Environment. Technology. Resources, Rezekne, Latvia
Proceedings of the 11th International Scientific and Practical Conference. Volume I, 28-34

Displacement Effects of Latvian Rural Development Programme 2007-2013

Elita Benga, Juris Hāzners, Zaiga Miķelsone
Institute of Agricultural Resources and Economics, Research Center of Economics, Division of Rural Development Evaluation. Address: Struktoru street 14, Riga, Latvia, LV-1039.

Abstract. Periodic evaluation of EU Member States Rural Development Programme (RDP) specific policy interventions is considered crucial in policy development. The main reasons for the evaluation of specific policy interventions are the assessment of a programme's impact, the improvement of programme management and administration, identification of necessary improvements in the delivery of interventions and meeting the accountability. The core question to be answered in programme evaluation is whether the stated objectives are accomplished by particular intervention (support or „treatment" provided to programme participants). The main problem in the process of evaluation is the assessment of the counterfactual outcome by modelling the situation where treatment is absent. The counterfactual outcome has to be estimated by statistical methods as it is usually not observed. General equilibrium effects occur when a programme affects units other than its participants. The most important possible impacts are the substitution effect and the displacement effect. Displacement effects are unplanned and indirect. They usually play a more important role in the evaluation at the programme level than in the evaluation of RDP individual measures. Displacement effect is the programme effect that occurs in a programme area at expense of another area. It takes place if farms located in one geographical area, which is not a subject to RD support, becomes adversely affected by a support provided to farms located in another geographically area. The existing study provides an assessment of the displacement effects on the employment in unsupported units at the programme level after the net effects on the employment calculated at the measure level are aggregated over the entire programme.

Keywords: policy evaluation, propensity score matching, counterfactual analysis, displacement effects.

I. INTRODUCTION

Periodic evaluation of EU Member States Rural Development Programme (RDP) specific policy interventions is considered crucial in policy development. The main reasons for the evaluation of specific policy interventions are the assessment of a programme's impact, the improvement of programme management and administration, identification of necessary improvements in the delivery of interventions and meeting the accountability. According to the EU definition, programme evaluation is a process that culminates in a judgment (assessment) of policy interventions according to their results, impacts and the needs. In the case of rural development (RDP) programmes, EU regulations distinguish between ex-ante, midterm, ex-post and ongoing evaluations. The existing study is considered a part of an ongoing evaluation which would provide the grounds for the ex-post evaluation of Latvian Rural Development Programme 2007-2013. The core question to be answered in programme evaluation is whether the stated objectives are accomplished by particular intervention (support or „treatment" provided to programme participants). The main problem in the process of evaluation is the assessment of the counterfactual outcome by modelling the situation where treatment is absent. The counterfactual outcome has to be estimated by statistical methods as it is usually not observed.

The core element of the EC evaluation framework are Common Evaluation Questions (CEQ) pre-defined by the EC and programme-specific questions defined by national programme authorities. The evaluation questions focus on a direct effect of the RD programme on specific result indicators. The answer to the crucial evaluation question on the contribution of the programme to the growth of the whole rural economy has to be provided by the measuring the net effects of the programme support on the Gross Value Added (GVA) in supported units.

As mostly the evaluation is focused on the assessment of the direct and planned effects of the policy interventions, the evaluation can produce biased results. Standard propensity score matching methods assume that outcomes for non-participants in the control group are not affected by the programme (no general equilibrium effects). If general equilibrium effects had occurred during the implementation of a given RD programme with substantial impact (positive or negative) on farms which did not participate in this programme, partial equilibrium evaluation techniques such as standard

ISSN 1691-5402

http://dx.doi.org/10.17770/etr2017vol1.2662

PSM would produce biased estimates of programme effects.

General equilibrium effects occur when a programme affects persons/enterprises other than its participants [15]. The most important possible impacts are the substitution effect and the displacement effect [2]. Displacement effects are unplanned and indirect. They usually play a more important role in the evaluation at the programme level than in the evaluation of RDP individual measures. Displacement effect is normally defined as the effect obtained in favour of direct programme beneficiaries but at the expense of units that do not qualify or participate in a given intervention. It occurs if, due to support provided from RDP employment shifts at the detriment of non-supported or non-eligible units usually located in close neighbourhood of units directly supported by a given programme.

