Evaluating the Effects of Farm Programs: Results from Propensity Score Matching

Andrea Pufahl*, Christoph R. Weiss**

***Institute of Rural Studies, Federal Agricultural Research Centre**
****Department of Economics, Vienna University of Economics and Business Administration**
Contact: andrea.pufahl@fal.de, cweiss@wu-wien.ac.at.

Paper prepared for presentation at the 107th EAAE Seminar *"Modelling of Agricultural and Rural Development Policies"*. Sevilla, Spain, January 29th -February 1st, 2008

Abstract

The paper applies a non-parametric propensity score matching approach to evaluate the effects of two types of farm programs (agri-environment (AE) programs and the less favoured area (LFA) scheme) on input use and farm output of individual farms in Germany. The analysis reveals a positive and significant treatment effect of the LFA scheme for farm sales and the area under cultivation. Participants in AE schemes are found to significantly increase the area under cultivation (in particular grassland), resulting in a decrease of livestock densities. Furthermore, participation in AE programs significantly reduced the purchase of farm chemicals (fertilizer, pesticide). We also find substantial differences in the treatment effect <u>*between*</u> *individual farms (heterogeneous treatment effects). Farms which can generate the largest benefit from the program are most likely to participate.*

Key words: evaluation, farm programs, propensity score matching

1. Introduction

The impact of government programs on agricultural output and farm structure is a key policy issues in the ongoing international trade negotiations on agriculture. Fostered by the fact that more and more data become available on a micro (individual farm) level, recent years have seen a substantial growth in the number of empirical studies on the consequences of farm policies for individual farms (Salhofer and Streicher, 2005; Shaik and Helmers, 2006), as well as for different regions (Ahearn, Yee and Korb, 2005; Kim et al., 2005). This literature mainly focuses on the consequences of policy measures for farm exit rates, farm output and growth as well as adjustments in on-farm and off-farm labour markets. Despite the fact that these topics now rank high on the agenda of economists and policy makers, Ahearn, Yee, and Korb (2005: 1182) conclude that 'our understanding of how government policies have affected the structure of agriculture, or how future policies could be designed to promote specific outcomes remains limited.'

In practice, policy interventions turn out to be difficult to evaluate. Government programs frequently have different objectives which are difficult to conceptualize, and each program often uses a large set of diverse instruments to accommodate these goals: programs might stimulate demand for agricultural output (export subsidies) or influence production processes (environmental programs or land retirement programs). Further more, policy measures not only impact individual farmers directly but also can trigger indirect effects through a variety of mechanisms. Key and Roberts (2006), for example, argue that farms receiving relatively high direct payments may be able to bid up the price of land and other fixed resources. Given the very complex effects and interactions, economic theory often provides only limited guidance with respect to the 'correct' specification of an econometric model.

Participation in farm programs typically is voluntary. An individual farmer will participate only if the additional benefits exceed the costs of participation. Costs and benefits will differ between individuals depending on specific characteristics of the farm as well as the farm family, some of which, however, may not fully be observed. The existence of systematic differences between program participants and non-participants requires separation of the 'true' effect of program participation ('causal effect') from the effect of initial differences in characteristics of the two groups (‘selection effect’). To distinguish between the two effects, an evaluator has to answer the question: ‘How much did farms participating in the program benefit compared to what they would have experienced without participating in the program?’ The fact that this counterfactual situation can not be observed constitutes the ‘classical evaluation problem’.

Finally, the optimal response to a particular governmental program does not only depend on the characteristics of the individual farm but also on the size and structure of the farm household. We should not expect to find the response to farm programs to be homogenous across individual farms.

The present paper applies a non-parametric propensity score matching approach to evaluate the effects of two types of farm programs (agri-evironmental programs and the less favoured area scheme) for individual farms in Germany. The matching approach is widely used when evaluating labour market policies (see e.g. Heckman, LaLonde and Smith, 1999). According to our knowledge, Lynch et al. (2007) is the only application evaluating agricultural policy measures. The key advantage of matching (over standard regression methods) is that it is less demanding with respect to the modelling assumptions. Specifically, matching does not require functional form assumptions for the outcome equation (it is non-parametric)[1]. Further, with matching, there is no need for the assumption of constant additive treatment effects across individuals. Instead, the individual causal effects are unrestricted and individual effect heterogeneity in the population is permitted. By applying a matching estimator, we thus hope to mitigate some of the difficulties of evaluating the consequences of farm policies mentioned above. Note, however, that the aim of this analysis is not to evaluate the effectiveness of a policy in terms of the degree to which a particular (often vaguely defined) policy objective has been realized. Instead, we follow previous studies and aim at assessing the effects of policy measures with respect to input use (land, labour, farm chemicals) and farm output (sales).

2. The Agri-environmental and Less Favoured Area Program

Total expenditures of the Common Agricultural Policy (CAP) reached 54.6 Billion Euros in 2006. Direct payments and price policies account for the largest share (78 %) of CAP expenditures. A minor but increasing part of the CAP budget (22%) is allocated to rural development policies. Agri-environment programs (AE-programs) and the less favoured area scheme (LFA) account for 57 % of total public expenditures for rural development in the EU (Agrar CEAS Consulting, 2005). These figures illustrate that both schemes became core instruments of the rural development policies within the EU and are no longer of solely marginal importance within the CAP framework.

