**85th Annual Conference of the Agricultural Economics Society**
**Warwick University**

**18 – 20 April 2011**

# Using a *matching* analysis to evaluate the structural effects of farm-investment support in Austria

**Stefan Kirchweger, Michael Eder, Martin Kapfer and Jochen Kantelhardt**[1]

## Abstract

The increasing demand for agricultural policy evaluation and the complexity of identifying the effects of policy measures are two challenges to agricultural economics. In this context the following paper shows the opportunities and limitations of the *direct radius-matching method* in evaluating EU rural development policies, as exemplified by the agricultural investment support program for Upper Austria. To determine the causal effects of investment support on the structure of participating farms, the matching approach is combined with the *Difference-in-Difference estimator.* The results show positive, but heterogeneous average effects on utilized agricultural area (UAA) and total livestock units (LU). As this approach is based on crucial assumptions, sensitivity analyses with modified calliper width for radius-matching are used. They show that wide callipers are preferable to narrow ones, when heterogeneous effects appear. The possible conclusion is that this approach is superior to the common method, such as before-/after-Comparison and regressions analysis, for evaluating the EU rural development when appropriate data and theoretical basis for the assumptions are available.

**Key words:** farm-investment support, causal effects, *direct radius-matching*, *difference-in-difference estimator*

**JEL classifications**: Q10, Q12, Q18

---

[1] Stefan Kirchweger, Michael Eder, Martin Kapfer and Jochen Kantelhardt are with the Institute of Agricultural and Forestry Economics at the University of Natural Resources and Life Sciences, Vienna, Feistmantelstrasse 4, A-1180 Vienna, Austria. E-Mail: stefan.kirchweger@boku.ac.at for correspondence.

## 1 Introduction

In the last decades several Common Agricultural Policy reforms in the European Union (EU) have increased the funds of rural-development programmes (2.pillar). Those payments accounted in 2008 for already 20% of EU agricultural expenditure. In Austria the co-financed rural-development programme is of much greater importance in its farm policy than in other countries. More than 2.3 million Euros – or 49% of the Austrian agriculture budget – were spent on this programme in 2009 (BMLFUW, 2010). The funded measures are categorised on four axes, where the agro-environmental programme (ÖPUL) and the compensation payments for disadvantaged areas (CP) which are part of Axis 2 (Environment and the countryside) are the most important ones. Through a change from the last funding period to the current period, money has been shifted from Axis 2 to Axis 1 (Competitiveness) and increased the funds in its main-measure farm-investment support. For this measure, 311 million Euros were spent in the last period and already 265 million Euros from 2007 to 2009 (Dantler et al., 2010). In 2009 the farm-investment support reached 128 million Euros, which is about 11% of the Austrian rural-development programme (see Figure 1).

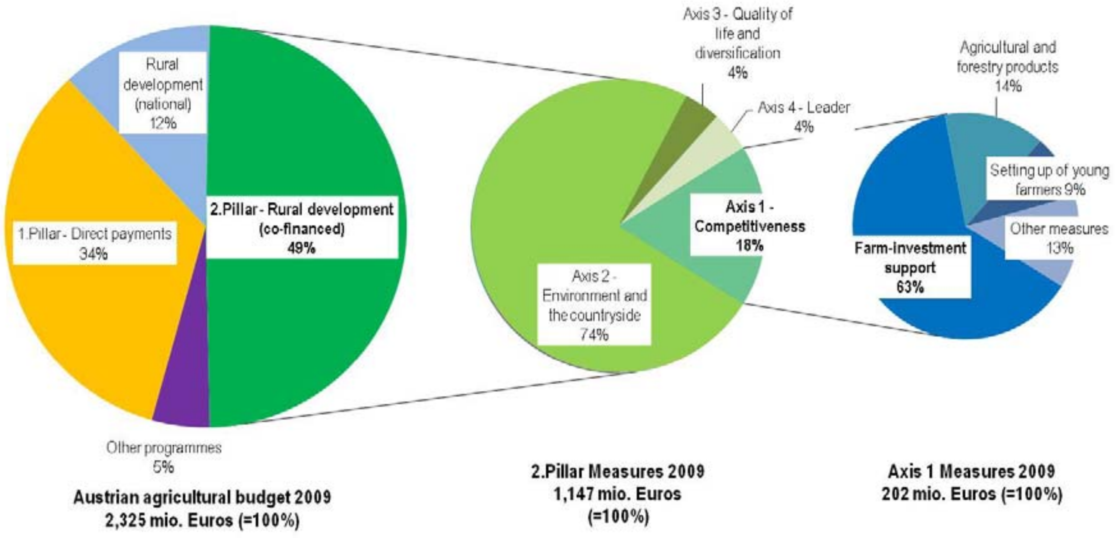

Figure 1: Austrian farm expenditures 2009 (Source: BMLFUW, 2010)

The fact of increasing expenditures and complex outcomes in rural-development measures requires consistent evaluation, which can be achieved by using quantitative ex-post analysis. Furthermore, evaluation is increasingly necessary to verify those expenditures to other societies. These challenges have been recognised by the EU and guidelines for a consistent evaluation have been set up but have not been carried out by practical evaluators (Henning and Michalek, 2008).

For policy evaluation it is necessary to identify effects which are directly caused by the programme. But the causal effects of farm-investment support are hard to define, as those payments are always in combination with an investment. A control group of farms who have invested but are not supported by the programme can hardly be found (Dirksmeyer et al., 2006). This study therefore considers the causal effects of support and investment jointly.