The existing study provides an assessment of the displacement effects on the employment in unsupported units at the programme level after the net effects on the employment calculated at the measure level are aggregated over the entire programme.

II. MATERIALS AND METHODS

To measure causal effects of programme or policy intervention, a potential outcome model is appropriate. The model was proposed by Roy [13] and further developed by Rubin [14] and Holland [5]. Using the potential outcome model, the causal effect of a given programme on unit can be expressed with basic evaluation formula:

$$e_i = Y_i(1) - Y_i(0), \quad (1)$$

where:

$Y_i(1)$ - potential outcome for unit i in case of participation in RDP (programme participants),

$Y_i(0)$ - potential outcome for unit i in case of non participation in RDP (counterfactual),

e_i - the effect of programme participation on unit i, relative to effect of non-participation on the basis of a response variable Y .

In evaluation it is relatively easy to obtain for programme beneficiaries the information about $Y_i(1)$ but it is very difficult to estimate $Y_i(0)$ which for programme beneficiaries is not directly observable.

The outcome for a participating unit can be observed directly and it is expressed by formula:

$$e_i = (Y2 - Y1), \quad (2)$$

where:

$Y1$ - value of the outcome variable at programme starting period for a participating unit,

$Y2$ - value of the outcome variable at programme ending period.

The outcome for the same unit without the participation can be interpreted as a result of other factors which may simultaneously affect observable impact variables and it is expressed by formula:

$$e_i = Y3 - Y1, \quad (3)$$

$Y3$ - value of the outcome variable for the same unit without a participation.

The unit can only be observed in one of two possible situations: being supported (participating) or not-supported (without a participation) which means that the real programme effect can be expressed as a difference between the outcome with a participation and outcome without a participation:

$$e_i = (Y2 - Y1) - (Y3 - Y1) = (Y2 - Y3), \quad (4)$$

The real programme effect (*Y2 – Y3)* cannot be directly observed.

The effectiveness of interventions on outcomes of interest can be evaluated by propensity score matching (PSM). Multiple regression is the most common method for estimating the programme support effect. PSM is a rigorous nonexperimental method. The data for PSM usually are pooled in a panel both from programme participants and non-participants. The non-participating or „untreated" units constitute the „control" group while participants are included in „treatment" group. The information from control group is used to assess what would be the outcome of interest for participants in the absence of the programme. The difference in outcomes for both groups is evaluated by comparison of relatively similar units in these groups. To successfully mitigate the potential bias, unit matching has to be based not on a single or a few characteristics but on a full range of available covariates that have potential impact. The propensity score is then defined as the probability of receiving the treatment by the given unit. Thus the matching is reduced to a single variable, and matching on entire set of covariates is no longer necessary. The method was developed by Rosenbaum and Rubin [11]. They introduced balancing score as a function of covariates that provides the same distributions of covariates in both groups. Imbens [6] suggested four step procedure for implementing the PSM:

1. selection of observational covariates and estimation of propensity scores,
2. stratification of propensity scores and testing of balancing properties in each block,
3. calculation of the Average Treatment on Treated (ATT) by matching,
4. sensitivity test for robustness of estimated ATT effects.

If the balancing properties of covariates are not satisfied in all strata, the test has to be repeated with different number of strata. If the balancing properties are not satisfied again, estimation of propensity scores has to be repeated with modified list of covariates by adding higher order (squared) covariates. After getting all covariates balanced in every stratum, causal effects can be estimated by nearest neighbor matching (NNM), radius matching (RM) or kernel matching (KM).

NN matching computes the ATT by finding the unit in the control group whose propensity score is nearest (absolute value of difference is minimal) for every unit in treatment group. Larger number of comparison units from control group decreases the variance of the estimator. At the same time, the bias of the estimator increases. Furthermore, one needs to choose between matching with replacement and matching without replacement [3]. When there are few comparison units, matching without replacement will force us to match treated units to the comparison ones that are quite different in propensity scores. This enhances the likelihood of bad matches (increase the bias of the estimator), but it could also decrease the variance of the estimator. Thus, matching without replacement decreases the variance of the estimator at the cost of increasing the estimation bias. In contrast, because matching with replacement allows one comparison unit to be matched more than once with each nearest treatment unit, matching with replacement can minimize the distance between the treatment unit and the matched comparison unit. This will reduce bias of the estimator but increase variance of the estimator.

