Jahrbücher f. Nationalökonomie u. Statistik (Lucius & Lucius, Stuttgart 2007) Bd. (Vol.) 227/3

An Empirical Assessment of the EU Agricultural Policy Based on Firm Level Data

By Carmen Murillo, Santander, Carlos San Juan, Getafe, Stefan Sperlich, Goettingen*

JEL Q11, Q18, C32, F02

Common Agricultural Policy, subsidies, efficiency, non-parametric methods, natural resources, environmental economics.

Summary

The study focuses on testing the hypothesis that the subsidy system of the Common Agricultural Reform in 1992 (CAP'92) drove to changes in farm efficiency towards the thereby claimed objectives. With sequential applications of semiparametric methods we succeed to identify the impact of the direct payments on environmental adaptation, productivity and efficiency before and after CAP'92 without restrictive model specifications. We find that the claimed objectives of the EU subvention policy were met only partly, but that the CAP'92 was, however, a step forward. Our case study uses large Spanish data sets of animal orientated farms.

This paper applies non parametric methods for policy evaluation at firm level. The study focuses on testing the hypothesis that the subsidy system of the Common Agricultural Reform in 1992 (CAP'92) drove to changes in farm efficiency towards the thereby claimed objectives. We concentrate here on animal oriented farms, in particular cattle, pig, sheep and goat farms. The correct quantification of efficiency and productivity differentials due to CAP'92 is crucial for such a policy analysis as different models can easily lead to different conclusions. Using non parametric methods we do not need to specify the production function of the farms. With sequential applications of semiparametric methods we succeed to identify the impact of the direct payments on environmental adaptation, productivity and efficiency before and after CAP'92 without restrictive model specifications. We find that the claimed objectives of the EU subvention policy were met only partly, but that the CAP'92 was – at least partly – indeed a step forward in that sense. Our case study uses large Spanish data sets of animal orientated farms. This is justified, among other reasons, by the relevance of these farms for Mediterranean forest and grazing land preservation in Spain.

1. Introduction

The main purpose of this paper is to exemplify a policy evaluation using non parametric methods to assess the European Union agricultural reform of 1992. The study focuses on the effects of Direct Payments (DP) on items such as efficiency before and after the Common Agricultural Policy reform of 1992 (CAP'92). The general hypothesis to test is that DP bias efficiency in different degrees depending of the firm characteristics. Policy reform objectives include the small farm protection by facilitation of efficiency increase.

* The authors thank an anonymous referee for helpful comments and Alois Kneip, Antonio Alvarez, and Carlos Arias for discussion. This research was supported by FUNCAS, the Spanish Ministry of Agriculture, the "Dirección General de Investigación del Ministerio de Ciencia y Tecnología", project number SEJ2005-08269/ECON and SEJ2004-04583/ECON and CAM 2007/04099/001.

Another hypothesis to test is if the reformed CAP succeeded in reaching the objective of increasing the efficiency and the environmental orientation of the farms simultaneously. The introduction of DP, reinforced strongly by the CAP'92, was actually motivated and warranted with its potential impact on efficiency and on environmental adaptation of farms. Given the increasing role that DP will play in the future distribution of resources to support the farmers of the European Union, a quantitative framework to measure the degree of distortion due to the subsidy policy is highly relevant.

Clearly, the quantification of the efficiency bias induced by the new policy instruments, applied to farms which are heterogeneous in size and taken from regions with divers bio-physical characteristics, is crucial. Therefore, this will be done in several steps applying rather sophisticated (non- and semiparametric) methodology. Further, we have chosen Spain as a case study for several reasons; the relevance of animal oriented farms for Mediterranean forest and grazing land preservation, the importance of agriculture economy in Spain, and because of the large number of farms with available accounting data.

Efficiency and environmental adaptation have become key issues in new European agricultural policy. A step towards an agreement in the WTO (World Trade Organization) and to the decoupling of income from prices was the agreement of the Council of Ministers in June 2003, the Mid-Term Review of the Common Agricultural Politic [CAP MTR]. Additionally, the CAP MTR introduces a modulation of the direct payment, for example limiting direct payments by size. In view of the new CAP reform which is in force since 2005, it would be interesting to look back and study the effects of the previous reform CAP'92, especially regarding those aspects tackled by the recent reform like the substitution of price support by a single payment.

We will concentrate on animal oriented farms. As said, for the preservation of large extensions of forest and grasslands in Spain the key farms are those oriented to animal production partially based on their own, including rented, land's vegetal output. In fact, the Mediterranean forest and the traditional techniques of livestock raising are frequently close to ecological farming. Note that animal density and environmental friendliness are strongly related, so we will use these expressions synonymously in the following. In fact, the EU defines environmental friendliness via animal density. We will discuss this point later in this paper. Under the CAP, especially after the 1992 reform, there was a potential contradiction between several policy targets: intervention price reduction and a rise in the environmental adaptation of farms and improvement of efficiency. We explore the ex-post effects on farms using a data-set of a representative sample of individual holdings. Farm technologies allow for shifting between different proportions of animal and vegetal products.

The current CAP 2003 reform is, in many ways, a step ahead on the basic principles and tools introduced in the CAP'92 reform to control over-production by reducing intervention price and to use DP to compensate farmers for their income losses.

The first step of the recent CAP reform was to introduce direct payments and cut the intervention price while trying to reduce intervention stocks. Just to give an idea of the importance of DP: currently, due to the CAP, 4.50 million farmers benefit from subsidies of 24.8 billion Euros at the EU 15 level. Part of that goes to Spain where 489 thousand beneficiaries received 2.98 billion Euros. On average, the farms in Spain receive fewer in DP per holding than the average EU farm. Note also that so far, the allocation of direct spending among farmers has been known to be unequal in two ways: First, the bulk of direct income support is concentrated in a few beneficiaries. For example, the Commission acknowledges that in the 2000-01 financial year, only 12.2% of beneficiaries received 69.21% of all payments in 14 EU[1] states. In the case of Spain, 9.3% of beneficiaries received 59.18% of all payments and 18.43% of the Spanish farmers received 75.86% of DP. This is mainly due to the direct or indirect link of direct payments to economic size. Second, the distribution of direct aid payments among animal type and/or crop is very asymmetrical and is spread over several regulations. The most important action under which farmers receive income is the set-aside program and animal premiums. As a result, the percentage of payments which were directed towards arable crops and livestock in the year 2000 amounted to 92.71 and 79.78% of all payments in the EU and Spain respectively. With the information available as to the distribution of direct income payments by crops, it is easy to contrast the unbalanced distribution for Spain: under livestock premiums 185.4 thousand beneficiaries received 974 million Euros while for arable crops 373.9 thousand of beneficiaries received 2.013 million Euros.

For our purposes it is important to point out that the typical holdings with vegetal and animal production can collect DP from several programs. We are interested in those types of farms because of their potential for cross-compliance. In fact, this type of land management provides important opportunities for a trade off between environment preservation and gains in efficiency using intensive techniques. The most obvious alternatives are the choice in the proportions of animal nutrients cropped or grazed on the farm versus feedstuff provided by the industry. There is also the choice between traditional and specialized (intensive) livestock breeds. Note that the spread of epidemics is often strongly related to the ratio of land per animal units. So for example the mad cow outbreak is a result of industrial feeding.

To study these questions, we decided to concentrate on only one country to guarantee certain homogeneity and reduce the variance. Note that the CAP has the same principles in all of the EU states, so a single country can be interpreted as a case study for various EU sates. However, we are aware of the fact that the CAP'92 has some measures which differently apply in different Member States and regions. Spain is an interesting case for its large size (in surface its 7.2 millions of Ha of grazing lands and 3.4 millions of hectares of cultivated lands), export oriented agriculture (more than 55% of exports over the total output), with 37.9% of its agricultural output in livestock raising in 2001 (MAPA, 2004 Anuario de Estadística Agraria 2003), and also because of the availability of large data samples for each type of animal farming. In 2001, Spain accounted for 14.32% of the vegetal output and 11.08% of animal output of the EU15 (Eurostat 2003. European Economic Accounts. SEC-95).