The Evaluation of farm-investment support is often based on *before/after analysis* where the outcome of participants in a before and after situation is compared (Beck und Dogot, 2006; Dirksmeyer et al., 2006; Pfefferli, 2006; Striewe et al., 1996). This approach has its main drawback in evaluating gross effects instead of net effects. Net effects can be assessed by using empirical *with/without treatment analysis.* When empirical data for evaluation of the effects of farm programmes is used, further problems arise (Pufahl and Weiss, 2009). One of the major difficulties is the identification of an adequate control group, which is required for measuring causal effects. Rural-development measures in particular show systematic differences in participants and non-participants (selection bias), caused by voluntary selection to programme participation. Salhofer und Streicher (2004) illustrate the evidence of selection bias in the agro-environmental measure ÖPUL. Consequently Forstner et al. (2008) contain the groups by economic and structural variables and compare the before and after situation of both groups. Further reduction of selection bias can be achieved by using *matching*, which compares participants pair wise with similar non-participants. Whereas the *matching* method is commonly applied in medicine and macroeconomics (Gensler et al., 2005), there have been a number of recent papers introducing *matching* as an innovative non-parametric method for the evaluation of agricultural policies (Henning and Michalek, 2008; Pufahl, 2009; Pufahl and Weiss, 2009).

The objective of this study is to apply a direct *radius-matching approach* to identify controls for participants in rural-development programme evaluation. The *matching* approach is combined with a d*ifference-in-difference estimator (DiD)* for measuring the *Average Treatment Effect on the Treated (ATT)* in the Austrian farm-investment programme. We use IACS-data (Integrated Administration and Control System) and focus on the structural effects on participants in the Upper-Austria region. Furthermore, the heterogeneity and robustness of results are investigated to show the opportunities and limitations of this approach in overcoming the challenges mentioned.

In Section 2 which follows, there is a brief description of the farm-investment support programme in Austria. Based on the work of Dantler et al. (2010), we illustrate the distribution of payments and participants as well as the structural characteristics of participants. Section 3 explains the evaluation problem in detail and the method and database used. The empirical results of the *matching* approach and its heterogeneity and robustness analysis are displayed in Section 4. In Section 5 a conclusion is drawn.

## 2 The farm-investment support programme in Austria

The farm-investment programme is part of the second pillar of Common Agriculture Policy and basically concerns improving competitiveness, work conditions, animal welfare and environmental conditions. Dantler et al. (2010) analyse the farm-investment programme for the term 2000 to 2009. They find in order to achieve these goals 576 million Euros have been spent in Austria. The number of fostered farms during this period is slightly above 37,000, all mainly located in mountainous regions (see Figure 1). Consequently, forage farms (including mainly dairy and suckler-cow farms) are the main beneficiary of farm-investment payments, with a share of more than 56%. By Contrast, in the distribution of farm type of all farms in Austria, forage farms have only a share of 37% (BMLFUW, 2010). In addition, there is an over-representation of granivore farms in contrast to field crop farms. It is therefore not surprising that more than 50% of these funds foster the construction of barns mainly for dairy farming. Even though participants are mainly mountainous farms, Figure 1 illustrates a low share of participants in the western federal states of Tyrol and Vorarlberg. This might be due to specific achievements by the federal states.

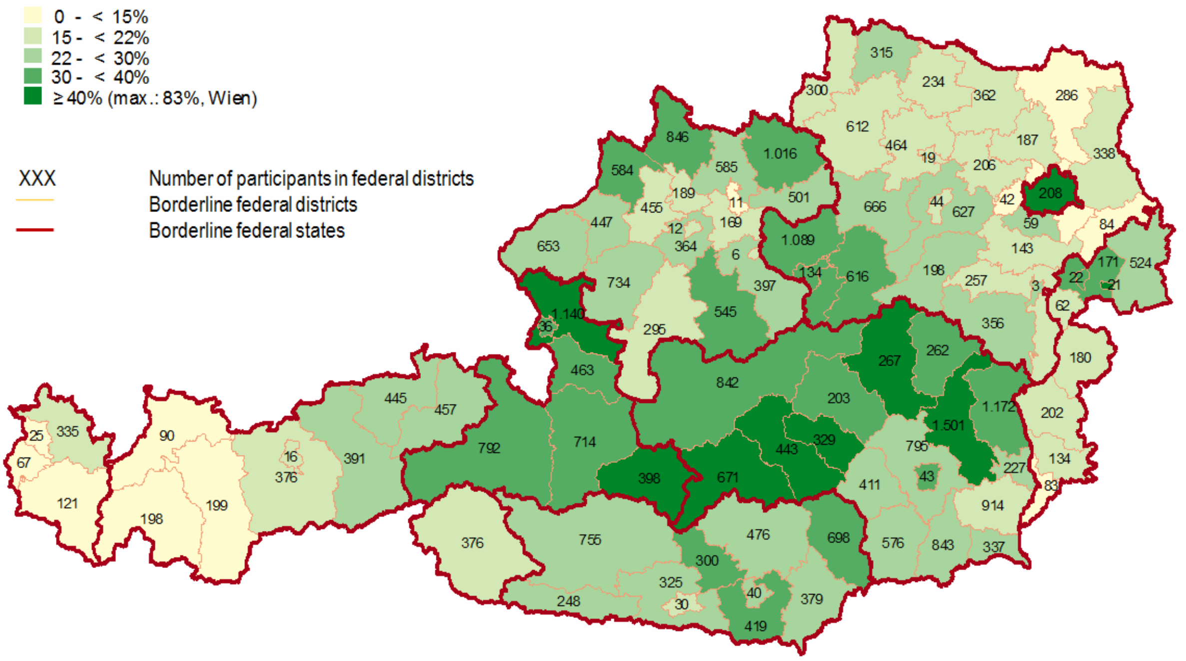

Figure 2: Share of participating farms in Austrian farm-investment programme (Source: Dantler et al., 2010)

Furthermore, Dantler et al. (2010) show that on average the share of participating farms increases for bigger farms. Hence the means of participants and non-participants differ, especially for the utilized agricultural area (UAA), total livestock units (LU) and milk quota. It is evident, therefore, that there has been a selection for participation based on structural and regional variables such as region, farm type and farm size.