In RM, the units in both groups are matched when the propensity scores in control group fall in the predefined radius of the units in treatment group. The larger the radius is, the more matches can be found. More matches typically increase the likelihood of finding bad matches, which raises the bias of the estimator but decreases the variance of the estimator.

In KM, all units in treatment group are matched with the weighted average of all units in control group. The weights are determined by distance of propensity scores, bandwidth parameter and a kernel function. Choosing an appropriate bandwidth is crucial because a wider bandwidth will produce a smoother function at the cost of tracking data less closely. Typically, wider bandwidth increases chance of bad matches so that the bias of the estimator will also be high. Yet, more comparison units due to wider bandwidth will also decrease the variance of the estimator.

In general, selection of the matching technique is empirical and it largely depends on the results obtained.

The PSM method first has been empirically applied by Heckman, Ichimura, Smith and Todd [4] in the estimations of training programmes on future income in the USA labor market. Subsequently, similar studies on the USA labor market were carried out by Dehejia and Wahba [3], and a few other researchers.

The modules for calculating propensity scores and matching for use in STATA software were developed by Becker and Ichino [1]. It is common first to run the pscore module which estimates the propensity scores and tests the satisfying of the balancing properties. If the balancing properties are satisfied then ATT can be estimated with one or more of the *att** modules. The modules attnd or attnw, attr and attk assume nearest neighbor, radius and kernel matching, respectively. After the calculation of ATT, the module *mhbounds* developed by Rosenbaum [12]provides sensitivity analysis with Rosenbaum bounds with Mantel and Haenszel [7] test statistic.

As the direct planned impact of the programme on employment is measured by the Annual Working Units (AWU), this result indicator is used also in measuring displacement effects. Michalek [8] proposes the following steps in the estimation of displacement effects:

- identification of supported units j in the area with high intensity of support;
- identification of non-supported units k in the area with high intensity of support, which match with units j;
- identification of non-supported units m in the area with low intensity of support, which match with units j;
- calculation of DiD-ATT between units j and units k as well as between units j and units m;

The lack of displacement effects would result in similar differences in DiD-ATT between units j and k compared with j and m.

The estimation of the indirect effects including displacement effects was effectively carried out following the procedure suggested above for Slovakian SAPARD programme [9] and farm support measures of programmes in selected countries [10].

III. RESULTS AND DISCUSSION

The data on participants and non-participants of Latvian Rural Development Programme are sourced from FADN database (Axis 1 and Axis 2 measures) and State Revenue (Axis 3 measures). The Axis 4 measures due to their specific support are not included in the assessment. The calculations are provided in two blocks for FADN data and State Revenue data as the all relevant information on programme participants and non-participants regarding their structure and performance from 2007 to 2013 differs depending upon the source.

First, as the information should cover periods before and after the implementation of the programme, certain number of relevant units was selected from both data sets. The employment was selected as a dependent variable.

As there were no districts without programme support, breakdown of districts by participation and non-participation was based upon the intensity of support.

The most common measure of the intensity of programme support is the average per capita public financing allocated to the statistical districts of the country.

The assessment of the displacement effects for the Axis1 and Axis2 measures is based upon the average per capita support provided to the districts within the Axis1 and Axis2 measures. The 119 districts are divided in two sets based on the 90% threshold of the average per capita support. The set of the high support intensity consists of 60 districts where the intensity exceeds 90% of the average. The set of the low support intensity consists of 59 districts where the intensity is below the 90% of the average.

The principal layout of the algorythm for assessment of the Axis1 and Axis2 displacement effects is mapped on Fig. 1.