As far as we know, despite of its relevance, few empirical studies counterpoint whether the CAP'92 increments in direct payments were attached to better levels in efficiency or environmental adaptation, or wether they have decreased the asymmetry between large and small holdings, see Bullock and Salhofer (2003). Critics even forecasted a negative incentive to improve productivity and income redistribution (Bryden/Hawkins 1992, Blandford/Dewbre 1994, Bureau 2005). Our aim is to analyze the distortions in efficiency due to CAP subsidies before and after the CAP'92, and ascertain whether the direct payments that farmers received actually resulted in more environmentally friendly farming. We also address the issue of the relationship between efficiency and size as it is theo-

[1] The 15 old Member States except Greece, for which no published data is available. Figures account for the directs aids pay to farmers under the Reg. (EC) 1259/1999. Commission, MEMO/02/198 and AGRI 63569/2002

retically unclear whether subsidies allow inefficient holdings to survive or help them to catch up the production frontier. A second (hypothetical) scenario (calculating efficiency without direct payments) allows us to compare whether (before and after the CAP'92) direct subsidies helped farms to reach the efficiency frontier.

The way of modelling is not neutral to the results; several papers assume functional forms that implicitly accept the existence of constant returns to scale or no limits for the scope of technical change (e.g. Ball et al. 2001, 2007, Jorgenson/Kuroda 1992, Jorgenson/Nishimizu 1978, Jorgenson et al. 1987). For our analysis we always use non- or semi parametric methods when (mis-)specifications could provoke serious disturbances in our conclusions. This greatly increases the econometric effort as well as the variance of our results but avoids any controversy as to the influence of subjective modelling.

One of our key results shows that, on average, absolute direct payments generally tend to increase efficiency (i.e. efficiency in monetary terms) but not productivity (i.e. efficiency in physical terms). However, in most of the cases the mean efficiency decreases as the percentage of direct payments rises. So the CAP'03 idea of *modulation or capping* of direct subsidies in the future will potentially increase the efficiency of the public expenses on DP.

Thus, the implications of this work are important for the future application of the recently approved CAP Reform 2003 on an historical basis. Applying the CAP reform on "historical basis" means translating the unequal distribution of subsidies throughout intervention prices and direct subsidies into a single payment to each farm. 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 previous year's level of protection. Thus potential implications of efficiency will affect agricultural competitiveness and have to be carefully analyzed.

The rest of the paper is organized as follows. In the next section we introduce the data and methods we will use for our study. After that we dedicate a large section to the presentation of the numerical results and their interpretations. Finally, we conclude. Additional information about the used data, numerical results and technical details of the procedures are given in the appendices of this article.

2. Data and procedure

The sample was obtained from the Farm Accounting Data Network (FADN) which provides homogeneous information for farms and classifies them in types of animal farming with positive plant production. We concentrate on cross sectional analysis for the years 1991, 1992 (before CAP'92 reform) and 1999, 2000 (after CAP'92 enforcement). We have chosen these years aiming to have data of average weather conditions and sufficiently delay to allow capturing the changes by the complete enforcement of the reforms and farmer reactions to the new policy environment. The FADN survey provides detailed information on input expenditures by farm. For the selected farms, livestock production (meat and animal products) is always greater than plant output (fodder, field crop, grain cereals, vineyards, potatoes, industrial crops, plants, fruits, dried pulses, olive groves and others) to ensure that we only include farms oriented to livestock raising. Plant production however is always positive in the selected sample to ensure that the production function remains homogenous by type of animal farming. We only want to include farms with similar production functions, e.g. oriented to animal production, but also with the possibility of harvesting plant products for re-use on the farm or for sale. We consider that farms with no-land (not even rented land) have a non-comparable production function and will therefore be excluded. For technical reasons when using DEA (Data Envelopment Analysis) the identification of an efficiency frontier is only possible if the individual production function of the farms is similar. Finally, as we are interested in the impact on efficiency at a farm level, instead of using aggregated data we use individual farm accounting data that include any kind of direct payment received.

We define the production function of farms with the two outputs and five inputs, listed in Table 1, in monetary values at the current prices as reflected in the farm accounts. Further details about the data can be found in the appendix, like the specific number of farms in each year and the proportion of farms without DP. There are also described some of the enumerated variables in more detail.

Table 1 Variables used in the DEA

OUTPUTS	pbveg	plant output
	pbanim	animal output
INPUTS	capital	capital, especially buildings and machinery
	costsg	fodder and other animal linked inputs
	costs	inputs crop linked (fertilizer, agro-chemicals, seeds, water and other crop specific inputs, fuels and lubricants)
	salary	wages
	land	Agricultural Utilized Area of farm aggregate adjusted for quality (thus including pasture and agricultural land adjusted for quality)
	SP	shadow price; the costs for producing without Direct Payments.

The efficiency cannot be directly observed and must be estimated in a first step by DEA. In the Appendix we give a brief introduction to the DEA method explaining in detail its exact definition and indicating how it is calculated in practice. We include here some useful remarks to understand the basic ideas of the procedure.

As animal and plant outputs cannot be substituted, we must not aggregate them, but estimate efficiency as a two dimensional output problem. There are two reasons why we preferred not to consider direct payments as a third output: First, many farms have zero subsidies and would thus form a non-interpretable hyper-plane in the DEA; furthermore, the subsidies are not actually produced by the inputs considered, so there is no reasonable argument for allowing them to come along on the left hand side of the production function. Alternatively, in DEA, including a variable as output or as negative input will give the same interpretation for the efficiency. Moreover, DP as negative inputs can be understood as including shadow prices representing the costs paid for not producing in a subsidized manner.

It could be discussed whether more input variables should be allowed to enter into the DEA production function, in order to get stronger results with respect to larger differences in the efficiency index for example. However, this question is nothing more than a discussion of the bias – variance trade off: more aggregating leads to more bias but less variance and vice versa. We have opted here for high resolution, in other words, high variance, small bias and thus, none of our results will suffer errors due to possible misspecification.

Until now we have used the word efficiency for both, economic efficiency and productivity. For a correct interpretation it will be helpful to distinguish them. This will be done as follows. In a first scenario we calculate efficiency with direct payments, in a second scenario without. We designate as E_W the efficiency without direct payments which corresponds to productivity, and E_{DP} to conventional (economic) efficiency which includes direct payments or, in other words, "to crop subsidies" (i.e. producing in a more subsidized way).

The economic behavior of the farm under CAP is a trade off between the choice of agricultural productions with a certain level of subsidies and other non-subsidized outputs. Thus, the differences we see when looking at E_W vs. when looking at E_{DP} indicate the level of efficiency distortion on the economic behavior of the farm. Under the current multifunction farming, on top of the income for selling on the market, farmers qualify for rewards to compensate market failures by pricing positive externalities. So the farmer incurs an opportunity cost of not meeting the conditions to receive DP (e.g. minimum land per livestock head).

We repeat efficiency calculations for two years before and two years after CAP'92, to test for whether the CAP reform had promoted efficiency and environmentally friendly practices. We took always two years to take into account the influence of random weather variability (e.g. pasture availability). Furthermore, the sample is divided into type of animal farming (cattle farming, pig farming, and sheep and goat) as these farms are neither uniform in the treatment by CAP nor in the production process. Summarizing: we carried out both estimations (with and without DP) for 1991, 1992, 1999 and 2000 for cattle, pig, and sheep and goat farms.

Once the efficiency indices are calculated, we use these results to analyze the changes caused by the CAP'92 reform by different methods (correlations and semi parametric regression). Key variables of our study now are the size measured by European Size Units (ESU) and a proxy for measuring how environmental friendly (EF) the farm is. As we analyze the subsidies policy here, for a fair evaluation, we use the same proxy the EU generally uses, i.e. livestock unit equivalents per agricultural utilized area, EF = (LU/AUA).[2] The LU/AUA is used in the European regulation and is generally considered a good proxy of the environmental performance of the farm, see also remarks in the Introduction. As we mentioned in the introduction, one could also say *animal adapted* but we use this synonymously for environmental adapted because of various reasons, among others: the animal density is proportional to the nitrogen production; on average, the extensive farms (low animal density) generate positive external economies like the preservation of the natural ecosystem. Note that EF is certainly inverse proportional to being environmental adapted. As the policy faces different targets simultaneously such as productivity, cross-compliance, and small farm protection, we need more sophisticated instruments to contrast these objectives.