# 3 Method and database

## 3.1 Microeconomic evaluation

The calculation of causal effects underlies the framework of potential outcome, which is also called the Roy (1951) - Rubin (1974) - model (Caliendo and Hujer, 2006). Through this the causal effect ($\Delta_i$) for one individual can be computed by comparing the outcome under the situation of participation ($Y_i^1$) and the outcome under the situation without participation ($Y_i^0$).

$$\Delta_i = Y_i^1 - Y_i^0 \tag{1}$$

To calculate the average effect of one measure, several evaluation parameters can be used. We use the Average Treatment Effect on the Treated (ATT) as evaluation parameter, which compares the mean actual outcome of participants and their outcome without participation (counterfactual outcome). The ATT is a very commonly used parameter and "focuses directly on the effects on those for whom the programme is actually intended" (Caliendo and Hujer, 2006). Furthermore the outcome might help to decide whether the programme is successful or not by comparing it to the programme costs (Heckmann et al., 1999).

With this a fundamental problem of microeconomic evaluation arises: it is not possible to observe the outcome of participants without participation and consequently a "*counterfactoral outcome*" (Henning and Michalek, 2008). Thus, in our case we know the income of a participating farm, but we do not know the income of this farm if it had not invested in new technology or buildings and received payments. To deal with the fundamental evaluation problem in experimental studies, the counterfactual outcome can easily be replaced by the outcome of non-participants. As experimental studies should not be used for ex-post evaluation of policy measures, we have to rely on an empirical framework (Henning and Michalek, 2008). Empirical studies carry the problem of selection bias, as individuals decide voluntarily on participation and systematic differences between both groups are evident. These differences must be controlled in order to identify the causal effects of political programmes.

There are several approaches available for solving these problems. Whereas the estimators *matching* and *regression* construct the counterfactual outcome based on observables, *difference-in-difference estimator*, *instrumental variables* and *selection models* allow for selection on unobservable variables (Caliendo and Hujer, 2006). In this paper a combination of *matching* and *difference-in-difference estimator*, the so-called "*conditional difference-in-difference estimation*" is applied.

### 3.2 Matching approach

*Matching* is a non-parametric approach and follows the Conditional Independence Assumption (CIA) to find an adequate control group. Based on Rubin (1977), the CIA assumes that under a given set of observable covariates (X), the outcome of one individual is independent of treatment or non-treatment. For this, it is necessary to identify covariates, which influence the outcome and the decision to participate but is not influenced by treatment (Reinowski, 2008). On this condition, pairs consisting of participants and controls are built, and a control group which is similar to the participant group is generated. This should lead to a reduction of systematic mean differences between these groups. Therefore the ATT with the reduced bias can be computed, as the difference of the mean outcome of participants and controls:

$$ATT=\sum_{i=1}^{n}(Y_i^1 \mid X)/n_i - \sum_{j=1}^{n}(Y_j^0 \mid X)/n_j \tag{2}$$

For *matching* a distance function and an algorithm are needed to identify similar controls for participants (Dettemann et al., 2010). In this paper we use a *direct radius-matching*, where participants are compared with all non-participants directly on selected covariates (X). Similarity is either established by total equity (nominal covariates) or by using a calliper (metric covariates). There can be several non-participants serving as control for one participant and one non-participant can also be found as control for more than one participant.

The used approach is favourable when a small number of covariates is used (Gensler et al., 2005; Schmidt, 1999) and has the advantage of self-defined similarity of participants and controls (Caliendo und Kopeinig, 2005). These advantages can also be seen as a critical point and is of high sensitivity, as the self-defined callipers rely on assumptions and might lead to dissimilarity or a big loss of controls and participants. Furthermore, problems arise with too few, wrong or too many covariates. As the chosen covariates should describe the decision to participation as good as possible, too few and wrong covariates would violate this assumption. Contrary, too many covariates lead to a reduced success in finding controls and an increase in time effort when this approach is applied. It is therefore of great importance to acquire the theoretical and practical background in order to choose the appropriate covariates. This can be done by analysing the distribution of the selected measure payments.

### 3.3 Difference-in-Difference Estimator

One of the drawbacks of the *matching* method is that it only conditions on observable covariates and leaves out hidden biases from unobservable covariates (Ankarali et al., 2009). To overcome this problem, Smith and Todd (2005) recommend the implementation of a *difference-in-difference (DiD)* estimator. The *DiD* relies on the assumption that the differences of participants and non-participants are similar every time and is computed as the difference of the progress of the participant and the non-participant from one point before (t') to one point after (t) the time of treatment ($t_T$) (Heckmann et al., 1998). By implementing the factor time and the before- and after-estimation in the analyses, we can monitor for unobservable, linear and time-invariant effects such as price fluctuations (Gensler et al., 2005). The combination of *matching* and *DiD* results in the *Conditional difference-in-*difference (CDiD) estimation and the used formula can be written as

$$ATT = \sum_{i=1}^{n}(Y_{i,t} - Y_{i,t'})/n_i - \sum_{j=1}^{n}(Y_{j,t} - Y_{j,t'})/n_j \qquad t' < t_T < t \qquad (3)$$

### 3.4 Assumptions and Database

Our analysis matches participants and non-participants on the following variables: organic farming, minor agricultural production areas, mountain farm zones, utilized agricultural area (UAA), share of arable land, livestock and milk quota. These covariates are chosen because of the specific distribution of farm-investment payments focusing on mountainous dairy farms (see Section 2). For the metric variables, a calliper is set to 15% or 5 hectare for UAA, 2 LU for livestock and 5 tons for milk quota (see Table 1). The other covariates are nominal or binary. As these callipers are of high sensitivity, we changed the callipers in further analyses to a narrower (5%) and a wider (25%) scenario to compare the quality of *matching* as well as the outcome using sensitivity analysis.