AE programs and the LFA scheme are directly targeted to farm enterprises. The average proportion of total farmland classified as LFA is 55 % in the EU-15. In Finland, Portugal, Luxembourg, Spain and Greece more than 70 % of the farmed land were classified as LFAs in 2003, while the share of LFAs is marginal in The Netherlands (0 %), Belgium (20 %) and Denmark (0 %) (IEEP, 2006: 153). Support for naturally less-favoured areas in Germany is available on 50 % of farmland. Farms located in

[1] Dehejia and Wahba (1999) and Smith and Todd (2005) directly compare the results of matching and regression estimates and show that avoiding functional form assumptions can be important to reduce bias.

designated LFAs are eligible for support. The core objective of the LFA scheme is the maintenance of the agricultural land use within these regions (Reg. (EC) No 1257/1999, Article 13a). The share of granted farmland on total farmland is highest in the southern part of Germany, followed by western and eastern states (Plankl et al., 2005). LFA support has little relevance in the north of Germany due to superior natural conditions for agricultural production.

The EU's agri-environmental programs were introduced as 'Accompanying Measures' of the 1992 Mac Sharry Reform of the CAP. Farmers receive compensation payments for the adoption of environmentally favourable production technologies. Agri-environment payments are meant to cover the income foregone and additional costs for compliance. Participation in the programs is voluntary and varies significantly between EU member states as well as between different regions within member states. While more than two thirds of the total agricultural area is covered by at least one AE-program in Austria, Finland, Sweden, and Luxemburg, the average share is around 25 % in Germany (Salhofer and Glebe, 2006: 3). Similar to the LFA scheme, participation in AE programs is very high in the South (70 % of total farm land), moderate in the West and East (20 %) and marginal in the North (5 %) of Germany. More than 100 different sub-programs[2] are available within the framework of the AE scheme on the state level. Support for reduced inputs on grassland and arable land and organic farming account for the largest share of AE expenditures in Germany (Osterburg, 2004).

3. Estimation method and data

2.1 Evaluation problem and matching

Evaluation studies attempt to estimate the mean effect of participating in a program (treatment). This requires making an inference about the outcome that would have been observed for the treated ('treatment group') if they had not been treated ('control group'). The key advantage of experimental studies (over non-experimental methods) is the ability to generate a control group that has the same distribution of characteristics as the treatment group. In this case, the treatment effect can be calculated as the difference of mean outcomes. In non-experimental studies on the other hand, subjects usually self-select into treatment groups. Treated and controls differ with respect to their participation status but also with respect to many other characteristics. Calculating the treatment effect as the difference of mean outcomes between the two groups would yield biased results (selection bias).

Matching is a widely-used non-experimental method of evaluation that can be used to estimate the average effect of a particular program (Caliendo, 2006; Heckman, LaLonde and Smith, 1999). This method compares the outcomes of program participants with those of matched non-participants, where matches are chosen on the basis of similarity in observed characteristics. Suppose there are two groups of farmers indexed by participation status $P = 0/1$, where 1 (0) indicates farms that did (not) participate in a program. Denote by Y_i^1 the outcome (performance of farm) conditional on participation ($P = 1$) and by Y_i^0 the outcome conditional on non-participation ($P = 0$).

[2] Unfortunately, our data do not allow us to distinguish among sub programs of the AE programs. Organic farms are identified by a specific code but are too few in numbers for a separate analysis.

The most common evaluation parameter of interest is the mean impact of treatment on the treated, $ATT = E\left(Y_i^1 - Y_i^0 \mid P_i = 1\right) = E\left[Y_i^1 \mid P_i = 1\right] - E\left[Y_i^0 \mid P_i = 1\right]$, which answers the following question: 'How much did farms participating in the program benefit compared to what they would have experienced without participating in the program?' Data on $E(Y_i^1 \mid P = 1)$ are available from the program participants. An evaluator's 'classic problem' is to find $E(Y_i^0 \mid P = 1)$, since data on non-participants enables one to identify $E(Y_i^0 \mid P = 0)$ only. So the difference between $E(Y_i^1 \mid P = 1)$ and $E(Y_i^0 \mid P = 1)$ cannot be observed for the same farm.

The solution advanced by Rubin (1977) is based on the assumption that given a set of observable covariates $\boldsymbol{X}$, potential (non-treatment) outcomes are independent of the participation status (conditional independence assumption-CIA): $Y_i^0 \perp P_i \mid \boldsymbol{X}$. Hence, after adjusting for observable differences, the mean of the potential outcome is the same for P = 1 and P = 0 ($E(Y_i^0 \mid P = 1, \boldsymbol{X}) = E(Y_i^0 \mid P = 0, \boldsymbol{X})$). This permits the use of matched non-participating farms to measure how the group of participating farms would have performed, had they not participated.

