 DOI: 10.22616/RRD.30.2024.020

IMPACT EVALUATION OF THE RURAL DEVELOPMENT PROGRAMME SUPPORT TO SMALL FARMS WITH REGRESSION DISCONTINUITY DESIGN

***Juris Hāzners , Armands Vēveris , Elita Benga**
Institute of Agricultural Resources and Economics, Latvia
*Corresponding author's e-mail: juris.hazners@arei.lv

Abstract
The viability of small farms and their participation in the market are demanding issues of the European agricultural policy. So it is important to appraise the most effective support measures to address these topics. The research objective is the evaluation of the net direct impact of the Latvian Rural Development programme (RDP) support to small farms on the farm economic indicators such as Balance, Utilized Agricultural Area, Intermediate Consumption, Labor and Productivity. To reach the research objective, Regression Discontinuity Design (RDD) method is applied. An outstanding advantage of the method is that the assessment can be carried out in conditions of limited data availability, when the larger number of farming indicators are unavailable and application of other methods is not appropriate. The results of the research prove that Program support contributes to the increase in all five indicators concerned. According to the specifics of the selected method, a single RDP sub-measure has been chosen for the study aimed directly at small farms. Applying the RDD method, a positive impact of support on indicators such as production subsidy and tax balance, agricultural land and productivity for a single beneficiary has been observed. A less pronounced positive impact on employment has been found. The impact on intermediate consumption expenses is small and positive. At the national level, the support has contributed to the substantial increase in the relevant indicators in the group of small-sized farms.
Key words: Rural Development Programme, Small Farms, Regression Discontinuity Design, Net Impact.

Introduction

The support and development of small farms recently has gained a momentum as a crucial issue within the Common Agricultural Policy (CAP). According to the European Commission, the problem of unequal income distribution among different size of farms is urgent (European Commission, 2021). So redistribution of support to smaller farms is considered essential. Therefore, the development of small farms is promoted through support programmes. Evaluating the effectiveness of support measures specifically tailored for small farms is challenging, as the availability of necessary data and information often is limited. Thus in the study the RDD method has been preferred for the assessment of the effectiveness of the support of small farms in Latvia Due to it's suitability in cases when data are scarce. The research objective is the evaluation of the net direct impact of the Latvian Rural Development programme support to small farms on the farm economic indicators such as Net Balance of Subsidies and Taxes, Utilized Agricultural Area (UAA), Intermediate Consumption, Labour and Productivity. The productivity is expressed as Net Turnover per Annual Working Unit (AWU). To reach the research objective, RDD method has been applied.

Materials and Methods

In this study the direct impact of sub-measure 6.3 (support for starting up a business by developing small farms) of Measure 6 'Farm and business development' of the Rural Development Programme 2014-2020 has been estimated. The sub-measure provides start-up aid of EUR 15 000 for farms with the net turnover or standard output (SO) in the year before receiving the support falling within the range from EUR 2 000 to EUR 15 000. As the unpublished data from Farm Accountancy Data Network (FADN) used for the study have the lowest threshold set at the EUR 4 000 EUR, only the upper eligibility limit was relevant. Total number of small farms for the extrapolation of the results at the farm size sub-sector level was obtained from Farm Structure Survey by National Statistics Office (CSP, 2020), where the lower threshold for the SO is set at the EUR 2 000. These lower threshold discrepancies would lead to certain selection bias. Moreover, the selection is based upon the Net Turnover which does not necessarily correspond to the respective SO. Hence, the extrapolated results have to be perceived with caution. The highest intensity of the implementation of the sub-measure occurred from 2016 to 2019, so the farms conforming to the upper threshold criteria in 2015 were selected. To avoid the selection bias due to the possible impact of support from other measures only 33 farms that received support exclusively in sub-measure 6.3 were included in the treatment group of the data panel. For controls, 113 eligible farms without any programme support were selected along with 79 farms with a Net Turnover falling within the EUR 15 000 - EUR 25 000 range limits.