Table 1: Used covariates and callipers

| Covariates (2000) | Callipers |
|---|---|
| Organic farming | Dummy |
| Minor agricultural production area | Dummy |
| Mountain farm zones | Dummy |
| UAA (ha) | +/-15%, or +/-5 ha |
| Share of arable land (ha) | +/-15% |
| Livestock (LU) | +/-15%, or +/-2 LU |
| Milk quota (t) | +/-15%, or +/-5 tons |

For our analysis, the pre-treatment situation is in 2000, post-treatment is 2008 and the treatment itself has taken place between 2002 and 2006. The applied pre-treatment estimation for *matching* ensures that the *matching* covariates are not influenced. The two-year gap before treatment can be verified, since the year of treatment is the year of payment and the investment usually happens one or two years before payment. The two-year gap after treatment gives the farm time to use the new investment to its full. We use a panel data from 2000 to 2009 of IACS database for the Upper Austria region. We find 3,106 farms who have participated in the farm-investment support programme at least once between 2002 and 2006 and 19,081 farms who have not participated between 2000 and 2009. Participants and non-participants are matched based on the year 2000.

## 4 Empirical Results

### 4.1 Results of Matching

The *matching* approach applied here identified 2,514 pairs of participants and controls, whereas the controls consist on average of 17 non-participating farms. Table 1 shows the mean differences of selected variables of participants and non-participants before *matching* and after *matching*. Through *matching* the differences are smaller than before and not statistically significant at the 10% level any more. *Matching* can be considered successful when the mean differences are not significant (Diwish et al., 2009). The last row in Table 2 illustrates the means of the 592 participants not selected. For those, no similar control was found because either all similar farms are participants or the calliper of matching was too small.

Table 2: Mean values of variables for participants and non-participants before and after matching

| | **Before matching** | | | **After matching** | | | |
|---|---|---|---|---|---|---|---|
| **Variable (2000)** | Participants | Non-participants | | Selected participants | Controls | | Not selected participants |
| Number of farms | 3.106 | 19.081 | | 2.514 | 2.514 | | 592 |
| Share of organic farming | 10% | 8% | *** | 6% | 6% | n.s. | 28% |
| Utilised agricultural area (ha) | 23.24 | 15.60 | *** | 21.85 | 21.47 | n.s. | 29.15 |
| Arable land (ha) | 12.83 | 8.80 | *** | 12.64 | 12.41 | n.s. | 13.64 |
| Livestock (LU) | 31.57 | 17.64 | *** | 30.13 | 29.67 | n.s. | 37.68 |
| Cattle, sheep, goats (LU) | 22.79 | 13.32 | *** | 20.43 | 20.59 | n.s. | 32.82 |
| Pigs (LU) | 8.11 | 3.75 | *** | 9.09 | 8.53 | n.s. | 3.95 |
| Dairy cows (LU) | 11.49 | 6.46 | *** | 10.18 | 10.27 | n.s. | 17.03 |
| Suckler cows (LU) | 0.62 | 0.66 | n.s. | 0.49 | 0.57 | n.s. | 1.20 |
| Milk quota (t) | 46.19 | 22.18 | *** | 38.03 | 37.44 | n.s. | 80.85 |
| Programme payments (Euro)[1] | 17,010 | - | | 15,959 | - | | 21,471 |

[1)] payment of farm-investment programme (measure 121) from 2000 to 2009
t-Test: * p < 0,1; ** p < 0,05; *** p < 0,01; n.s. = not significant

## 4.2 Results of Conditional Difference-in-Difference Estimation

These 2,514 pairs, therefore, are used for computing the ATT based on the CDiD. A positive (negative) and statistically significant value for ATT shows that the farm-investment programme has a positive (negative) effect on structural farm growth. The results in Table 3 show different developments between participants and control. While participants could increase their utilised agricultural area (UAA) by 3.2 hectares, controls increased only by 0.6 hectares. A more dramatic difference can be found in livestock, where participants increased their livestock by 6.3 livestock units (LU) and controls reduced it by 2.6 LU. Even though both groups reduced their number of dairy cows from 2000 to 2008, the effect was still positive as the reduction in participating farms is less. The computed ATT is in all variables positive and statistically significant at the 1% level and shows an acceleration of structural change through investment programme in Upper Austrian farming.

Table 3: Mean developments of participants and controls from 2000 to 2008 for selected variables and the ATT-effect

| **variable (2000)** | **selected participants** (1) | **Controls** (2) | **ATT** (1)-(2) | |
|---|---|---|---|---|
| Number of farms | 2.514 | 2.514 | | |
| UAA (ha) | 3.2 | 0.6 | **2.6** | *** |
| Arable land (ha) | 2.3 | 0.7 | **1.6** | *** |
| Livestock (LU) | 6.3 | -2.6 | **8.9** | *** |
| Cattle, sheep, goats (LU) | 1.9 | -2.0 | **3.9** | *** |
| Pigs (LU) | 4.0 | -0.5 | **4.5** | *** |
| Dairy cows (LU) | -0.7 | -1.5 | **0.8** | *** |
| Suckler cows (LU) | 1.5 | 0.4 | **1.1** | *** |
| Milk quota (t) | 8.2 | 0.8 | **7.4** | *** |
| Programme payments (Euro)[1] | 15,959 | - | | |