This procedure assumes that after conditioning on a set of observable characteristics, outcomes are conditionally mean independent of program participation. Heckman, Ichimura and Todd (1997) stress that, for a variety of reasons, there may be systematic differences between participant and non-participant outcomes, even after conditioning on observables. Such differences may occur, for example, because of program selectivity on unmeasured characteristics or because of level differences in outcomes ($(E(Y_i^1 - Y_i^0 \mid P_i = 1))$ that might arise when participants and non-participants reside in different regions. To improve the results of the matching procedure, the authors suggest a conditional difference-in-difference matching estimator (d-i-d). Let *t* represent a time period after the program start date and *t'* a time period before the program. The conditional d-i-d estimator compares the conditional before-after outcomes of program participants with those of non-participants: $E\left(Y_{it}^1 - Y_{it'}^0 \mid P_i = 1, \boldsymbol{X}\right) - E\left(Y_{it}^0 - Y_{it'}^0 \mid P_i = 0, \boldsymbol{X}\right)$. The d-i-d is attractive because, unlike conventional matching estimators, it permits selection to be based on potential program outcomes at time *t'* and allows for selection on unobservables (Smith and Todd, 2005).

Instead of conditioning on $\boldsymbol{X}$, Rosenbaum and Rubin (1983) suggest conditioning on a propensity score ('propensity score matching'). The propensity score is defined as the probability of participation for farm *i* given a set $\boldsymbol{X} = \boldsymbol{x}_i$ of farm characteristics $p(\boldsymbol{X}) \equiv \Pr\left(P_i = 1 \mid \boldsymbol{X} = \boldsymbol{x}_i\right)$. In the present context with multiple treatments (AE programs and LFA scheme), the propensity scores are derived from two logit models where participation in the AE and LFA program serve as endogenous variables. The estimated propensity scores are then used to construct the comparison groups. A Greedy algorithm employing calliper pair matching without replacement is applied (Parson, 2001). To assess the stability of our results we compare the outcomes with those from alternative estimators (Heckman, Ichimura and Todd, 1997; Smith and Todd, 2005).

3.2 Data and definition of variables

The empirical analysis is based on a panel data set ('LAND-Data') of more than 32,000 bookkeeping farms in Germany for the period 2000 to 2005[3]. 'LAND-Data' provides information on farm characteristics (area under cultivation, sales, labour inputs, capital endowment and expenditures for farm chemicals ...) and also includes information on the participation in the AE and LFA program. From these 32,000 farms in the original data set, roughly one third had to be eliminated due to incomplete and missing data. To evaluate the effect of programme participation with the conditional d-i-d estimator, we focus only on those farms, which did not participate in the program in the initial time period (2000). The selection of data and the definition of the participation variables are described in Table 1.

Table 1
Sample Selection Criteria and Program Participation

	AE programs	*LFA scheme*
Total number of farms with continuous records from 2000 to 2005	*32,503*	
Omitted due to missing observations for some variables	*10,390*	*8,594*
Number of remaining farms	*22,113*	*23,909*
Program Participation in base year (2000)	*557*	9,695
Non-participation in base year (2000)	*21,556*	14,214
Continuous program participation (2001 – 2005):		
Dummy variable (P_{AE}) is set equal to 1 for farms which continually participate in the AE program from 2001 until 2005 (for five years)	***9,138***	
Dummy variable (P_{LFA}) is set equal to 1 for farms which continually participate in the LFA scheme starting from 2001, 2002 or 2003 until 2005 (at least for three years)		***502***
Program participation in some years only (these observations will not be used for the empirical analysis):		
Number of farms participating in the AE program for some years only (less than five years)	*5,223*	
Number of farms participating in the LFA scheme for few years only (less than three years)		*637*
Non-participation (2001 – 2005):		
Dummy variable (P_{AE}) is set equal to 0 for farms which never participated in the AE program between 2001 and 2005	***7,195***	
Dummy variable (P_{LFA}) is set equal to 0 for farms which never participated in the LFA scheme between 2001 and 2005		***13,075***

[3] The sample comprises 8 % of all farm enterprises in Germany. Note that the sample is not representative for Germany as large-scale and full-time farm enterprises are over represented.

The basis for the empirical analysis (propensity score difference-in-difference matching estimator) of AE programs are those 21,556 farms that did not participate in AE programs in the base year 2000. From those farms, 9,138 farms (42.4 %) continually participate in AE programs during the following five year period from 2001 until 2005 (the dummy variable P_{AE} is set equal to one). The dummy variable is set equal to zero for the 7,195 farms (33.4 %) which never participate in AE programs between 2001 and 2005. Note that 5,223 farms (24.2 %) participate in some years only. These farms will not be used for the empirical analysis.

A dummy variable for participation in the LFA program is defined in a similar way. In the initial period 2000, 14,214 farms did not participate in the LFA program. Note that in the case of the LFA program, the number of farms continually participating in the program in all five years (from 2001 until 2005) and not participating in the base year 2000 is very small (only 109 farms). Since this number is too small to carry out a matching analysis, we have chosen a less restrictive classification criterion in this case. A dummy variable (P_{LFA}) is set equal to one for those farms (502 or 3.5 %) which participate in the program from 2001, 2002 or 2003 until 2005 (for at least for three years). The majority of farms (13,075 or 92.0 %) never participate in the LFA scheme. The dummy variable P_{LFA} is set equal to zero in this case. The remaining 637 farms, which participate in a few years only, are eliminated from the empirical analysis. Given that the selection into the treatment group is less restrictive for the LFA scheme, we expect to find a weaker causal effect of this program. Whether program participation (P_{AE} = 1 or P_{LFA} = 1) has significant effects on farm performance rates will be evaluated in the following section.