More specifically, we use two regression models to study the level of compatibility between different targets of the new CAP, including environmental adaptation and competitiveness at the farm level. The objective is to quantify the impact on efficiency when the CAP'92 increased the direct payments. Our model has efficiency as the dependent

[2] 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.

variable and the explicative variable is environmental adaptation, filtering out the regional and size effects:

$$E = g\left[\ln(EF), \ln(ESU)\right] + \beta^T R + e \ , \tag{1}$$

where E is first economic efficiency E_{DP} (efficiency with direct payments), then conventional efficiency (productivity in monetary terms) E_W (efficiency without direct payments). For a more detailed explanation see Kleinhanß et al. (2006) where similar methods are used though for a different study comparing countries in one subsidy system instead of comparing systems (before and after the CAP'92) for one country. *EF* indicates the degree of animal density, *ESU* is the European Size Unit, and *R* is a vector of dummy variables for agricultural region divided into North, Center, Northeast, South and East. Recall that we use *EF* as an inverse proxy for animal adaptation (or environmental friendliness, see discussion above). Note that the smaller the EF, the more "environmentally friendly" the farm. In the next section it will be seen how the comparison of these two regressions (i.e. one with E_{DP}, one without subsidies E_W) helps us to better understand the impact of subsidy policy in practice.

Note that the function $g \colon \mathbb{R}^2 \to \mathbb{R}$ is nonparametric, i.e. not specified further. As we will see, the impact of $\ln(EF)$ and $\ln(ESU)$ show strong nonlinearities and severe interactions. The term e stands for the not further specified heterogeneity. As $g(\cdot, \cdot)$ is non-parametric, the logarithm does not impose any model specification here. This transformation is only due to smoothing necessities, see Appendix for further details.

3. Empirical results, interpretation and comments

3.1. Calculation of efficiency and productivity with DEA

We first calculate efficiency and productivity with the aid of DEA. These results will be used for most of our further conclusions relating them with different economic and policy factors. In this sense it is mainly an auxiliary step. As a byproduct, based on these results we are also able to check the effect of certain agricultural extension programs on efficiency (and productivity). Note that we will not separate the two scenarios (with and without subsidies) into two subsections because we are not so much interested in the result of each individually but in the differences between them.

In order to address the question "what are the subsidies related to in practice?" we must first clarify the question of modelling.

The correlation target is to quantify the relationship between productivity (E_W), the farm size, the animal density and the fact that the farmer qualify for direct payments. To account of both, absolute and relative mean increases, we have estimated the following correlations and their p-values: $corr(DP, E_W)$, $corr(\ln(1+DP), E_W)$, $corr(DP, ESU)$, $corr(\ln(1+DP), \ln(ESU))$, $corr(DP, EF)$, $corr(\ln(1+DP), \ln(EF))$. It is conspicuous that efficiency calculated with DP will be (positively) related to DP. Therefore we considered here only E_W wich we call simply productivity. The numerical results for these correlations can be found in the appendix, Tables A3 to A5, separated only by animal-type.

For an easier interpretation, let us briefly summarize the signs we see in the tables. In Table 2 we have summarized first the pair $\{corr(DP, E_W), corr(\ln(1+DP), E_W)\}$ in the first two lines, the signs of the pair $\{corr(DP, ESU), corr(\ln(1 + DP), \ln(ESU))\}$ in lines 3 and 4, and of the pair $\{corr(DP, EF), corr(\ln(1 + DP), \ln(EF))\}$ in the last two lines.

Table 2 Signs of correlations: the pair $\{corr(DP, E_W), corr(\ln(1 + DP), E_W)\}$ in lines 1,2; pair $\{corr(DP, ESU), corr(\ln(1 + DP), \ln(ESU))\}$ in lines 3,4; pair $\{corr(DP, EF), corr(\ln(1 + DP), \ln(EF))\}$ in lines 5,6.

	cattle	pig	sheep and goat
Before CAP'92	0 –	0 –	0 –
After CAP'92	+ +	– –	+ –
Before CAP'92	+ +	+ +	+ +
After CAP'92	+ +	+ +	+ +
Before CAP'92	– –	0 –	+ –
After CAP'92	– –	– –	– –

As to the relation between subsidies and productivity we find that before CAP'92 there is no positive impact of subsidies on productivity (E_W), see Table 2. Looking at the effect of relative subsidy increase, the impact is even negative. Something similar happens during the period following CAP'92, except in the case of cattle farms. However, the impact of absolute increase of subsidies on productivity is positive for cattle, and sheep & goat, but negative for pig farms after '92. This may support the argument that subsidies improve productivity but evidently in a quite regressive way. This means that a combination of direct subsidies and economic size (ESU) would be counterproductive. Comparing with the numerical results, say levels, given in Tables A3 to A5 in the Appendix, we see that the relation between DP and productivity (E_W) hardly changes with CAP'92 in any type of farm with almost all of them being close to zero. When subsidies rise in relative terms we detect significant changes only for cattle farms.

Turning to the relationship between direct payments and economic size, we see that this is always (clearly) positive, see Table 2. This means that no matter whether we measure in absolute or relative terms, the policy always benefits the larger farms more strongly. This policy did not change with the CAP'92. Looking at the tables in the Appendix, the correlations even strongly increase for cattle and sheep & goat after the CAP'92, actually over 90% in some cases. In other words, since the CAP'92, the level of subsidies can mainly be linked to farm size. That is not surprising since set-aside payments and animal premia are related to the area and number of animals respectively. So, after CAP'92, direct payment correlation with farm size shows the level of real "modulation" of the post CAP'92 subsidies. Thus, our results are congruent with the generally accepted hypothesis, see e.g. OECD (2001) and references therein, that direct subsidies are basically (even if indirectly) linked to output level especially after CAP'92. But therefor it is difficult to defend the presumable decoupled characteristic of these aids. Moreover, our findings indicate a strong coupling of size and premia since CAP'92 what seems to be counterproductive, see the last paragraph.

Regarding the relationship between DP and EF (recall that the higher the EF the more intensively), we see no change of signs for cattle farms (always negative), but some for pigs, and sheep & goat farms. There, along the detected signs the relationship between subsidies and extensive farming has increased after CAP'92, compare again Table 2. In particular, to see more detailed effects of the supporting extensive farm claimed policy target, we compare Tables A3 to A5. Indeed, the results differ somewhat on the measure, see e.g. Table A4, in sign as well as in the level. Therefore we will use a more sophisticated approach to deal with this problem (see the following nonparametric regression study).

In summary, evaluating the policy reform along the targets sketched in the introduction, we can say so far that:

- productivity at firm level is indeed biased due to the policy reform.
- (but) the farms are affected in different degrees by the introduction of DP, depending on the type of farm.
- in particular, productivity is positively correlated with DP in absolute, but negative in relative terms. This does not contradict literally the claimed objectives (DP increase productivity) but the practice contradicts their spirit (larger farms get more DP what is counterproductive).
- farm size is strongly positively correlated with DP, after CAP'92 more than before, what contradicts the idea of helping small farms.
- environmental friendliness is positively but weakly correlated with DP.
- the target of supporting extensive farming (i.e. animal adapted farms), we do not find as clear a trend as we saw for economic size.

3.2. A nonparametric regression analysis of the efficiency

For a further analysis we need to relate the efficiency and/or productivity to size (ESU), *EF*, and regions via a proper regression model. For this purpose we consider now equation (1) with E being E_{DP} (*DP* included in model) or E_W respectively (*DP* not included in the model). By comparing the results of these two regressions (i.e. using two different dependent variables), we will see how the CAP policy distorts the efficiency of an individual farm. The estimation procedures applied here are explained in the Appendix.