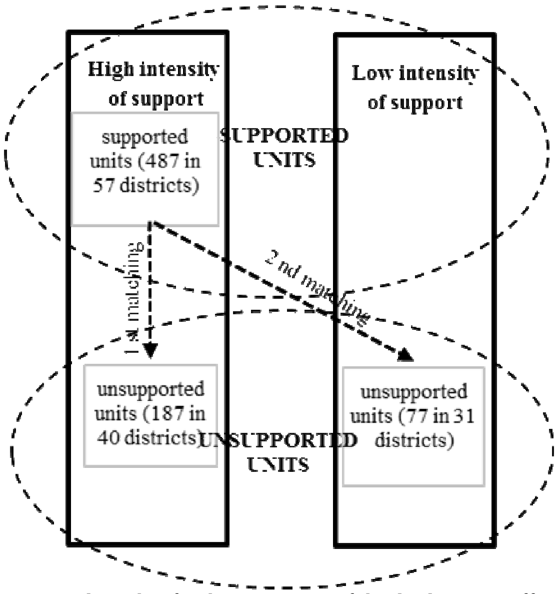

Fig. 1. Algorythm for the assessment of the displacement effects for the Axis1 and Axis2 measures

First, the 487 supported units in 57 districts with high intensity of support were identified. Second, the 187 unsupported units were identified in 40 districts with high intensity of support. Third, the 77 unsupported units were identified in 31 districts with low intensity of support.

The first matching was provided for the supported and unsupported units in the area of high support intensity. With respect to propensity score matching (PSM-DiD method), in total, 52 variables related to unit structure which were considered critical for comparability of economic performance were selected for use in matching process.

Although only 2 and 3 variables proved statistically significant at 5% and 10% level, respectively, after Logit regression, dropping the variables with lower significance levels caused a loss of balancing properties in one or more blocks. Similarly, adding of higher order covariates caused the loss of balancing properties. Therefore, the original specification of Logit function was preferred.

For the treated units, control units and total for each of iterated five blocks the computed z-value does not exceed the critical value for the 5% confidence interval.

The common support option has been selected. This restriction implies that the test of the balancing property is performed only on the observations whose propensity score belongs to the intersection of the propensity scores in both groups. With the given specification the balancing property was satisfied.

Matching with the radius (R=0.1) method was selected based upon the t-test results.

The results of the matching supported and unsupported units in the area with high intensity of support are shown in Table 1.

Table 1

Average changes in employment of supported (T=1) and non-supported (T=0) units of Axis1 and Axis2 measures in the area of high programme support

	Employment (AWU)		
	2007	2013	DiD (2014-2007)
Unmatched supported units in high intensity region (P=1) (487)	6.04	5.64	-0.39
Unmatched unsupported units in high intensity region (P=0) (187)	5.86	5.18	-0.68
Total Ø (674)	5.99	5.51	-0.47
Difference (1-0)	0.18	0.46	0.28
Difference (1-Ø)	0.05	0.13	0.08
Matched supported units in high intensity region (M=1) (487)	6.04	5.64	-0.39
Matched unsupported units in high intensity region (M=0) (187)	2.77	3.05	0.29
ATT	3.27	2.59	-0.68

The ATT effect on the employment of supported units in the area of high intensity of programme support evaluated by PSM-DiD method is negative at 0.68 AWU per unit. It means that unsupported units in this area are affected positively in terms of employment. Using the simple difference-in-differences estimator without matching would lead to an erroneous assumption that the effect on supported units is positive at 0.28 AWU.

The second matching was provided for the supported units in the area of high support intensity and the unsupported units in the area of low support intensity. With respect to propensity score matching (PSM-DiD method), in total, 52 variables related to unit structure which were considered critical for comparability of economic performance were selected for use in matching process.

Although only 9 and 11 variables proved statistically significant at 5% and 10% level, respectively, after Logit regression, dropping the variables with lower significance levels caused a loss of balancing properties in one or more blocks. Similarly, adding of higher order covariates caused the loss of balancing properties. Therefore, the original specification of Logit function was preferred.

For the treated units, control units and total for each of iterated five blocks the computed z-value does not exceed the critical value for the 5% confidence interval.

The common support option has been selected. This restriction implies that the test of the balancing property is performed only on the observations whose propensity score belongs to the intersection of the propensity scores in both groups. With the given specification the balancing property was satisfied.