4. Empirical results

4.1 Propensity Scores and Matching

Conditional probabilities for participation in AE and LFA programs are computed by estimating two logit models. The estimated models are statistically significant at the 1 % level or better, as measured by the likelihood ratio test. The empirical model for the AE-program (the LFA scheme) correctly classifies 87.79 % (96.61 %) of all observations. From the parameter estimates of the logit models, the unbounded propensity scores are calculated for every farm which is then used for the matching analysis[4].

Matching is considered successful when significant differences of covariates among participants and non-participants are removed. Table 2 reports unadjusted and adjusted mean differences of covariates among participants and non-participants of AE and LFA programs in the pre-treatment year (2000).

[4] We use the unbounded $x_i^{'}\hat{\beta}$ rather than the bounded propensity score $\Phi\left(x_i^{'}\hat{\beta}\right)$ because of its preferable distribution properties (Hujer, Mauerer and Wellner, 1997).

Table 2
Mean comparison of selected variables (Frequencies for Dummies) in the pre-treatment year 2000

	Agri-Environmental Programs			Less Favoured Area Program		
	(1)	(2)	(3)	(4)	(5)	(6)
Variable	Selected Treatments	Potential Controls	Selected Controls	Selected Treatments	Potential Controls	Selected Controls
Ln farm sales	4.781	**4.783**	4.776	4.742	**4.807**	4.783
Ln on-farm labour	0.365	**0.307**	0.363	0.427	**0.326**	0.464
Ln off-farm labour	1.080	1.118	1,083	0.723	**1.159**	0.740
Ln area under cultivation	4.053	**3.932**	4.039	3.955	3.941	3.971
Ln share of grassland	3.066	**2.920**	3.047	2.850	**2.304**	2.783
Ln share of rented land	3.792	**3.698**	3.798	3.745	3.712	3.749
Ln farm sales (per ha)	0.728	**0.850**	0.737	0.787	**0.866**	0.812
Farm income	20.159	**18.466**	18.554	21.140	18.836	19.950
Ln farm capital (per ha)	2.301	**2.349**	2.319	2.521	2.414	2.493
Ln cattle livestock units	3.247	**3.135**	3.227	2.995	**2.553**	2.926
Ln cattle livestock density	0.489	**0.545**	0.494	0.511	**0.474**	0.501
Ln fertilizer expenditures (per ha)	-2.522	**-2.443**	-2.532	-2.925	**-2.409**	-2.641
Ln pesticide expenditures (per ha)	-2.970	**-2.952**	-2.992	-2.640	**-2.617**	-2.871
Dummy North	593	**2,970**	541	6	**4,865**	5
Dummy West	711	**3,545**	751	108	**4,313**	108
Dummy South	451	**581**	463	334	**3,718**	337
Dummy East	52	**95**	52	4	179	2
Number of observations	1,807	7,195	1,807	452	13,075	452

Notes: For variable definition and abbreviations see Table A1. Bold numbers indicate significantly different means between observation from the potential (selected) treatment group and from the potential (selected) control group in a t-test for equality of means at the 5 % level.

Prior to the matching analysis, farms participating in AE and/or LFA programs significantly differ from non-participants with respect to nearly all characteristics shown in Table 2. A comparison between column (1) and (2) indicates that farms enrolled in AE programs are characterized by a larger area under cultivation and higher farm incomes, for example. These differences in farm characteristics between program participants and non-participants are significantly different from zero. Table 2 also reports significant differences between treatments and potential controls in the case of participation in the LFA program (compare column (4) and (5)).

Columns (3) and (6) report the means of the relevant variables for the control group after the matching procedure has been applied. From the 9,138 (502) farms with participation in AE (LFA) programs, 1,807 (452) were matched to farms with no participation but similar propensity scores. The differences to columns (1) and (4) are now much smaller and in no case significantly different from zero at the 5 % level. We can thus conclude that all differences in means between treatments and controls have been removed through matching in the initial period 2000 (before program participation).

4.2 Treatment Effects

The average effect of the participation in AE and LFA programs is estimated by comparing the changes in individual outcomes (farm characteristics) between participants ($\Delta Y_i^1 = Y_{i,2005}^1 - Y_{i,2000}^1$) and their matched counterparts ($\Delta Y_i^0 = Y_{i,2005}^0 - Y_{i,2000}^0$) between 2000 and 2005 (d-i-d analysis). The impact of treatment on the treated ('causal effect' of program participation) is estimated by computing mean differences across both groups:

$$ATT = \frac{1}{N_1}\left(\sum_{i=1}^{N_1} \Delta Y_i^1 - \sum_{i=1}^{N_1} \Delta Y_i^0\right).$$

A positive (negative) value of ATT suggests that farms with participation in AE and/or LFA programs have higher (lower) growth rates of variable *Y* than non-participants. Table 3 displays mean growth rates for the treatment and control group as well as the difference between both (the ATT).