First, let us look at the regional effects before and after CAP'92, i.e. the estimates of the β-parameter in regression model (1). We split up Spain into 5 regions: North, Center, Andalusia, Ebro (along the Ebro river), and Levante. The last one has been employed as a normalizing region. Note that Levante could correspond for "East", and Andalusia for "South". Ebro stands for the northeastern Spanish region including the northeastern Mediterranean coast and the Ebro river valley both with a chiefly Mediterranean climate that traditionally has been conceived them as an homogenous agricultural region. All numerical results are given in Tables A6 and A7 of the appendix.

Before CAP'92 (Table A6): The North was slightly less efficient in cattle farming whereas the center is more efficient than other regions. In pig farms as well as with sheep and goats, all regions seem to be very close except for the ones in Levante.

After CAP'92 (Table A7): The North and Ebro regions seem to be less efficient. These results are only insignificant for pig farming in 2000 and in 1999 when looking at E_W (productivity). Levante seems to be best for cattle farming, though not significantly better than the center and the south. These aspects scarcely alter in both, the model with dependent variables E_{DP} and its counterfactual opposite, using E_W.

These findings might intuitively surprise experts in Spanish animal farming. However, there are actually several explanations for these findings. First we must point out that in Levante we have less than 10 farms in the sample, which turned out to be technically quite advanced holdings (sample effect). Next, it is important to know that even though the North is the so called "green Spain", the rather poor farms are situated right there in the mountains whereas in the center we have hardly cattle farms in mountain areas. This is different for sheep & goats. Finally, cattle farms are mainly milk producer, a product that has to be processed rapidly (different from meat and wool). It is well known that factories (farms in our case) of those kinds of products are in average technically more advanced close to huge consumption areas like Madrid. This may also explain why cattle farms in the center are in average more productive and efficient. This argument would not hold for pig or sheep & goats holdings what is indeed in accordance with our findings.

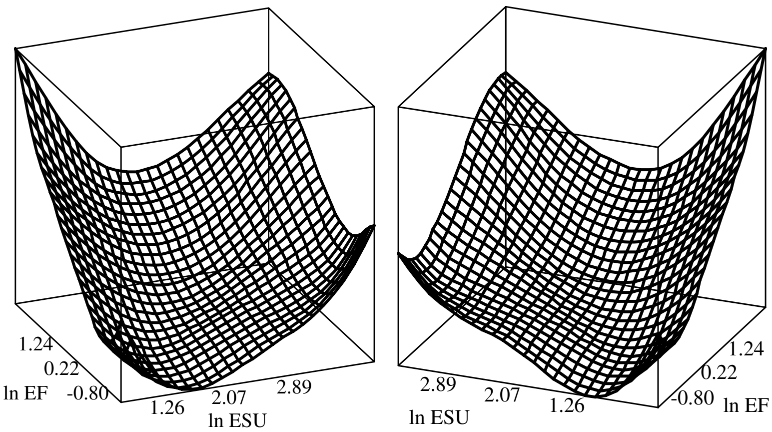

Figure 1 Cattle farms in 1991, with dependent variable E_{DP} (left), respectively E_W (right)

As the functional form of $g(\cdot,\cdot)$ in model (1) is non-parametric, the results are given graphically, see Figures 1 to 7 where we displayed the numerical results for $g(\cdot,\cdot)$. Note that $g(\cdot,\cdot)$ is a three dimensional graph throughout. Along our empirical results, the two regressors $\ln(EF)$, $\ln(ESU)$ unfortunately do not enter additively, i.e. we can not write $g\{\ln(EF), \ln(ESU)\} = g_1\{\ln(EF)\} + g_2\{\ln(ESU)\}$. That means, as can be seen in Figure 1, that it is not enough to look at the purely marginal effects of $\ln(EF)$, and $\ln(ESU)$ respectively, because they have strong interaction. For example, in Figure 1 is plotted function $g(\cdot,\cdot)$ (vertical axis) on its arguments $\ln(EF)$, $\ln(ESU)$. There we can see that for small farms with $\ln(ESU) \approx 1.0$ the effect of $\ln(EF)$ on efficiency and productivity as about 50% stronger than for middle-sized farms with $\ln(ESU) \approx 1.9$. For any given $\ln(EF)$ the impact of size is always U-shaped although not symmetric, and the steepness of the U-borders changes over $\ln(EF)$.

For ease of presentation we decided not to show the three dimensional graphs but present the marginal impacts of $\ln(EF)$ (on E_{DP} and E_W) of the median sized (measured in ESU) farms (solid line), the large farms (the upper 95% quantile farms with respect to ESU, dotted line), and the small farms (the lower 5% quantile farms, dashed line). Note that when looking at marginal impacts, the single impact of one variable on efficiency can be greater than one or also be negative. The three resulting (2 dimensional) functions represent three slices of the three dimensional plot. For a better understanding compare the three dimensional plot in Figure 1 with its three slices in Figure 2. In all these graphs, the outer 2% are cut off (i.e. not plotted) to avoid interpreting the boundary effects.

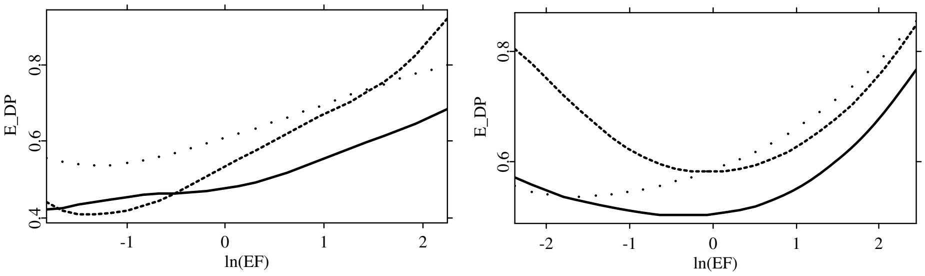

Figure 2 Cattle farms in 1991 (left) and 1992 (right), E_{DP} on ln(EF). median sized farms: solid lines; large farms: dotted lines; small farms dashed lines

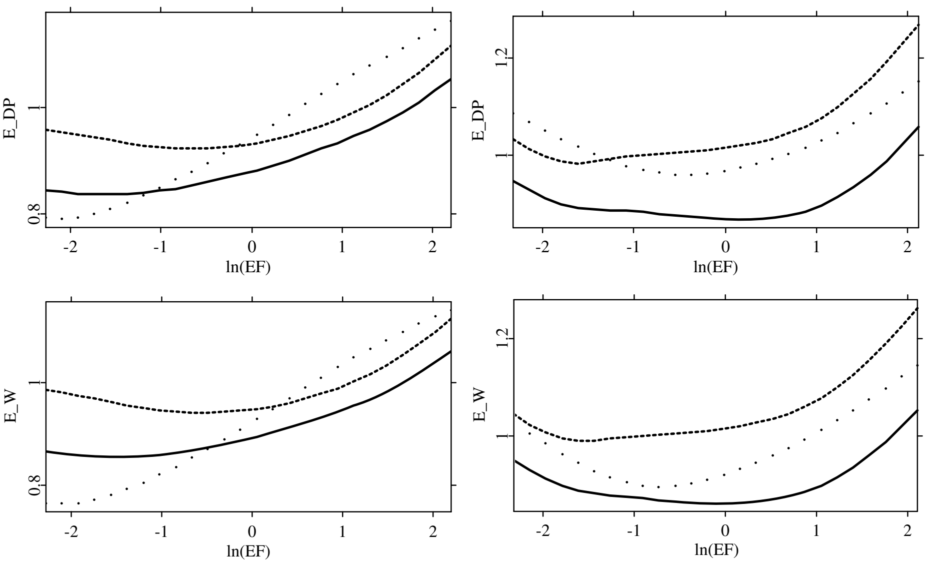

Figure 3 Cattle farms in 1999 (left) and 2000 (right), E_{DP} on ln(EF) (upper row), and E_W on ln(EF) (lower row). median sized farms: solid lines; large farms: dotted lines; small farms dashed lines

Let us have a look at the differences between the regression of E_{DP} compared to the regression of E_W. Before the CAP'92 reform (Figures 2, 4, 6) we could not find differences between the two regressions, and therefore have given here only the results for E_{DP}. This indicates that there was no effect of direct payments on efficiency versus productivity.