The RDD first was applied in the evaluation of the USA national college student scholarship programs by Thistlewaite and Campbell (1960) by matching two groups of nearwinners in a competition on several background variables. After that the method was somewhat disregarded until it was enlivened by Goldberger (2008) in the analysis of compensatory school educational programs. The method compared to other quasi-experimental design methods has its advantages as additional pre-treament variables are not needed. Moreover, the values of the dependent variable are measured only once at a single point of time. Nevertheless, similarly to other methods data on untreated units are necessary which almost always prove to be a problem. Traditionally, support programmes require only data on participants. The method also could yield results with low statistical significance if there is no marked change in the dependent variable at the cutpoint. Usually analysis begins with an examination of the scatterplot of the outcome variable and rating variable. In most cases, the relationship between these variables is non-linear. With respect to populations two types of design can be distinguished (Battistin & Rettore, 2008). In the 'sharp' design all units on both sides of the cutoff either receive or do not receive their treatment, thus the treatment variable is binary. In the Type I 'fuzzy' design, there are units in the treatment group which do not receive treatment referred to as 'no-shows'. In the Type II 'fuzzy' design, there are both 'no-shows' and units in the control group which receive treatment referred to as 'crossovers'. In 'fuzzy' design treatment is assigned based upon the probability of receiving the support. Then the probabilities are calculated as the share of receivers within the treatment or control groups. As a rule, RDD analysis begins with an examination of a scatterplot with an outcome variable plotted on the vertical axis and the independent or rating variable plotted on the horizontal axis. The scatterplot shows whether there is a discontinuity in the outcome variable at the cutoff point. The observed discontinuity justifies further analysis. Bloom (2012) suggests two types of strategies for the correct specification of such functional form. Parametric or global strategy uses all observations in the sample. Nonparametric or local strategy uses only the observations that lies in the vicinity of the cutpoint (called a bandwidth). The rating variable can be centered on the cutpoint by including a new variable $x_i - c$ in the model. Then the most common approach to estimation using an RDD can be expressed with the equation:

$$y_i = \beta_0 + \sum_{j=1}^{m} \beta_{1j}(x_i - c)^j + \beta_2 t_i + \sum_{j=1}^{m} \beta_{3m}(x_i - c)^j t_i + \varepsilon_i \tag{1}$$

where β_0, β_1, β_2, β_3 – regression coefficients; $y_1, \ldots, y_n$ – vector of the dependent variable: $x_1, \ldots, x_n$ – vector of the independent variable; $t_1, \ldots, t_n$ – vector of the treatment variable; $\varepsilon_1, \ldots, \varepsilon_n$ – vector of residuals; c – cutoff point; n – number of observations; m – polynomial degrees.

The coefficient β_2 shows the average treatment effect on the treated (ATT). Usually, only linear and second-order polynomial models are applied. Gelman and Imbens (2019) think that higher order polynomial regressions are a poor choice in regression discontinuity analysis because imprecise estimates due to noise, sensitivity to the polynomial's degree, and inadequate coverage of confidence intervals. Huntington-Klein (2021) recommends the use of local regression to obtain 'smoothed' values of the dependent variable for the estimation with the linear or polynomial model. Locally weighted polynomial regression method (LOESS) was originally proposed by Cleveland (1979). The set of the independent variable is divided into subsets using a 'smoothing parameter' selected by the user, which shows the size of the subsets. For every value of the independent variable, a respective number of nearest neighbours are included in a subset. For each localized subset, weighted least squares regression (WOLS) introduced by Aitken (1935) is performed to find the coefficients for calculation of adjusted values of the dependent variables by simple regression. First, the distances from each point in a subset to the point of estimation are calculated. Then the distances are scaled by the maximum distance between all possible pairs of points in a subset. For calculation of the weights from scaled distances, most commonly Tukey's tri-cube weight function (Tukey, 1977) is used:

$$w(x) = \begin{cases} (1 - |x|^3)^3 for\ |x| < 1 \\ 0\ for\ |x| \geq 1 \end{cases}, \tag{2}$$

After the obtaining the weights, WOLS regression is performed by the matrix equation:

$$B = (X^T W X)^{-1} X^T W Y\ , \tag{3}$$

where X - matrix with the independent variable in second column and first column set to one; X^T - transposed matrix of X; W - square matrix with weights on the diagonal and other elements set to zero; Y - vector of the dependent variable; B - vector of regression coefficients. Then the vector of predicted Y values $\hat{Y}$ can be expressed as:

$$\hat{Y} = B^T X, \tag{4}$$

where X - matrix with the independent variable in second column and first column set to one; B^T – transposed matrix of B; $\hat{Y}$ – vector of predicted values of the dependent variable. After the obtaining the predicted 'smoothed' values of the dependent variable from the equations (3) and (4), regression is performed with the equation (1) to calculate the regression coefficients. If several models are employed, Akaike Information Criterion (AIC) (Akaike, 1974) is used. AIC aims to select the model which best explains the variance in the dependent variable with the fewest number of independent variables (parameters). So it helps select a simpler model (fewer parameters) over a complex model (more parameters). AIC measures the information lost, so the model with a lower AIC score indicates a better fit. AIC is calculated by formula:

$$AIC = Nlog\left(\frac{SS_e}{N}\right) + 2K, \tag{5}$$

where N - number of observations; K - number of parameters; SS_e - sum square of errors.

Results and Discussion

The scatterplot with the Balance after treatment as the dependent variable and net turnover before treatment as the independent variable is mapped on 'Figure 1'.