1) payment of farm-investment programme (measure 121) from 2000 to 2009
t-Test: * p < 0,1; ** p < 0,05; *** p < 0,01; n.s. = not significant

Even though those effects are of positive value, mean values do not necessarily count for all participating individuals. Given the reality of a heterogeneous Austrian farming structure and the heterogeneous goals of the programme, heterogeneous effects are expected. This means that each participant carries important information about the treatment effect and each loss of participants might increase the bias (Augurzky und Kluve, 2004). The *matching* approach allows the differentiation of effects for different groups of farms. For example, we analyse the different effects on different farm types. Figure 3 displays the causal effects of the farm-investment programme on forage, granivore and all farms, as these are the most common farm types (1,630 and 609 farms) in the sample. Whereas on granivore farms slightly higher effects can be observed for UAA (2.71 and 2.85 ha), the effects on total livestock (6.66 and 18.76 LU) are significant greater (see Table 5). Therefore these facts indicate higher intensification on granivore farms. Unsurprisingly, at farm-type specific variables, such as pigs or cows, the effects differ more dramatic. Most of the analysed variables show statistically significant effects for both farm types.

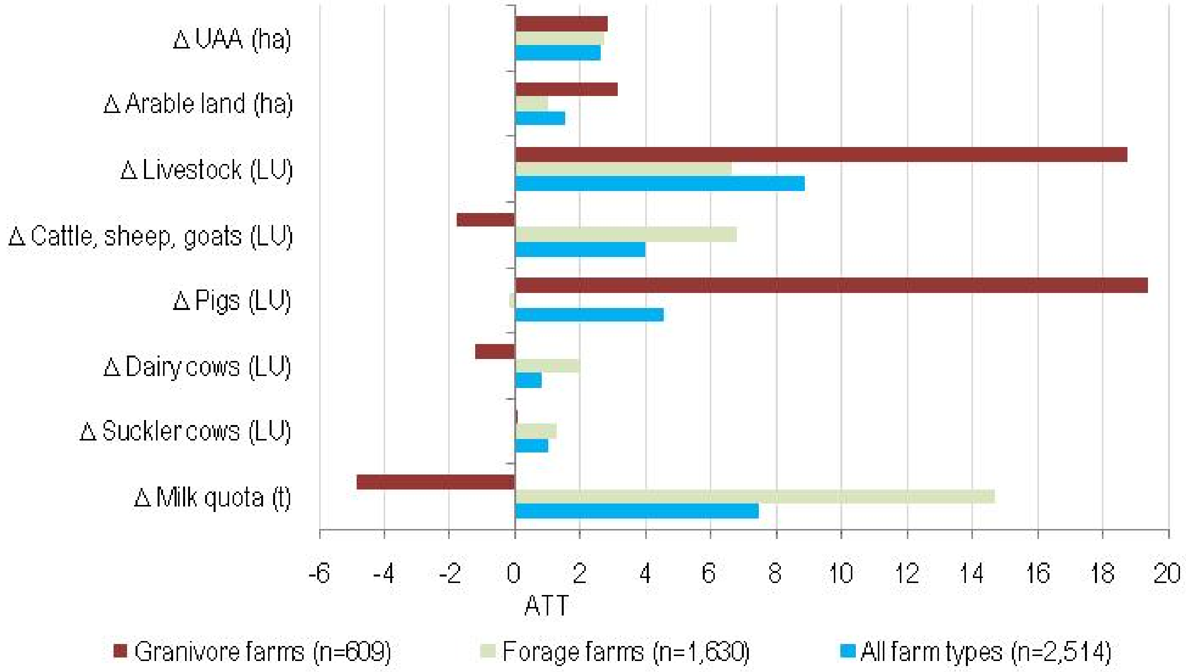

Figure 3: ATT-values for selected variables and different farm types

### 4.3 Results of Sensitivity Analysis

The loss of participants is strongly influenced by the applied calliper for selecting non-participants as controls in the *matching* approach. Using the method of sensitivity analysis, we change the calliper and create next to the base-scenario (15% callipers) one narrower (5% callipers) and one wider (25% callipers) scenario. Through changing the calliper, the number of *matching* pairs goes down in “narrow” to 1,003 (40%) and rises in “wide” up to 2,914 (116%). Augurzky und Kluve (2004) argue that callipers which are not too narrow are preferable when the heterogeneous effects of treatment are expected. But this leads to the trade-off with the similarity between participants and controls describing the quality of *matching*. To measure the quality of *matching* we use *percental bias reduction* (PBR) and the t-Test (Reinowski, 2008). The PBR can be computed for all used metric covariates by dividing the bias (mean differences between participants and controls) before *matching* with after *matching:*

$$[PBR=(bias\ before/bias\ after)*100]. \quad (4)$$

The results in Table 4 show a small increase of the quality of *matching* in all covariates with narrow callipers but a big reduction in the wider scenario. The average PBR increases from 96% to 99% under narrow and goes down to 89% under the wide scenario. When the number of pairs increase, the mean of UAA, livestock and milk quota rises as well. This is caused through the fact that the share of participants rises with farm size and therefore it is hard for big participants to find controls with narrow callipers.