This changes with the CAP'92 reform. For cattle farms (Figure 3) we again obtain the same results for the two regressions, whereas for pig holdings the DP now favor the extensive and in particular the large holdings (compare Figure 5 left with right side). Since DP do not exist per pig head, this is possibly due to DP related to crop and environmental issues which would explain that the large farms benefit especially. E.g. Iberian pigs grazing in the Mediterranean forest need large plots of land for grazing but are quite profitable

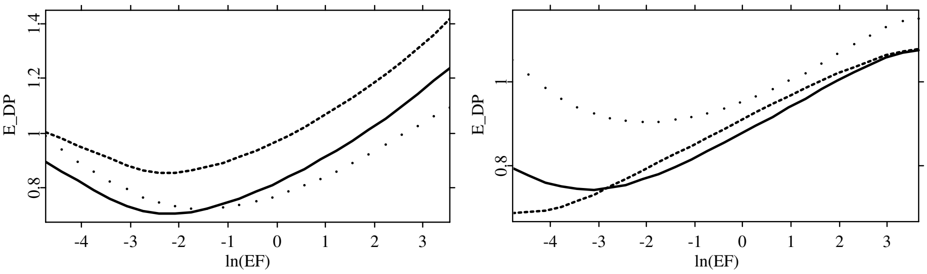

Figure 4 Pig farms in 1991 (left) and 1992 (right). median sized farms: solid lines; large farms: dotted lines; small farms dashed lines

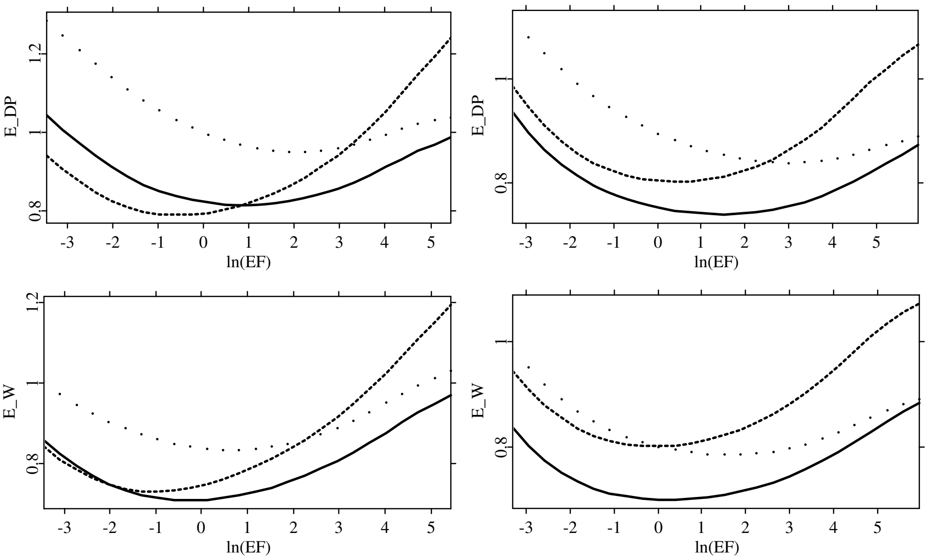

Figure 5 Pig farms in 1999 (left) and 2000 (right), E_{DP} on ln(EF) (upper row), and E_W on ln(EF) (lower row). median sized farms: solid lines; large farms: dotted lines; small farms dashed lines

due to the high prices their meat yields on the market. Also for sheep & goat farms, we see after the CAP'92 that now including DP in the efficiency changes the regression, see Figure 7. For any size of farm it seems that the DP especially favor farms that are identified as being intensive holdings (right tail).

Comparing the results from before with those after the CAP'92 reform we detect that extensive farms are not better situated (compared to intensive holdings) in terms of relative efficiency levels. Also the efficiency rankings by size remain unchanged by the '92 reform. In any case, it is hard to make clear statements because the results vary greatly with the years. Focusing only on the median farms, one might say that after the CAP reform the efficiency difference between intensive and extensive farms has become smaller. The scenario without direct subsidies upholds all our results.

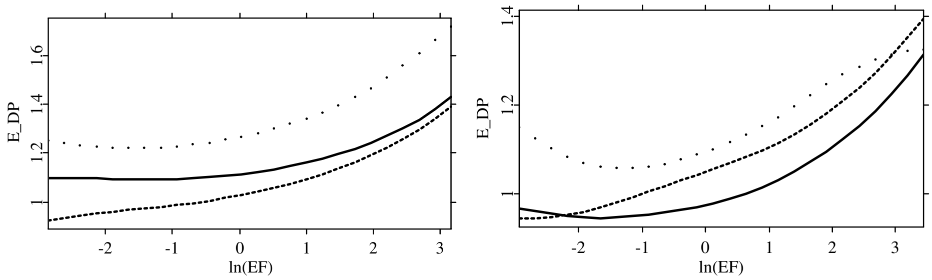

Figure 6 Sheep and goat farms in 1991 (left) and 1992 (right). median sized farms: solid lines; large farms: dotted lines; small farms dashed lines

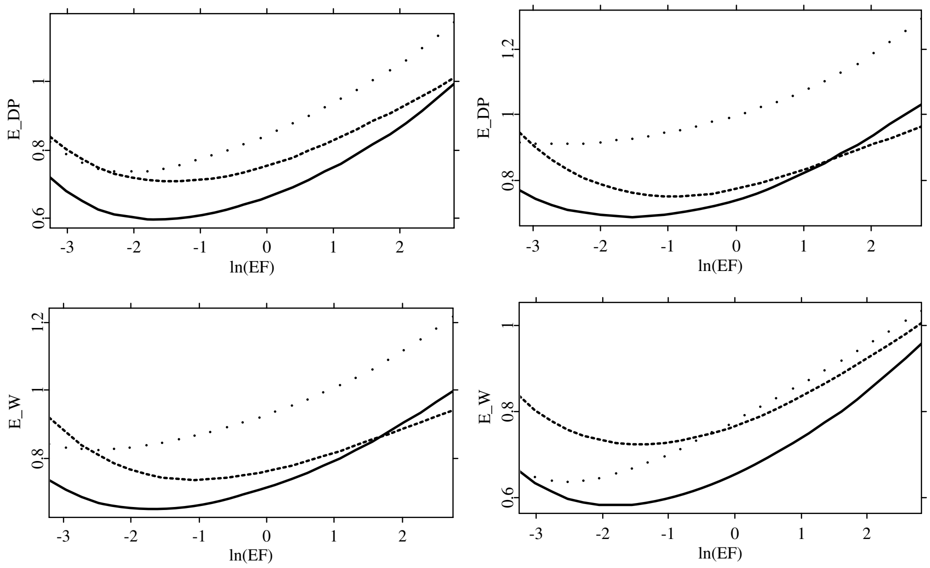

Figure 7 Sheep and goat farms in 1999 (left) and 2000 (right), E_{DP} on ln(EF) (upper row), and E_W on ln(EF) (lower row). median sized farms: solid lines; large farms: dotted lines; small farms dashed lines

The efficiency of the intensive pig farms is clearly above that of the environmentally friendly farms before, and also after the CAP reform. However, after '92 the distance to the more environmental friendly farms has also become smaller, i.e. the plotted functions have flattened. The scenario ignoring the subsidies when calculating the efficiency, upholds these results as well.

Finally, when looking at the sheep & goat we find the following. Before CAP'92 there is a clear positive impact of ln(EF) on efficiency (both, E_{DP} and E_W) for any size of farm, whereas this impact becomes strongly U-shaped after CAP'92. This means that extensive holdings are now relatively better off than before. However, this statement is only true for small and median sized farms. Furthermore, the rankings by size changed after '92 putting the small farms in a better situation than before. In any case, the large, intensive farms remain at all times the most productive and efficient ones even though these differences narrowed after CAP'92.