Table 3
Average treatment effect (ATT) of the treated for the AE and LFA programs (2000 to 2005)

	Treatments [1]	Controls [2]	ATT = [1] - [2]	t-value (Significance)	
Agri-Environmental Programs					
Ln farm sales	0.073	0.048	0.025	1.72	(*)
Ln on-farm labour	0.007	-0.012	0.019	1.95	(*)
Ln off-farm labour	-0.003	-0.001	-0.002	-0.55	
Ln area under cultivation	0.077	0.042	0.035	5.32	(***)
Ln share of grassland	-0.046	-0.098	0.052	3.13	(***)
Ln share of rented land	0.004	-0.018	0.022	1.52	
Ln cattle livestock units	-0.187	-0.187	0.001	0.03	
Ln cattle livestock density	-0.108	-0.048	-0.060	-3.30	(***)
Ln farm sales (per ha)	-0.004	0.006	-0.010	-0.70	
Ln farm capital (per ha)	0.035	0.047	-0.012	-0.79	
Ln fertilizer expenditures (per ha)	0.037	0.131	-0.094	-4.57	(***)
Ln pesticide expenditures (per ha)	-0.025	0.022	-0.047	-1.97	(**)
Less Favoured Area Program					
Ln farm sales	0.144	0.056	0.088	2.96	(***)
Ln on-farm labour	0.001	-0.012	0.013	0.83	
Ln off-farm labour	-0.008	0.000	-0.008	-1.11	
Ln area under cultivation	0.114	0.060	0.054	3.64	(**)
Ln share of grassland	-0.041	-0.048	0.007	0.28	
Ln share of rented land	0.043	0.011	0.032	1.06	
Ln cattle livestock units	-0.102	-0.147	0.046	1.09	
Ln cattle livestock density	-0.095	-0.088	-0.007	-0.26	
Ln farm sales (per ha)	0.030	-0.004	0.034	1.17	
Ln farm capital (per ha)	0.000	-0.011	0.011	0.34	
Ln fertilizer expenditures (per ha)	0.132	0.169	-0.037	-0.91	
Ln pesticide expenditures (per ha)	-0.012	0.037	-0.049	-0.92	

Notes: For variable definition and abbreviations see Table A1. Asterisks denote statistical significance in a t-test for equality of means at 1 % (***), 5 % (**), or 10 % (*) level.

The d-i-d estimator suggests a significant and positive causal impact of program participation on farm sales. During the period of investigation (from 2000 until 2005) sales of farms participating in AE programs have been growing by 7.3 %, while non-participants report a positive growth rate in sales of 4.8 % on average. The difference (ATT = 2.5 %) is different from zero at the 10 % level of significance. The slight positive effect of AE programs on farm sales is surprising insofar as participation in this program requires the adoption of less intensive production methods which could be expected to reduce farm output and thus farm sales, ceteris paribus. No such adjustment in production methods is required for participation in the LFA program. Consistently, we observe that the causal effect of program participation in the LFA program is much stronger. Sales of non-participants have been growing by 5.6 % on average during the period from 2000 until 2005, while the growth rate of participating farms is 14.4 % on average. The average treatment effect on the treated (ATT = 8.8 %) is significantly different from zero at the 1 % level. Where does this significant increase in farm sales come from?

Results in Table 3 suggest that the increase in farm sales is paralleled by an increase in the area under cultivation. These results comply with findings of Key et al. (2005: 1217f) and Osterburg (2004). Average growth rates of the area under cultivation differ significantly among program participants and non-participants. On average, participation in AE programs causes farm growth rates to double. Whereas the area under cultivation for non-participants has been growing by 4.2 %, participants report a growth rate of 7.7 % on average. The average treatment effect on the treated (ATT) of 3.5 % is significantly different from zero at the 1 % level.

Higher farm land growth rates of participants in AE programs can be explained by the adjustment process of farms induced by program eligibility criteria. Farms with participation to certain AE programs (low input grassland management, for example) are, among others, required not to exceed a certain cattle livestock density (livestock units per forage area). In order to meet these criteria, farm operators predominantly choose to expand the forage area, while total cattle livestock units per farm are kept stable. Results in Table 3 illustrate this adjustment process. The number of livestock units is not affected by programme participation (ATT = 0 %). The cattle livestock density is, on average, reduced by 10.8 % in farms with program participation compared to a decrease of 4.8 % in farms with non-participation. The ATT with respect to the cattle livestock density is –6 % and significantly different from zero at the 1 % level.

The causal effect of the LFA scheme on farm land growth is of similar magnitude as for the AE programs. Scheme participation increases growth in the area under cultivation from 6.0 % to 11.4 % (ATT = 5.4 %). No significant causal effect is observed with respect to the amount of cattle livestock units or density. For the LFA scheme, the changes in farm land are very similar in magnitude to the figures reported for farm sales. Given the fact that LFA payments are granted on a per-acreage base, the increase of farmed land eligible for LFA payments seems to be a reasonable strategy to maximize benefits from participation.

Table 3 does not suggest a significant treatment effect of AE programs on productivity (sales per hectare), the causal effect of the LFA program is positive but not significantly different from zero. This result corresponds to Salhofer and Streicher (2005) who also observe an insignificant productivity effect of participation for ten different farm programs in Austria. The same holds for the capital endowment on farms with program participation, which does not change significantly compared to the control group.