Matching with the radius (R=0.01) method was selected based upon the t-test results.

The results of the matching supported units in the area of high intensity of support to unsupported units are shown in Table 2.

Table 2
Average changes in employment of supported (T=1) units in the area with high level of support and non-supported (T=0) units in the area of low level of support of Axis1 and Axis2 measures

	Employment (AWU)		
	2007	2013	DiD (2014-2007)
Unmatched supported units in high intensity region (P=1) (61)	6.04	5.64	-0.39
Unmatched non-supported units in low intensity region (P=0) (23)	7.61	5.05	-2.56
Total Ø (84)	6.25	5.56	-0.69
Difference (1-0)	-1.57	0.59	2.16
Difference (1-Ø)	-0.21	0.08	0.30
Matched supported units in high intensity region (M=1) (58)	6.04	5.64	-0.39
Matched non-supported units in low intensity region (M=0) (23)	2.85	2.31	-0.54
ATT	3.19	3.33	0.14

The ATT effect on the employment of supported units in the area of high intensity of programme support evaluated by PSM-DiD method is positive at 0.14 AWU per unit. It means that unsupported units in this area are affected negatively in terms of employment. Using the simple difference-in-differences estimator without matching would lead to an erroneous assumption that the effect on supported units is positive at 2.16 AWU.

As it is assumed that the lack of displacement effects would result in similar calculated effects from both matchings, the difference in the results points to considerable displacement effects. The employment situation in unsupported farms located in the areas with high intensity of support is not deteriorating relative to unsupported farms in the areas with low intensity of support.

The estimated displacement effect on the employment of supported units is calculated as a difference between ATT effects for the first and second matching. The estimated effect is positive at 0.82 AWU.

The estimated displacement effect at the axis level (Axis1 and Axis2) is calculated by multiplying the calculated difference in ATT to the total number of unsupported farms in areas with low intensity of programme support. The number of such farms amount to a 3,409. The calculated displacement effect at the axis level (Axis1 and Axis2) is positive at 2,795 AWU. It means the programme support in regions with high intensity of programme support has positively affected the employment in non-supported units in direct neighborhood of supported units.

The assessment of the displacement effects for the Axis3 measures is based upon the average per capita support provided to the districts within the Axis3 measures. The 119 districts are divided in two sets based on the 90% threshold of the average per capita support. The set of the high support intensity consists of 60 districts where the intensity exceeds 90% of the average. The set of the low support intensity consists of 59 districts where the intensity is below the 90% of the average.

First, the 202 supported units in 45 districts with high intensity of support were identified. Second, the 95 unsupported units were identified in 41 districts with high intensity of support. Third, the 123 unsupported units were identified in 43 districts with low intensity of support.

The first matching was provided for the supported and unsupported units in the area of high support intensity. With respect to propensity score matching (PSM-DiD method), in total, 12 variables related to unit structure which were considered critical for comparability of economic performance were selected for use in matching process.

Although only 1 variable proved statistically significant at 5% level after Logit regression, dropping the variables with lower significance levels caused a loss of balancing properties in one or more blocks. Similarly, adding of higher order covariates caused the loss of balancing properties. Therefore, the original specification of Logit function was preferred.

For the treated units, control units and total for each of iterated five blocks the computed z-value does not exceed the critical value for the 5% confidence interval.

The common support option has been selected. This restriction implies that the test of the balancing property is performed only on the observations whose propensity score belongs to the intersection of the propensity scores in both groups. With the given specification the balancing property was satisfied.

Matching with the nearest neighbor method was selected based upon the t-test results.

The results of the matching supported units and unsupported units in the area of high intensity of support are shown in Table 3.

The ATT effect on the employment of supported units in the area of high intensity of programme support evaluated by PSM-DiD method is positive at 0.39 AWU per unit. It means that unsupported units in this area are affected negatively in terms of employment. Using the simple difference-in-differences estimator without matching would lead to an erroneous assumption that the effect on supported units is negative at 0.39 AWU.