Comparing the results found in the graphs explicitly with the initial hypotheses discussed in the introduction, we can summarize that:

- a distortion of efficiency caused by DP is visible in our plots.
- there is evidence that less environmentally friendly farms (intensive holdings) are more productive and more efficient regardless the DP.
- however, after CAP'92, for some type of animal farms extensive holdings are a little bit better off (relatively to the intensive ones) thanks to DP.
- the "efficiency and productivity ranking" changes more over the different years than when comparing before and after CAP'92.
- the strong DP of the biggest farms cannot be empirically justified with their hypothetical above average productivity or low animal density.

4. Conclusions

The results clearly show the real difficulties of the reformed CAP in reaching the objective of increasing the productivity (with special focus on the preservation of small farms because of the cross-compliance argument) and the environmental orientation of the farms simultaneously. The empirical evidence shows a positive correlation between subsidies and productivity when looking at the absolute amounts. However, the mean productivity decreases or stagnates as the percentage of direct payments rise. This means that a combination of direct subsidies and size is counterproductive. Our results also show that this is what the subvention policy after CAP'92 was doing much more than before. The heavy subsidization of large farms cannot be justified with their presumably high productivity nor environmental friendliness. Neither our graphical nor our numerical results confirm the hypothesis that DP, in particular DP per size, encourage productivity or environmental friendliness; they may, in fact, even contradict.

We have found that economies of scale seem to be important before and after CAP'92, but we have found several exceptions, see e.g. the pig holdings. The small intensive pig farms performed rather well before and after CAP'92. Also the small cattle farms are above the mean efficiency index. This reflects the shortcomings of the managers in running the farm and solving technical problems when the economic size increases above a certain threshold, a finding that concords with Alvarez and Arias (2003).

After CAP'92 the environmentally friendly hog holdings reach efficiency levels similar to those of the most conventional (intensive) farms when accounting for direct payments. This was not the case before CAP'92. In contrast, for the cattle farms the order of efficiency between small to big and intensive to extensive farms does not change with DP, neither before nor after the CAP reform. For sheep & goats finally, we cannot detect an effect of DP before, and only a marginal one after CAP'92. Small extensive farms seem to benefit a little bit since the reform what would be in accordance with one of the claimed objectives.

However, all together the less environmentally adapted farms are still the most efficient ones. Thus, together with the numerical results on correlations, we conclude that there is empirical evidence that the subsidy system after CAP'92 was somewhat more "environmentally oriented" for all types of farms studied than it was before CAP'92. Notice further that our results show a significant positive correlation of the DP and EF, i.e. direct payment were negatively correlated with animal density, for the years after CAP'92. That makes a difference with the situation before the CAP reform where we can not find a significant correlation between direct payments and environmental adaptation.

So the conclusion that the DP structure after '92 played a role in helping farms to preserve the natural environment under competitive conditions is empirically supported by our results. Nevertheless, we have found also that the DP system introduced by CAP'92 was not sufficient to correct the fact that the less environmentally friendly farms as well as the big ones are more efficient. Moreover, the empirical results show especially for sheep & goat and cattle farming that there is room for improvement of the environmental implementation of the CAP.

Summarizing, our comparison of the situation before with that of after the CAP'92 reform, cannot lead us to conclude that the claimed objectives would have been reached; partly the effects have gone in the right direction, partly not. Neither the small nor the more animal adapted farmers benefit from the subsidy policy whereas the large ones mainly do. This, moreover is counterproductive what is in contradiction with the claimed objectives of the subsidy policy. The changes we have observed for Spain due to the CAP'92 are marginal, although visible in some cases, concerning the claimed objectives, but strong concerning the non-claimed objectives (e.g. linked to economic size).

As for the general assessment of the CAP reform it is important to remark that future tools must guaranty an effective link between the declared objectives and the ex-post results of the policy which is not the case for the analyzed period. In particular, environmental adaptation of farms needs to be promoted in a more effective way. The same recommendation holds for the promotion of the efficiency for smaller farms, an objective systematically repeated and never achieved.

Finally, the empirical results ask for future research using similar methodology for evaluating the 2003 CAP reform, now on the way to be enforced in the EU Member States. Specifically, the complete substitution of the price support by the so called "single payments" per farm, (which in fact is a single direct payment) will potentially affect in a significant way the different types of farm. It will be highly interesting to continue the research in this direction once the new data were available. The here used methodology, in our opinion, has proven to be rather helpful for such a policy evaluation. It is quite flexible, extremely robust against model specifications, and facilitates clear interpretations.

Appendix I: Data and further empirical results

In Tables A1 and A2 are summarized the number of farms used for all our estimations, separated by year and animal type.

In the list of variables used for the DEA, i.e. for calculating the efficiency indices, SP, land, and capital are defined as follows:

The so called "shadow price" (SP) is calculated from the total amount of direct payments a farm receives (including premiums), denoted by DP, but with a negative sign: therefore we set SP = -DP. Direct payments include any amount of cash received under the CAP or national regulation and not linked to an amount of production (e.g. set aside

Table A1 Number of farms used for all our calculations

	'91	'92	'99	'00
sheep and goat	391	373	553	679
cattle	2230	1787	1435	1543
pig farming	126	161	255	249

Table A2 Number and percentages of farms used for all our calculations not receiving direct payments

	in absolute numbers				in percentages			
	'91	'92	'99	'00	'91	'92	'99	'00
sheep and goat	82	35	0	0	21.0	9.4	0.0	0.0
cattle	1373	1198	439	230	61.6	67.0	30.6	14.9
pig farming	72	66	22	17	57.1	41.0	8.6	6.8

payments, livestock premia, etc.). Capital input includes machinery, transport equipment and structures (not dwellings). For simplicity all animals are treated as variable stock. Land is considered as a separate input variable. Land area is adjusted by quality. This means that we calculated the value of input "land" by

$$\text{land value} = \text{sauirr} * \text{psauirr} + \text{saudry} * \text{psaudry}\ ,$$

where sauirr: Agricultural Utilized Area (AUA) irrigated (Ha.); saudry: AUA non-irrigated (Ha.); psauirr: price AUA irrigated by region (Euros/Ha); psaudry: price AUA non-irrigated by region (Euros/Ha).

Appendix II: Methodologies

Even though these methodologies are not completely new, part of the readership might not be familiar either with DEA or with nonparametric regression. Therefore we give here a brief insight in order to achieve a better understanding and interpretation of the results presented later.

Estimation of (technical) efficiency

Data envelopment analysis (DEA) is a non-parametric approach for evaluating the performance ("the best practice") of a set of peer units called DMUs (decision making units) by using linear programming methods. We introduce here some basic concepts of the method (see Cooper/Seiford/Tone (2000) for a complete description of the methodology).

Let us consider an economic sector where firms produce q outputs with p inputs in which we may define, following Simar and Wilson (2000)'s notation, the next set of feasible input-output combinations:

$$\Psi = \{(x, y) \in I\!R^{p+q} : \quad \text{x can produce y}\}. \tag{A1}$$

Table A3 Correlations for Cattle farms; $\underline{\text{underlined}}$ indicates not significant at 5% level

Cattle	'91	'92	'99	'00
$corr(DP, E_W)$	$\underline{-0.023}$	0.050	$\underline{0.011}$	0.135
$corr(\ln(1+DP), E_W)$	-0.158	-0.059	0.110	0.090
$corr(DP, ESU)$	0.100	0.254	0.657	0.616
$corr(\ln(1+DP), \ln(ESU))$	$\underline{0.024}$	0.085	0.226	0.164
$corr(DP, EF)$	$\underline{-0.012}$	-0.115	-0.127	-0.181
$corr(\ln(1+DP), \ln(EF))$	-0.079	-0.166	-0.185	-0.196

Table A4 Correlations for Pig farms; $\underline{\text{underlined}}$ indicates not significant at 5% level