An important objective of agri-environmental policy in Germany is the maintenance of grassland. Land eligible for AE support is mainly grassland, whereas both, arable land and grassland are eligible for LFA support. Neither AE nor LFA support resulted in an increase of the share of grassland in farms with program participation. We find that the share of grassland decreases significantly less in farms participating in AE programs (-4.6 %) than in those with non-participation (-9.8 %). The ATT of 5.2 % is significantly different from zero. The effect of the LFA program on the share of grassland is almost zero. We conclude that current AE programs slow down the decrease of grassland while they are not able to stop or reverse this process.

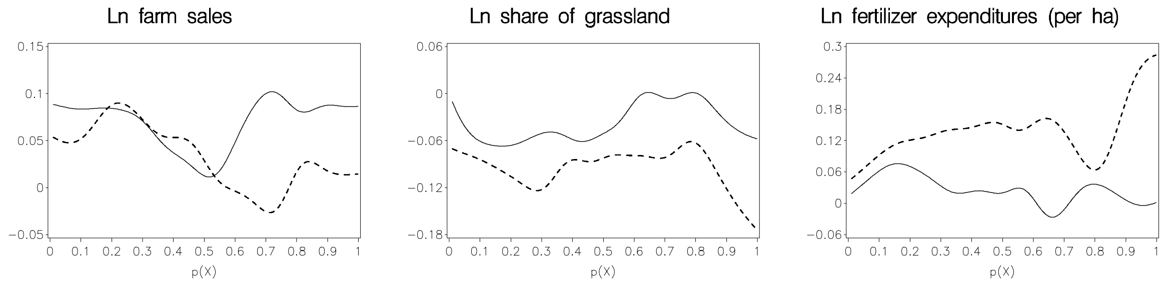
Participants in the AE program are required to reduce or abandon the use fertilizers and pesticides, while no such eligibility criteria are in place for the LFA scheme. Consequently, the causal effect of program participation with respect to expenditures for fertilizer and pesticides per hectare differs remarkably between the two programs. The amount of expenditures for fertilizers (pesticides) per hectare changes in farms with participation in AE programs by 3.7 % (-2.5 %). At the same time, expenditures for fertilizers (pesticides) in the control group increase by 13.1 % (2.2 %). The ATT of -9.4 % (fertilizers) and -4.7 % (pesticides) indicate that farms participating in AE significantly reduced the purchase of farm chemicals compared to the control group. The ATT is significantly different from zero at the one percent (fertilizers) and five percent (pesticides) levels. No significant treatment effect is observed for the LFA program with respect to expenditures for farm chemicals.

4.3 Heterogeneity of Effects and Robustness of Results

The response to a specific policy differs between individual observations (heterogeneity of treatment effects) for various reasons. First, it is plausible to expect that the treatment effect increases with the probability of participation in the program; that is, farmers who can generate the largest benefit from the program are most likely to participate. Second, the magnitude of the treatment effect might be influenced by the amount of program payments per hectare, by farm size and the duration of program participation.

To check these hypotheses, we follow and extend the approach suggested in Lechner (2002). The expectation of the outcome variable conditional on the conditional selection probability ($p(\boldsymbol{X})$) in the pool of participants and non-participants is shown in Figure 1. The comparison is based on kernel-smoothed regressions for program participants in AE programs (solid line) versus non-participants (dotted line) for selected variables. Figure 1 clearly supports the idea of heterogeneous treatment effects. The causal effect of the farm program, which is the difference between the two curves at any point, fluctuates over the support of participation probabilities. The outcomes for the program participants (solid line) are higher for farm sales and the share of grassland; and lower for fertilizer expenditures at (almost) all points, which is consistent with the significant average treatment effect of AE programs for these variables reported in Table 3.

Figure 1
Nonparametric regression of the conditional participation probabilities ($p(\boldsymbol{X})$)
on the outcome variable for the AEP



Remarks: Nadaraya-Watson estimate using a Gaussian kernel and the rule-of-thumb bandwidth. The solid (dotted) line represents the outcome variable for participants (non-participants) in AE programs.

For the LFA program, we found a positive relationship between the probability of program participation ($p(\boldsymbol{X})$) and the treatment effect on farm sales and the area under cultivation. Our results further support the hypothesis, that for both programs, the magnitude of the treatment effect is influenced by the duration of program participation. No clear relationship is found between the causal effect and the amount of program payments and farm size (as measured by the area under cultivation)[5].

The robustness of our results is tested by applying alternative matching estimators. From these results we can conclude, that our results are relatively independent of the choice of matching estimator. Finally, we compared the results from propensity-score matching with those from a naive estimator (without controlling for differences in pre-treatment characteristics). The calculated effect of AE programs on farm sales (the area under cultivation) is overestimated by the naive model by a factor of 2.2 (1.2). Distinguishing between a 'selection effect' and a 'causal effect' is key for an appropriate evaluation of farm programs.