Table 3
Average changes in employment of supported (T=1) and non-supported (T=0) units of Axis3 measures in the area of high intensity of support

	Employment (AWU)		
	2007	2013	DiD (2014-2007)
Unmatched supported units in high intensity region (P=1) (96)	3.04	3.72	0.68
Unmatched unsupported units in high intensity region (P=0) (61)	0.27	1.34	1.07
Total Ø (157)	1.96	2.80	0.83
Difference (1-0)	2.76	2.37	-0.39
Difference (1-Ø)	1.07	0.92	-0.15
Matched supported units in high intensity region (M=1) (96)	3.04	3.72	0.68
Matched unsupported units in high intensity region (M=0) (61)	0.98	1.26	0.29
ATT	2.06	2.45	0.39

The second matching was provided for the supported units in the area of high support intensity and the unsupported units in the area of low support intensity. With respect to propensity score matching (PSM-DiD method), in total, 10 variables related to unit structure which were considered critical for comparability of economic performance were selected for use in matching process.

Although only 2 and 4 variables proved statistically significant at 5% and 10% level, respectively, after Logit regression, dropping the variables with lower significance levels caused a loss of balancing properties in one or more blocks. Similarly, adding of higher order covariates caused the loss of balancing properties. Therefore, the original specification of Logit function was preferred.

For the treated units, control units and total for each of iterated five blocks the computed z-value does not exceed the critical value for the 5% confidence interval.

The common support option has been selected. This restriction implies that the test of the balancing property is performed only on the observations whose propensity score belongs to the intersection of the propensity scores in both groups. With the given specification the balancing property was satisfied.

Matching with the radius (R=0.1) method was selected based upon the t-test results.

The results of the matching supported units in the area of high intensity of support to unsupported units in the area of low intensity of support are shown in Table 4.

The ATT effect on employment of supported units in the area of high intensity of programme support evaluated by PSM-DiD method is positive at 0.32 AWU per unit. It means that unsupported units in this area are affected negatively in terms of employment. Using the simple difference-in-differences estimator without matching would lead to an erroneous assumption that the effect on supported units is positive at 0.23 AWU.

Table 4
Average changes in employment of supported (T=1) units in the area with high level of support and non-supported (T=0) units in the area of low level of support of Axis3 measures

	Employment (AWU)		
	2007	2013	DiD (2014-2007)
Unmatched supported units in high intensity region (P=1) (61)	3.04	3.72	0.68
Unmatched non-supported units in low intensity region (P=0) (23)	0.32	0.78	0.45
Total Ø (84)	2.33	2.95	0.62
Difference (1-0)	2.71	2.94	0.23
Difference (1-Ø)	0.71	0.77	0.06
Matched supported units in high intensity region (M=1) (58)	0.80	1.70	0.90
Matched non-supported units in low intensity region (M=0) (23)	0.44	1.01	0.58
ATT	0.36	0.68	0.32

The estimated displacement effect on AWU of supported units is calculated as a difference between ATT effects for the first and second matching. The estimated effect is negative at 0.07 AWU. The employment situation in unsupported farms located in the areas with high intensity of support is slightly deteriorating relative to unsupported farms in the areas with low intensity of support.

The estimated displacement effect at the axis level (Axis3) is calculated by multiplying the calculated difference in ATT to the total number of unsupported enterprises in areas with low intensity of programme support. The number of such enterprises amounts to a 13,721. The calculated displacement effect at the axis level (Axis3) is negative at 968 AWU. It means the programme support in regions with high intensity of programme support has negatively affected the employment in non-supported units in direct neighborhood of supported units.

The estimated displacement effect at the programme level is calculated as the aggregate of the effects calculated ar the axis level. Total net impact of the programme on employment considering the previously estimated direct impact at sector level (policy intervention focus area) is shown in Table 5.

Table 5
Total programme net impact on employment (AWU)

	Direct impact	Displacement effects	Net impact
Farm modernization	-2705	2795	2369
Forestry&environment	1144		
Food processing	771		
Subsistency farming	364		
Rural diversification	1055	-968	251
Rural tourism	82		
Rural communities	82		
Total	793	1827	2620

The net programme effects calculated in the previous research on the employment at the national level amount to 793 AWU. Taking the displacement effects into account increases the net effects up to 2620 AWU.