Pigs	'91	'92	'99	'00
$corr(DP, E_W)$	$\underline{0.011}$	-0.310	$\underline{-0.053}$	-0.146
$corr(\ln(1+DP), E_W)$	-0.208	-0.460	-0.309	-0.237
$corr(DP, ESU)$	$\underline{0.133}$	0.309	0.489	0.168
$corr(\ln(1+DP), \ln(ESU))$	$\underline{0.110}$	0.298	$\underline{0.119}$	0.028
$corr(DP, EF)$	0.301	-0.196	$\underline{-0.089}$	-0.143
$corr(\ln(1+DP), \ln(EF))$	$\underline{-0.119}$	$\underline{-0.114}$	-0.131	-0.384

Table A5 Correlations for sheep and goat farms; $\underline{\text{underlined}}$ indicates not significant at 5% level

Sheep and Goat	'91	'92	'99	'00
$corr(DP, E_W)$	$\underline{0.034}$	$\underline{-0.048}$	$\underline{0.048}$	0.092
$corr(\ln(1+DP), E_W)$	$\underline{-0.057}$	-0.160	-0.205	-0.096
$corr(DP, ESU)$	0.223	0.298	0.968	0.973
$corr(\ln(1+DP), \ln(ESU))$	0.099	0.159	0.596	0.571
$corr(DP, EF)$	$\underline{0.014}$	$\underline{0.023}$	$\underline{-0.082}$	-0.086
$corr(\ln(1+DP), \ln(EF))$	$\underline{-0.072}$	$\underline{-0.090}$	-0.127	-0.130

For any y we specify the input requirement set as

$$X(y) = \{x \in I\!R^p : (x, y) \in \Psi\}, \tag{A2}$$

where the input efficient frontier is defined by:

$$\delta X(y) = \{x \in X(y) : \theta x \notin X(y) \quad \forall \theta < 1\} . \tag{A3}$$

Efficiency measures for each firm (Farrell 1957) $\theta(x, y)$ are then obtained as the following maximum contraction of inputs along a fixed ray:

$$\theta(x, y) = \inf\{\theta : \theta x \in X(y)\} . \tag{A4}$$

Note that in our text θ is E (Efficiency with or without DP). A value of $\theta = 1$ means that the producer is input efficient while a value of $\theta < 1$ indicates technical inefficiency and

Table A6 Estimates (upper lines) with standard error (lower lines) for the regional effects before CAP'92. In 1991 we have neither pig nor cattle farms observed in Levante, the reference region is Ebro, or else Levante. In 1992 we have no cattle farms observed in Levante, the reference region is Ebro, or else Levante

	1991				1992			
dep.var.	North	Center	Andal.	Ebro	North	Center	Andal.	Ebro
				Cattle				
E_{DP}	-0.046	0.174	-0.029	-	-0.053	0.139	-0.058	-
	0.029	0.029	0.038	-	0.029	0.029	0.033	-
E_W	-0.050	0.165	-0.028	-	-0.051	0.111	-0.054	-
	0.028	0.028	0.037	-	0.028	0.028	0.033	-
				Pig				
E_{DP}	-0.056	-0.006	-0.135	-	-0.084	-0.167	-0.017	-0.101
	0.156	0.038	0.064	-	0.135	0.125	0.145	0.122
E_W	-0.054	-0.004	-0.133	-	-0.079	-0.211	-0.069	-0.120
	0.156	0.038	0.065	-	0.132	0.122	0.141	0.119
				Sheep and Goat				
E_{DP}	-0.455	-0.462	-0.459	-0.452	-0.357	-0.342	-0.335	-0.374
	0.169	0.162	0.170	0.162	0.189	0.178	0.181	0.179
E_W	-0.474	-0.472	-0.456	-0.470	-0.378	-0.390	-0.382	-0.399
	0.164	0.158	0.165	0.158	0.183	0.172	0.176	0.174

the producer may reduce inputs in that proportion while upholding the output level.[3] As the model is non-parametric, the estimation of all the above unknown concepts by DEA requires to assume convexity and free disposability of inputs and outputs for the production possibility set, see Färe, Grosskopf and Lovel (1994) for a more detailed description of the characterization of the technology. So the estimate of equation (A1) under the least restrictive returns to scale assumption (i.e. variable returns[4]) for a sample of n producers is:

$$\widehat{\Psi} = \{(x, y) \in \Re^{p+q} : \ x \geq \sum_{i=1}^{n} \gamma_i x_i, \quad y \leq \sum_{i=1}^{n} \gamma_i y_i, \quad \sum_{i=1}^{n} \gamma_i = 1, \quad \forall \gamma_i \geq 0\}, \qquad \text{(A5)}$$

where: γ_i is the intensity vector of firm i and it defines its *best practice or benchmark firm* by a linear combination of all the firms observed in the sample. Note that here we are using (x, y) as a context variable and not with the same meaning as in Section 6 above.

3 Alternatively, one could formulate (A3) to (A4) as an output oriented problem. However, in practice, the interpretation is then often more complicated, in particular considering how to include direct payments in the production function.

4 The assumption of variable returns to scale is suitable when not all firms are operating at the optimal scale and it ensures that an inefficient firm is only "benchmarked" against firms of similar size.

Table A7 Estimates (upper lines) with standard error (lower lines) for the regional effects after CAP'92. The reference region is Levante

	1999				2000			
	North	Center	Andal.	Ebro	North	Center	Andal.	Ebro
				Cattle				
E_{DP}	-0.462	-0.219	-0.376	-0.491	-0.419	-0.179	-0.359	-0.407
	0.104	0.104	0.115	0.106	0.089	0.089	0.100	0.091
E_W	-0.467	-0.254	-0.330	-0.506	-0.423	-0.200	-0.342	-0.403
	0.102	0.102	0.113	0.104	0.088	0.088	0.099	0.090
				Pig				
E_{DP}	-0.346	-0.061	0.083	-0.093	-0.033	-0.016	0.208	-0.046
	0.158	0.045	0.083	0.037	0.110	0.058	0.091	0.047
E_W	-0.215	-0.016	0.123	-0.041	-0.095	-0.020	0.057	-0.043
	0.157	0.045	0.083	0.037	0.104	0.055	0.086	0.044
				Sheep and Goat				
E_{DP}	-0.120	-0.023	-0.090	-0.138	-0.105	-0.073	-0.232	-0.216
	0.064	0.041	0.055	0.045	0.051	0.030	0.037	0.034
E_W	-0.146	-0.033	-0.099	-0.185	-0.097	-0.067	-0.218	-0.263
	0.059	0.038	0.051	0.042	0.049	0.029	0.035	0.033

Equally, estimates of equations (A2) and (A3) are then

$$\hat{X}(y) = \{x \in \Re^p \mid (x, y) \in \hat{\Psi}\}, \quad \delta\hat{X}(y) = \{x \in \hat{X}(y) \mid \theta x \notin \hat{X}(y) \ , \forall\theta < 1\}, \tag{A6}$$

while the efficiency measure (equation A4) is estimated by linear programming techniques as follows:

$$\hat{\theta}(x_j, y_j) = \min\{\theta : \sum_{i=1}^{n} \gamma_i x_i \leq \theta x_j \ , \quad y_j \leq \sum_{i=1}^{n} \gamma_i y_i \ , \quad \sum_{i=1}^{n} \gamma_i = 1 \ , \quad \forall \gamma_i \geq 0\}. \tag{A7}$$

Firm j is technically efficient if and only if $\hat{\theta}(x_j, y_j) = 1$ and it is located on the frontier while a value as $\hat{\theta}(x_j, y_j) < 1$ means that the firm is inefficient and is located under the frontier. Technical efficiency is then calculated for each unit without needing to specify a particular functional form for the production frontier, though the main withdraw of the method is the absence of a random error term in the estimation. In any case, and under some regularity assumptions on the data generating process, DEA provides consistent estimation of all the above concepts (see Kneip, A., L. Simar and P. Wilson (2003) for a review of DEA statistical properties).

Regression analysis

As mentioned in the main text, we do not want to assume any particular functional form on $g(\cdot, \cdot)$ except that it is a smooth function, i.e. has continuous second derivatives.