5. Conclusions

Evaluating the effects of farm programs on farm output is a key policy issues since this determines whether programs are condemned as trade distorting or can be classified as 'decoupled' and conform with WTO regulations. An empirical evaluation of the effects of farm programs, however, faces a number of challenges: First, economic theory often provides limited guidance with respect to the appropriate specification of an econometric model. Second, farms self-select into program participation; participants and non-participants thus differ significantly in important characteristics (selection bias). Third, factors that determine the selection into the program and/or influence outcome variables may not fully be observed (unobserved heterogeneity). Further it remains unknown how participants would have performed if they had not participated in the program, as counterfactuals cannot be observed in non-experimental studies. Finally, the optimal response to governmental programs will not be homogenous across individual farms (heterogeneity in response).

[5] Results are available from the authors on request.

The present paper addresses these issues by applying a non-parametric propensity score matching approach (difference-in-difference estimator). The method turns out to be a useful technique for the empirical evaluation of farm programs. Specifically, we investigate the effects of two farm programs – agri-environment (AE) programs and the less favoured area (LFA) scheme – with respect to input use and farm output in Germany for the period 2000 to 2005.

The analysis reveals a positive and significant treatment effect of the LFA programs on farm sales. Growth rates in farm sales are significantly higher (on average) for farms participating in the LFA program compared to non-participants. The increase in farm sales observed is paralleled by an increase in the area under cultivation. Since LFA payments are granted on a per-acreage basis, an increase in land eligible for support seems to be a reasonable strategy to maximize benefits from participation. We also observe a significant positive effect of the AE program on the area under cultivation. The increase in farm size can be explained by the need to reduce livestock densities (livestock units per forage area) in order to become eligible for AE payments. Compared to non-participation, AE participants reduce expenditures for farm chemicals (fertilizer, chemicals). The share of grassland per farm continues to decrease in farms with participation in AE programs, although at a lower rate than in farms with non-participation. No significant effect on farm productivity (sales per ha), capital endowment per ha, off-farm labour, total cattle livestock units (per farm) can be found for AE or LFA programs.

The analysis reveals substantial differences in the treatment effects between individual farms (heterogeneous treatment effects). Farmers who can generate the largest benefit from the program (in terms of additional sales, for example) are most likely to participate. Other sources of heterogeneous treatment effects are found to be the duration of program participation and. Heterogeneous treatment effects have not yet been addressed in greater detail in empirical evaluation studies. This remains an important area to be explored in future research.

Literature

Council Regulation (EC) No 1257/1999 of 17 May 1999 on support for rural development from the European Agricultural Guidance and Guarantee Fund (EAGGF) and amending and repealing certain Regulations.

Agrar CEAS Consulting (2005): Synthesis of rural development mid-term evaluations. Lot 1, EAGGF Guarantee. European Commission: http://ec.europa.eu/agriculture/eval/index2_en.htm#rep2.

Ahearn, M. C.; Yee, J. and Korb, P. (2005): Effects of differing farm policies on farm structure and dynamics. *American Journal of Agricultural Economics* 87: 1182-1189.

BBR, Bundesamt für Bauwesen und Raumordnung (2004): Indikatoren und Karten zur Raumentwicklung, INKAR Ausgabe 2004. Bonn.

Caliendo, M. (2006): Microeconometric Evaluation of Labour Market Policies. Lecture Notes in Economics and Mathematical Systems. Springer.

Dehejia, R. and Wahba, S. (1999): Causal effects in nonexperimental studies: reevaluation the evaluation of training programs. *Journal of the American Statistical Association* 94: 1053-1062.

Heckman, J.; Ichimura, H. and Todd, P. E. (1997): Matching as an econometric evaluation estimator: evidence from evaluating a job training program. *Review of Economic Studies* 64: 605-654.

Heckman, J.; LaLonde, R. J. and Smith, J. (1999): The economics and econometrics of active labor market programs. In: Ashenfelter, O. and Card, D. E. (eds.): *Handbook of Labor Economics. Vol. III*. Amsterdam, 1865-2097.

Hujer, R.; Mauerer, K. O. and Wellner, M. (1997): The impact of training on unemployment duration in West Germany. Discussion Paper 74. Department of Economics, Johann Wolfgang Goethe-University Frankfurt.

IEEP, Insitute for European Environmental Policy (2006): An evaluation of the Less Favoured Area measure in the 25 member states of the European Union. European Commission, DG Agri: http://ec.europa.eu/agriculture/eval/reports/lfa.

Key, N.; Lubowski, R. N. and Roberts, M. J. (2005): Farm-level production effects from participation in government commodity programs: Did the 1996 federal agricultural improvement and reform act make a difference? *American Journal of Agricultural Economics* 87: 1211-1219.

Key, N. and Roberts, D. (2006): Government payments and farm business survival. *American Journal of Agricultural Economics* 88: 382-392.

Kim, C. S.; Schluter, G.; Schaible, G.; Mishra, A. K. and Hallahan, C. (2005): A decomposed negative binominal model of structural change: A theoretical and empirical application to U.S. agriculture. *Canadian journal of agricultural economics* 53: 161-176.

Lynch, L.; Gray, W. and Geoghegan, J. (2007): Are farmland preservation program easement restrictions capitalized into farmland prices? What can a propensity score matching analysis tell us? *Review of agricultural economics* 29: 502-509.

Osterburg, B. (2004): Assessing long-term impacts of agri-environmental measures in Germany. OECD Workshop on evaluation agri-environmental policies. Session 4. Evaluation of payments. Paris, 6-8 December 2004.