IV. CONCLUSIONS

Use of „naïve" estimators in evaluation of programme effects on economic variables without matching can lead to the erroneous overestimation or underestimation of unplanned indirect effects on changes in employment attributed solely to the programme. Propensity score matching has to be considered a more suitable method in establishing a sound counterfactual.

The previously calculated direct programme impact on the changes in employment can not be considered as a correct result without the estimation of displacement effects that are unintended, and, in cases, can substantially change the final estimation.

The displacement effects can be either positive or negative. As the study shows, the indirect and unintended impacts of the programme support on non-participants can be rather substantial and comparable to direct intended impacts.

The size and sign of the impact varies depending upon the activities enclosed in the measures of programme axis. As Axis1 support mainly focuses on farm modernization, the employment would shift to non-supported farms in the close neighborhood of these supported farms. In turn, Axis3 support with the focus on diversification in the areas other than agriculture can provide job opportunities to neighboring farmers.

V. ACKNOWLEDGEMENTS

The authors would like to express their special gratitude and thanks to our colleagues from Farm Accountancy Data Network Department whose assistance in retrieving the data necessary for the research proved to be especially valuable.

REFERENCES

[1] S. Becker and A. Ichino, "Estimation of average treatment effects based on propensity scores," The Stata Journal. 2, pp. 358-377, 2002.

[2] L. Calmfors, "Active Labour Market Policy and Unemployment: A Framework for the Analysis of Crucial Design Features," OECD Economic Studies, 1994, 22, pp. 8-47.

[3] R. Dehejia and S. Wahba, "Propensity score-matching methods for nonexperimental causal studies," Review of Economics and Statistics, 84, pp. 151-161, DOI: 10.3386/w6829, 2002.

[4] J.J. Heckman, H. Ichimura, J. Smith, and P. Todd, "Characterizing selection bias using experimental data," Econometrica. 66(5), pp. 1017-1098. DOI: 10.2307/2999630, 1998.

[5] P.W. Holland, "Statistics and causal inference (with discussion)," Journal of the American Statistical Association, 81, pp. 945-970, DOI: 10.2307/2289064, 1986.

[6] G.W. Imbens, "Nonparametric estimation of average treatment effects under exogeneity: A review," The Review of Economics and Statistics. 86, pp. 4-29, DOI: 10.1162/003465304323023651, 2004.

[7] N. Mantel and W. Haenszel, "Statistical aspects of the analysis of data from retrospective studies of disease," Journal of the National Cancer Institute, 22(4), pp. 719-748, DOI: 10.1093/jnci/22.4.719, 1959.

[8] J. Michalek, "Quantitative tools for ex-post evaluation of EU RD programmes," in Advanced-Eval. Working Paper Series, University of Kiel, April 2007. http://www.advanced-eval.eu.

[9] PCM, "Ex-post evaluation of the SAPARD programme in the Slovak Republic. Final Report," in P.C.M. Group, December 2007.

[10] A. Pufahl, and Ch. Weiss, "Evaluating the effects of farm programs: results from propensity score matching," in Department of Economics Working Paper Series 113, Wien: Vienna University of Economics and Business Administration, 2007.

[11] P.R. Rosenbaum and D.B. Rubin, "The central role of propensity score in observational studies for causal effects," Biometrika, 70, 41- 55, DOI: 10.1093/biomet/70.1.41, 1983.

[12] P.R. Rosenbaum, *Observational Studies.* New York: Springer, 2nd edition, 2002.

[13] A. Roy, "Some thoughts on the distribution of earnings," Oxford Economic Papers, 3, pp. 135-146, 1951.

[14] D.B. Rubin, "Estimating causal effects of treatments in randomized and nonrandomized studies," Journal of Educational Psychology, 66, pp. 688-701. DOI: 10.1037/h0037350, 1974.

[15] J. Smith, "A Critical Survey of Empirical Methods for Evaluating Active Labor Market Policies," Schweiz, Zeitschrift für Volkswirtschaft und Statistik, 2000, 136, pp. 1-22.