We explain the estimation of the parameters β and the asymptotic covariance of the estimators, as well as the estimation of the non-parametric function of $g(\cdot, \cdot)$ in a semiparametric model of the form as described in equation (1). We assume $E[e|EF, ESU, R] = 0$, $Var[e] < \infty$. The estimation of $g(\cdot)$ and β will be made in two steps: first the estimation of β 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 modelling see also Härdle, Müller, Sperlich, and Werwatz (2004).

The basic idea is to construct an estimator that gives simply a smooth surface (or hyper-plane), e.g. in the one dimensional case a smooth line, that fits best into the point cloud of real observations. The smoothness of that surface can be (pre-) determined by choosing a respectively large smoothing parameter h, called bandwidth. Actually, this parameter can also often 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 smooth function. Imagine we aim to estimate $G(x_0)$ for some point $x_0 \in \Re^d$. Having observed $\{X_i, Y_i\}_{i=1}^n$, this can be done by local least squares:

$$\begin{pmatrix} \hat{G}(x_0) \\ \widehat{\nabla G}(x_0) \end{pmatrix} = \underset{a_0, a_1}{argmin} \sum_{i=1}^{n} \left\{ Y_i - a_0 - a_1^T (X_i - x_0) \right\}^2 K_h(X_i - x_0) \ , \tag{A8}$$

$a_0 \in \Re$, $a_1 \in \Re^d$ 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 $K(v) = \frac{15}{16}(1 - v^2)^2 1\!\!1\{|v| \le 1\}$. 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 the vector a_1 in equation (A8) 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:

$$\tilde{G}(x_0) = \frac{\sum_{i=1}^n K_h(X_i - x_0) Y_i}{\sum_{i=1}^n K_h(X_i - x_0)} \ . \tag{A9}$$

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 such as ours, the choice of the smoothing parameter should be considered as degrees of freedom which would be chosen, 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.

Coming back to our model (1), we will apply the local linear estimation method, i.e. equation (A8), on $\{W_i := (\ln(EF_i), \ln(ESU_i)), (E_i - \hat{\beta} R_i)\}_{i=1}^m$. The remaining question is how to get $\hat{\beta}$. The estimator of β is defined as

$$\hat{\beta} = S^{-1}_{R\text{-}\tilde{R},R\text{-}\tilde{R}} S_{R\text{-}\tilde{R},E\text{-}\tilde{E}} \tag{A10}$$

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 (A9). It is easy to see that the variance of $\hat{\beta}$ can be estimated by $\hat{\sigma}^2 S^{-1}_{R\text{-}\tilde{R},R\text{-}\tilde{R}}$ 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).

Furthermore, note that as $g(\cdot,\cdot)$ is non-parametric, we could have directly used the covariates *ESU* and *EF* in the model (1). As mentioned above, the logarithm therefore does 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 a 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.

In non- and semi-parametric regression, the choice of smoothness controlled via the bandwidth (named h in Section 4.) and chosen by the empirical researcher, is often either not discussed or quite controversial. 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 β of model (1) 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 for example the regional dummies are almost uncorrelated with the other covariates (ln(*ESU*) and ln(*EF*) in our case). More details on the methodology are available on request.

References

Alvarez, A., C. Arias (2003), Diseconomies of Size with Fixed Managerial Ability. American Journal of Agricultural Econonomics 85: 134-142.

Ball, V.E., J.-C. Bureau, J.-P. Butault, R. Nehring (2001), Levels of Farm Sector Productivity: An International Comparison. Journal of productivity Analysis 15: 5-29.

Ball, V.E., W.A. Lindamood, R. Nehring, C. San Juan (2007), Capital as a Factor of Production in OECD Agriculture: Measurement and Data. Applied Economics, forthcoming.

Ball, V.E., C.A.K. Lovell, H. Luu, R. Nehring (2004), Incorporating Environmental Impacts in the Measurement of Agricultural Productivity Growth. Journal of Agricultural and Resource Economics 29: 436-460.

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, A. Spilimbergo (2002), Who's Afraid of the Big Enlargement? CEPR Policy Paper 7, www.cepr.org/pubs/books/cepr/booklist.asp

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

Blandford, D., J. Dewbre (1994), Structural adjustment and learning to live without subsidies in OECD countries. American Journal of Agricultural Economics 76: 1047-1052.

Bryden, J., E. Hawkins (1992), Economic and Social Impact of Direct Payments in Europe: Some issues arising from the results of 12-country research programme. 30th EAAE Seminar, Chateau d'Oex (Swizerland), mimeo.
Bullock, D.S., K.K. Salhofer (2003), Judging agricultural policies: a survey. Agricultural Economics 28: 225-243.
Bureau, J.-C. (2005), The EU Common Agricultural Policy: Enlargement, reform of and impacts on the western hemisphere countries. forthcoming as an IADB publication. Inter-American Development Bank, Washington DC.
Cooper, W.W., L.M. Seiford, K. Tone (2000), Data Envelopment Analysis. A Comprehensive Text with Models, Applications, References and DEA-Solver Software. Kluwer Academic Publishers. London.
Färe, R., S. Grosskopf, C.A.K. Lovell (1994), Production Frontiers. Cambridge University Press. Cambridge.
Fan, J., I. Gijbels (1996), Local polynomial regression. Chapman and Hall. London.
Farrell, M.J. (1957), The Measurement of Productive Efficiency. Journal of the Royal Statistical Society: Series 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.
Härdle, W., M. Müller, S. Sperlich, A. Werwatz (2004), Nonparametric and Semiparametric Models. Berlin/Heidelberg: Springer Verlag.
Jorgenson, D.W., M. Kuroda (1992), Productivity and International Competitiveness in Japan and the United States. Pp. 203-249 in: B.G. Hickman (ed.), International Productivity and Competitiveness. New York: Oxford University Press.
Jorgenson, D.W., M. Nishimizu (1978), U.S. and Japanese Economic growth, 1952-1974: An International Comparison. Economic Journal 83: 707-726.
Jorgenson, D.W., M. Kuroda, M. Nishimizu (1987), Japan-U.S. Industry-level Productivity Comparisons, 1960-1979. Journal of the Japanese and International Economics 1: 1-30.
Kneip, A., L. Simar, P. Wilson (2003), Asymptotics for DEA estimators in nonparametric frontier models. Discussion Paper 0317, Institut de Statistique, UCL.
Mora, R., 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., C. San Juan, J.E. Torre (2003), The representativeness of the Spanish RICA Survey. Pp. 115-141 in: K. Poppe, A. Povellato, K. Krijgsman (eds.), European farmers and the growing of data. LEI. The Hague.
OECD (2001), Decoupling: A conceptual overview. OECD papers, 10.
Robinson, P. (1988), Root-N-Consistent Semiparametric Regression. Econometrica 56: 931-954.
Ruppert, D., M.P. Wand (1994), Multivariate locally weighted least squares regression. Annals of Statistics 22: 1346-1370.
Kleinhanß, W., C. Murillo, C. San Juan, S. Sperlich (2006), Efficiency, subsidies and environmental adaptation of animal farming under CAP. Agricultural Economics 36: 49-65.
Simar, L., P. Wilson (2000), Statistical Inference in Non-parametric Frontier Models: The State of the Art. Journal of Productivity Analysis 13: 49-78.

Carmen Murillo, Departamento de Economía, Universidad de Cantabria, Avda. Los Castros s/n, 39005 Santander, Spain. Phone: +34/942/202006. E-Mail: carmen.murillo@unican.es

Carlos San Juan, Departamento de Economía, Universidad Carlos III de Madrid, c/ Madrid 126, 28903 Getafe, Spain. Phone: +34/91/6249577. E-Mail: csj@eco.uc3m.es

Stefan Sperlich, Institut für Statistik und Ökonometrie, Georg August Universität Göttingen, Platz der Göttinger Sieben 5, 37073 Göttingen, Germany. Phone: +49/(0)551/39-7350.
E-Mail: stefan.sperlich@wiwi.uni-goettingen.de