Table 4: Quality of matching in three different scenarios

| | Base | | | | Narrow | | | | Wide | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Part | Contr | PRB | | Part | Contr | PRB | | Part | Contr | PRB | |
| UAA (ha) | 21,9 | 21,5 | 95 | n.s. | 20,4 | 20,4 | 99 | n.s. | 22,5 | 21,5 | 87 | *** |
| Share of arable land (%) | 51 | 51 | 94 | n.s. | 61 | 61 | 99 | n.s. | 50 | 49 | 86 | ** |
| Livestock (LU) | 30,1 | 29,7 | 97 | n.s. | 27,8 | 27,7 | 99 | n.s. | 30,8 | 29,5 | 91 | *** |
| Milk qouta (t) | 38,0 | 37,4 | 98 | n.s. | 17,6 | 17,7 | 100 | n.s. | 42,7 | 40,6 | 92 | * |
| Mean | | | 96 | | | | 99 | | | | 89 | |

PRB=Percental Bias Reduction

Furthermore, the ATT is computed as mentioned above for the two new scenarios. The results show no statistically significant differences on the 10% level to base scenario in UAA, arable land and livestock under both scenarios. Under the narrow scenario, a significant higher effect can be observed in pig production but significant lower effect in ruminant production. The opposite happens under the wider scenario, where the effects increase particularly for dairy cows and milk quota but only statistically significant on the 10% level for the latter (see Figure 4 and Table 6). Next to the increase in dairy-farming specific variable means in the wider scenario, higher effects in dairy farming can be observed, since narrow callipers exclude large and structurally growing dairy farms, which are the main receiver of investment payments. Hence the effect of the farm-investment programme would be biased downwards under narrow callipers and would lead to inaccurate conclusions about the farm-investment programme. Callipers should also not be too wide, as the quality of matching is still very important.

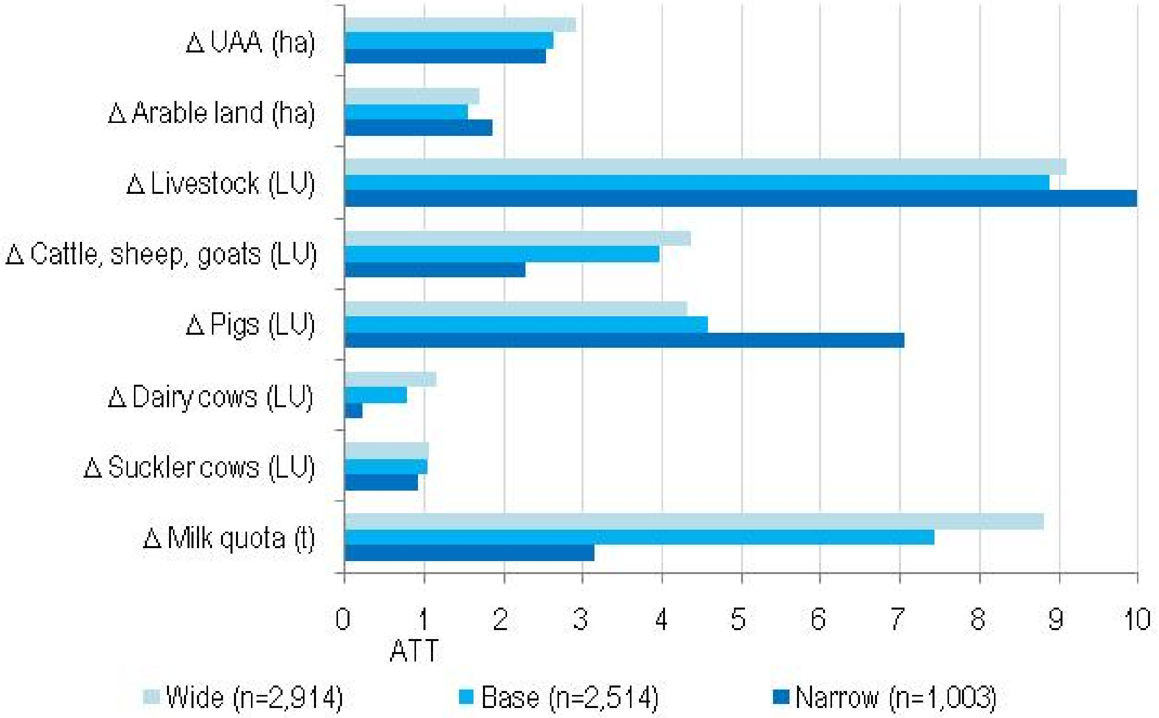

Figure 4: ATT-values for selected variables in different scenarios

## 5 Discussion and conclusions

Based on an increasing need for the quantitative evaluation of farm policy measures, we analyse the application of the *matching* method to evaluate rural-development measures. Whereas the *matching* method is commonly applied in medicine and macroeconomics (Gensler et al., 2005), up to now there have been only a few studies concerning agricultural policy (e.g. Pufahl, 2009). In this paper we use a *direct radius-matching* approach in combination with the *difference-in-difference estimator* to illustrate the structural effects of the farm-investment programme in the Upper-Austria region. As the approach we use was originally applied to an evaluation for the Austrian life ministry, it was our intention to obtain transparency for the methodology used and to communicate the results for our non-scientific client.

The methodology identifies for all participants similar controls and compares their development from a before- to an after-situation. Gensler et al. (2005) argue that this approach is strongly based on important and sensitive assumptions. As the pairs are build directly on covariates, one of the assumptions is the selection of those. It is necessary to identify those variables which have the greatest influence on the decision to participate. But the number of covariates used is restricted, since matching success decreases and time effort increases with each variable, to a dramatic extent. This drawback can be faced by pooling information and applying covariates which are plausible for the institutional environment, in which the study is carried out (Lechner, 2002). One further assumption is the need to define the similarity of participants and controls. For this we implement a calliper (upper and lower bounds) for each metric covariate, which allows the adapting to individual data and the aim of study. However, this approach shows a trade-off regarding the similarity within the pairs and the drop-out of participants. The latter might lead to a loss of information, especially when effects are heterogeneous. Therefore we apply additional analyses to demonstrate the quality of *matching* and sensitivity of outcome in one narrower and wider scenario.

