimation techniques. *Statistics & Probability Letters*, **52**, 215–224.

Charnes, A., W.W., Cooper, and, E. Rhodes (1978) Measuring the Inefficiency of Decision Making Units. *European Journal of Operational Research*, **2**, 429–444.

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

Decimavilla, E. and C. San Juan (2002) Encuesta de precios de la tierra 2001. Base 1997. Metadología. *Boletin Mensual de Estadística Agraria*, November **11**, 7-76.

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

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

Farrell, M.J. (1957) The Measurement of Productive Efficiency. *Journal of the Royal Statistcal Society: Series A (General)*, **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. *The Journal of Productivity Analysis*, **7**, 283–302.

González Fidalgo, E. (2001) La estimación de la eficiencia con métodos no paramétricos, pp. 139-166 in A. Álvarez (ed.) *La medición de la eficiencia y la productividad*, Pirámide, Madrid.

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

Kneip, A., B.U. Park, and L. Simar (1998) A note on the convergence of non-parametric DEA estimators for production efficiency scores. *Econometric Theory*, **14**, 783–793.

MCKay, L., D. Lawrence, and C. Vlastuin (1983) Profit, Output Supply, and Input Demand Functions for Multiproduct Firms: The Case of Australian Agriculture. *International Economic Review*, **24**, 323–339.

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

Mora, R., and C. San Juan (2004) Product concentration and Farm Specialization in Spain after implementation of the CAP and its reform, pp. 105-121 in K. Poppe (ed.) *New Roads for Farm Accounting and FADN*, LEI. The Hague

Robinson, P. (1988) Root-N-Consistent Semiparametric Regression. *Econometrica*, **56**, 931–954.

Ruppert, D., and M.P. Wand (1994) Multivariate locally weighted least squares regression. *Annals of Statistics* **22**, 1346–1370.

Simar, L., and P. Wilson (2000) Statistical Inference in Non-parametric Frontier Models: The State of the Art. *Journal of Productivity Analysis*, **13**, 49–78.