Parson, L. S. (2001): Reducing bias in a propensity score matched-pair sample using greedy matching techniques. Proceedings of the 26th Meeting of the SAS International User Group (SUGI 26): www2.sas.com/proceedings/sugi26/p214-26.pdf.

Plankl, R.; Brand-Saßen, H.; Daub, R.; Doll, H.; Pohl, C. and Rudow, K. (2005): Aktualisierung der Halbzeitbewertung der Ausgleichszulage in benachteiligten Gebieten 2001 bis 2004. Länderübergreifender Bericht. Braunschweig (FAL; Institut für Betriebswirtschaft, Agrarstruktur und ländliche Räume).

Rosenbaum, P. R. and Rubin, D. B. (1983): The central role of propensity score in observational studies for causal effects. *Biometrica* 70: 41-55.

Salhofer, K. and Glebe, T. W. (2006): EU agri-environment programs and the "restaurant table effect". Discussion Paper 04-2006, Technical University of Munich, Environmental Economics, Resource Economics and Agricultural Policy.

Salhofer, K. and Streicher, G. (2005): Self-selection as a problem in evaluation agri-environmental programs. Paper presented at the 87th EAAE-Seminar: Assessing rural development of the CAP.

Shaik, S. and Helmers, G. (2006): An examination of farm program payments on farm economic structure. Research in Agricultural and Applied Economics: http://agecon.lib.umn.edu/.

Smith, J. A. and Todd, P. E. (2005): Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics* 125: 305-353.

STAT, Statistical Office of Germany and of German States (2005): Statistic regional. Data for districts and towns in Germany. Düsseldorf.

ZAV, Federal Employment Agency (2005): Unemployed persons on district level. Federal Employment Agency: http://www.pub.arbeitsamt.de/hst/services/statistik/000000/html/start/index.shtml.

Appendix

Table A1:
Variable definition and data sources

Variables	Unit	Year	Source
Participation in AE programs	0=no, 1=yes	2000-2005	LAND-Data
Participation in the LFA scheme	0=no, 1=yes	2000-2005	LAND-Data
Farm characteristics			
Area under cultivation	ha	2000-2005	LAND-Data
Farm sales	1000 Euro	2000-2005	LAND-Data
Share of grassland	%	2000-2005	LAND-Data
Share of rented land	%	2000-2005	LAND-Data
On-farm labour units	LU	2000-2005	
(1 LU = 2720 working hours per year)			LAND-Data
Off-farm labour (farmer couple)	LU	2000-2005	LAND-Data
Farm capital (per ha)	1000 Euro	2000-2005	LAND-Data
Fertilizer expenditures (per ha)	1000 Euro	2000-2005	LAND-Data
Pesticide expenditures (per ha)	1000 Euro	2000-2005	LAND-Data
Commodity payments, livestock (per ha)	1000 Euro	2000-2005	LAND-Data
Commodity payments, arable (per ha)	1000 Euro	2000-2005	LAND-Data
AE payments (per ha)	Euro	2000-2005	LAND-Data
LFA payments (per ha)	Euro	2000-2005	LAND-Data
Livestock units (all livestock)	LSU	2000-2005	LAND-Data
(1 LSU = 1 milk cow)			
Cattle livestock units	LSU	2000-2005	LAND-Data
Cattle livestock density (LSU per ha grassland, fodder crops)	LSU		
Livestock farm	0=no, 1=yes	2000-2005	LAND-Data
Pig & poultry farm	0=no, 1=yes	2000-2005	LAND-Data
Soil index (< 30 = very poor, 100 = best)	Index	2000-2005	LAND-Data
Regional characteristics			
Unemployment rate	%	2000-2005	ZAV (2005)
Farmland rent (per ha)	Euro	2000-2005	LAND-Data
Land price (per sqm)	Euro	Ø 2000-2002	BBR (2004)
Share of rural population	%	2001	BBR (2004)
Share of farms ≤ 20 ha	%	1999	STAT (2005)
Share of farms ≥ 20 < 50 ha	%	1999	STAT (2005)
Change in the number of farms ≥ 20 < 50 ha	%	1999, 2003	STAT (2005)
(between 1999 and 2003)			
Gross value added in agriculture	1000 Euro	2000-2004	STAT (2005)
Share of gross value added in agriculture	1000 Euro	2000-2004	STAT (2005)
Gross domestic product (per capita)	1000 Euro	2000-2004	STAT (2005)
Share of livestock farms	%	1999	STAT (2005)
Share of arable farms	%	1999	STAT (2005)
Share of pig & poultry farms	%	1999	STAT (2005)
Share of mixed farms	%	1999	STAT (2005)
Dummy North	0=no, 1=yes	2000-2005	LAND-Data
Dummy West	0=no, 1=yes	2000-2005	LAND-Data
Dummy South	0=no, 1=yes	2000-2005	LAND-Data

Abbreviations: AE = Agri-environmental programs, LFA = Less favoured areas program, ha = hectare, sqm = square meter, LU = Labour units, LSU = Livestock units
Notes: 'Regional characteristics' refer to the characteristics of the 440 administrative districts of Germany.