The results of our study indicate an acceleration in farm growth caused by the farm-investment support measure. We find significant differences in the development from participants against their controls in UAA, livestock production (ruminants and pigs) and dairy production. In particular, the average effects of total livestock production (8.9 LU) are enormous. Alongside the average effects on all participants, we also examine the effects regarding different farm types. This analysis emphasises the heterogeneous effects of rural-development measures as granivore farms show higher effects regarding total livestock production than forage farms. We can conclude, therefore, that heterogeneous effects are evident due to different farm structure and strategies, as well as the diversity of goals in the analysed measure, and might also appear in other rural-development measures. Pufahl (2009) finds that this also occurs in the German agro-environment programme. This fact has to be considered in further evaluation studies and in policy making.

In our sensitivity analysis we find under narrow callipers a loss of big dairy farms and therefore a reduced effect regarding dairy cows and milk quota. In contrast, wider callipers reduce the quality of matching especially for UAA and livestock production. Consequently, the base scenario seems to be preferable as it has an overall percental bias reduction of 96% and a loss of almost 20% of participants. We would only recommend wider callipers for the covariate milk quota, as there are fewer non-participants in the group of bigger farms. Therefore we suggest that in farm-policy evaluation, callipers should definitely not be too narrow, as this leads to significant loss of information. Individual callipers can be installed, depending on the variance of the covariant.

Even though this approach is dependent on several assumptions, next to individual adjustments it allows transparency for non-scientific stakeholders in the evaluation process. This is particular necessary as practical information is important to find covariates. Furthermore, it shows the advantage of easily communicated results. We would like to stress that policy evaluation must be carried out with and for stakeholders and not only for scientists.

We acknowledge that further research has to be done on identifying covariates and their influence on participation as well as sensitivity analysis to other *matching* approaches like *propensity score matching*. As we use IACS data, only structural variables are available. The model and results can be improved by using economic and qualitative data, but hidden bias might still remain, as the decision to participate in the farm-investment support programme is often due to the need of investment. We would also point out that we never know if a farmer would have invested in, for example, new building without federal support. This illustrates the complex effects of this measure and challenges for evaluation. However, we find that the approach used, in combination with pre-studies and stakeholder information, can help towards a consistent farm-policy evaluation in rural-development programmes.

## Acknowledgements

We are thankful to the Austrian Federal Ministry of Agriculture, Forestry, Environment and Water Management (Division II/5) for funding the project on farm-investment support programme evaluation.

## 6 Literature

ANKARALI, H.C., V. SUMBULOGLU, A.C. YAZICI, I. YALUG , M. SELEKLER (2009): Comparison of Different Matching Methods in Observational Studies and Sensitivity Analysis: The Relation Between Depression and STAI-2 scores. In: Expert Systems with Applications 36, 1876-1884.

AUGURZKY, B., J. KLUVE (2004): Assessing the Performance of Matching Algorithms - When Selection into Treatment Is Strong. Forschungsinstitut zur Zukunft der Arbeit, Diskussionspapier No. 1301.

BECK, M., T. DOGOT (2006): The Use of Impact Indicators for the Evaluation of Farm Investment Support – A Case Study Based on the Rural Development Programme for Wallonia (2000 – 2006). In: BERGSCHMIDT A., W. DIRKSMEYER, J. EFKEN, B. FORSTNER, I. UETRECHT (ed.) Proceedings of the European Workshop on the Evaluation Farm Investment Support, Investment Support for Improvement of Processing and Marketing of Agricultural Products. Arbeitsberichte des Bereichs Agrarökonomie 03/2006, 69-77.

BMLFUW (2010): Grüner Bericht 2010. Wien.

CALIENDO, M., R. HUJER (2006): The Mircoeconometric Estimation of Treatment Effects – An Overview. In: Allgemeines Statistisches Archiv 90, 199-215.

DIRKSMEYER, W., B. FORSTNER, A. MARGARINA, Y. ZIMMER (2006): Aktualisierung der Zwischenbewertung des Agrarinvestitionsförderprogramms (AFP) in Deutschland für den Förderzeitraum 2000 bis 2004. Länderübergreifender Bericht. Bundesanstalt für Landwirtschaft (FAL), Braunschweig.

DIWISCH, D.S., P. VOITHOFER, CH.R. WEISS (2009): Succession and Firm Growth: Results from a Non-parametric Matching Approach. In: Small Business Economics 32:45-46

DANTLER, M., S. KIRCHWEGER, M. EDER, J. KANTELHARDT (2010): Analyse der Investitionsförderung für landwirtschaftliche Betriebe in Österreich. Universität für Bodenkultur, Institut für Agrar- und Forstökonomie, Wien.

DETTEMANN, E., C. BECKER, CH. SCHMEIßER (2010): Is there a Superior Distance Function for Matching in Small Samples? Institut für Wirtschaftsforschung Halle, Diskussionspapier Nr. 03.

FORSTNER, B., W. DIRKSMEYER, A. BERGSCHMIDT, H. EBERS, A. FITSCHEN-LISCHEWSKI, A. MARGARIAN, J. HEUER (2008): Ex-Post-Bewertung des Agrarinvestitionsförderungsprogramms (AFP) für den Förderzeitraum 2000 bis 2006. Baden-Württemberg. http://www.vti.bund.de/de/ institute/bw/publikationen/sonstige/ex_post/ badenwuerttemberg_bericht_de. pdf (25.05.2010).

GENSLER, S., B. SKIERA, M. BÖHM (2005): Einsatzmöglichkeiten der Matching Methode zur Berücksichtigung von Selbstselektion. In: Journal für Betriebswirtschaft 55, 37-62.

HECKMAN, J. J., H. ICHIMURA, J. A. SMITH, P. E. TODD (1998): Characterizing Selection Bias Using Experimental Data. In: Econometrica 66 (5), 1017–1098.

HECKMAN, J. J., R. J. LALONDE, J. A. SMITH (1999): The Economics and Econometrics of Active Labor Market Programs. In: Ashenfelter, O., D. E. Card (Hrsg.): Handbook of Labor Economics. Band III, Amsterdam: Elsevier Science B.V., 1865–2097.

HENNING, C.H.C.A., J. MICHALEK (2008): Ökonometrische Methoden der Politikevaluation: Meilensteine für eine sinnvolle Agrarpolitik der 2. Säule oder akademische Finderübung. In: Agrarwirtschaft 57(3/4), 232-243.

LECHNER, M. (2002): Mikroökonomische Evaluation arbeitspolitischer Maßnahmen. University of St.Gallen, Department of Economics, Diskussionspapier No. 2002-20.

PFEFFERLI, S. (2006): Impact Analysis of Investment Support for Agricultural Buildings in Switzerland. In: BERGSCHMIDT A., W. DIRKSMEYER, J. EFKEN, B. FORSTNER, I. UETRECHT (ed.) Proceedings of the European Workshop on the Evaluation Farm Investment Support, Investment Support for Improvement of Processing and Marketing of Agricultural Products. Arbeitsberichte des Bereichs Agrarökonomie 03/2006, 69-77.

PUFAHL, A. (2009): Empirische Wirkungsanalyse direkter Transferzahlungen – am Beispel von Agrarumweltmaßnahmen und der Ausgleichszulage für benachteiligte Gebiete. Dissertation, Georg-August Universität Göttingen.

PUFAHL, A., CH.R. WEISS (2009): Farm Structure and the Effects of Agri-Environmental Programs: Results from a Matching Analysis for European Countries. 111 EAAE-IAAE Seminar ‚Small Farms: decline or persistence‘. University of Kent, Canterbury;UK 26th – 27th June 2009

REINOWSKI, E. (2008): Matching kleiner Stichproben. Ein Vergleich verschiedener Verfahren. Dissertation, Martin-Luther-Universität Halle-Wittenberg.

ROY, A. (1951): Some Thoughts on the Distribution of Earnings, Oxford Economic Papers, 3, 135-145.

RUBIN, D.B. (1974): Estimating Causal Effects to Treatments in Randomised and Nonrandomised Studies, Journal of Educational Psychology, 66, 688-701.

RUBIN, D.B. (1977). Assignment to Treatment Group on the Basis of a Covariate. Journal of Educational Statistics, 2, 1–26.

SALHOFER, K., G. STREICHER (2005): Self-selection as a Problem in Evaluation Agri-environmental Programs. In: Ortner, K.M. (ed.): Assesing Rural Development Policies of the Common Agricultural Policies. Selection of Papers from the 87th EAAE-Seminar, Kiel: Vauk, 203-213

SCHMIDT, CH.M. (1999): Knowing what Works - the Case for Rigorous Program Evaluation. Forschungsinstitut zur Zukunft der Arbeit, Bonn, Diskussionspapier Paper No. 77.

SMITH, J.A., P.E. TODD (2005): Does Matching Overcome LaLonde’s Critique of nonexperimental Estimators? In: Journal of Econometrics 125, 305-353.

STRIEWE, L., J.P. LOY, U. KOESTER (1996): Analyse und Beurteilung der einzelbetrieblichen Investitionsförderung in Schleswig-Holstein. In: Agrarwirtschaft 45 (12), 423-434.

## Appendix

Table 5: ATT-values for selected variables and different farm types

| | Forage farms | | Granivore farms | |
|---|---|---|---|---|
| Number of farms | 1,630 | | 609 | |
| UAA (ha) | 2.71 | *** | 2.85 | *** |
| Arable land (ha) | 1.03 | *** | 3.15 | *** |
| Livestock (LU) | 6.66 | *** | 18.76 | *** |
| Cattle, sheep, goats (LU) | 6.77 | *** | -1.80 | *** |
| Pigs (LU) | -0.16 | ** | 19.35 | *** |
| Dairy cows (LU) | 1.99 | *** | -1.22 | *** |
| Suckler cows (LU) | 1.28 | *** | 0.10 | n.s. |
| Milk quota (t) | 14.67 | *** | 4.88 | *** |

t-Test: * p < 0,1; ** p < 0,05; *** p < 0,01; n.s. = not significant

Table 6: ATT-values for selected variables in different scenarios

| | Base | Narrow | | Wide | |
|---|---|---|---|---|---|
| Number of farms | 2,514 | 1,003 | | 2,914 | |
| UAA (ha) | 2.6 | 2.5 | n.s. | 2.9 | n.s. |
| Arable land (ha) | 1.6 | 1.9 | n.s. | 1.7 | n.s. |
| Livestock (LU) | 8.9 | 10.0 | n.s. | 9.1 | n.s. |
| Cattle, sheep, goats (LU) | 4.0 | 2.3 | *** | 4.4 | n.s. |
| Pigs (LU) | 4.6 | 7.0 | *** | 4.3 | n.s. |
| Dairy cows (LU) | 0.8 | 0.2 | ** | 1.2 | * |
| Suckler cows (LU) | 1.0 | 0.9 | n.s. | 1.1 | n.s. |
| Milk quota (t) | 7.4 | 3.1 | *** | 8.8 | n.s. |

t-Test (mean differences to base): * p < 0,1; ** p < 0,05; *** p < 0,01; n.s. = not significant