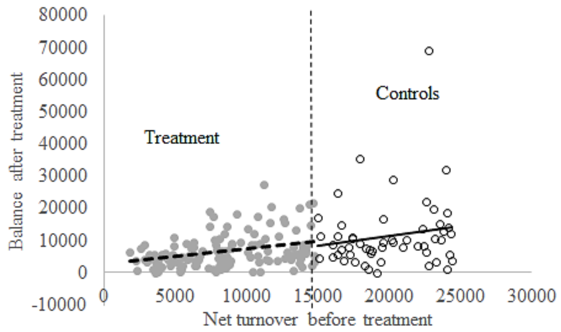

Figure 1. Balance after treatment and net turnover before treatment in treatment and control groups.

As there is a discontinuity in the outcome variable at the cutoff point (€15000 threshold), the research continues with the construction of several regression models. As some treatment group members do not receive treatment, Type I fuzzy design is chosen. Similarly with the Balance, discontinuity at the cutoff point can be seen for the other indicators - Productivity, Labour, UAA and Intermediate Consumption. The parametric or global strategy is selected using all observations for the modelling the outcome as a function of the rating and treatment variables. The study uses two specifications of the equation (1) - linear with interactions (m=1) and quadratic with interactions (m=2). Three smoothing parameters are used with values 0.5, 0.33 and 0.16. First, for all four indicators and every specification of the equation and value of the smoothing parameter, weights are calculated by formula (2), WOLS regression is performed by the equation (3) and predicted (smoothed) values of the dependent variable are calculated by the equation (4). Then the regression is performed with the equation (1) to calculate the regression coefficients. After that AIC criterion is calculated by formula (5). The statistically significant values of the ATT (coefficients β_2) from all 24 models along with respective AIC criterion values are shown in Table 1.

Table 1

Models with statistically significant ATT effects and respective AIC values

Indicator	Model specification	Smoothing parameter	ATT effect	AIC
Balance	Quadratic	0.5	2112 (5.09)***	874
Balance	Linear	0.5	6301 (13.04)***	968
Balance	Quadratic	0.33	3482 (4.24)***	987
Balance	Linear	0.33	9191 (12.75)***	1034
Balance	Linear	0.16	6161 (4.4)***	1144
Productivity	Quadratic	0.5	1415 (2.85)***	904
Productivity	Quadratic	0.33	4097 (6.08)***	954
Productivity	Linear	0.33	1518 (1.95)*	1047
Productivity	Quadratic	0.16	5502 (2.94)***	1123
Labour	Linear	0.5	0.09 (8.17)***	-798
Labour	Quadratic	0.33	0.11 (2.85)***	-651
Labour	Linear	0.33	0.24 (8.66)***	-644
Labour	Quadratic	0.16	0.61 (5.37)***	-479
Labour	Linear	0.16	0.48 (6.23)***	-476
UAA	Quadratic	0.5	7 (6.33)***	-99
UAA	Quadratic	0.33	11 (4.37)***	24
UAA	Linear	0.5	28 (13.82)***	66
UAA	Linear	0.33	41 (13.63)***	129
UAA	Linear	0.16	38 (7.93)***	207
Intermediate consumption	Quadratic	0.5	516 (2.33)**	770
Intermediate consumption	Linear	0.5	2674 (13.47)***	821
Intermediate consumption	Quadratic	0.33	5632 (9.16)***	939
Intermediate consumption	Linear	0.33	7023 (16.22)***	950
Intermediate consumption	Linear	0.16	3170 (1.74)*	1187
Intermediate consumption	Quadratic	0.16	5094 (1.81)*	1190

For the further assessment, values of the treatment effect with the lowest AIC criterion scores are retained. The ATE for four indicators are shown in Table 2.

According to the data of the paying agency, during the entire period of operation of the RDP 2014-2020, the support of the relevant measure has been provided to 3464 beneficiaries. The total impact for each indicator is calculated by multiplying the respective ATE effect by the total number of beneficiaries. Then the average indicators for the farms with the net turnover less than €15 thousand from the FADN data panel are calculated. By dividing the ATE effect with the respective indicator value from the FADN data panel, the shares of the impact in the values of the indicators are estimated at this farm size group level.

Table 2

Indicator values with the lowest AIC scores

Indicator	ATT effect
Balance	EUR 2 112
Productivity	EUR 1 415
Labour	0.09 AWU
UAA	7 ha
Intermediate consumption	EUR 516

The total number of farms with the net turnover less than €15 thousand nationwide is 12,270. Then the total national aggregate values for each indicator are calculated by multiplying the respective indicator value from the FADN data panel by the total number of small farms. Finally, shares of the impact in the aggregate indicator values are calculated dividing the total impact for each indicator by the respective total national aggregate values. The results of these estimations are shown in Table 3.

Table 3

Estimated aggregate impact of the support on the economic indicators

Indicator	Balance (EUR)	UAA (ha)	Intermediate consumption (EUR)	Labour (AWU)	Productivity (EUR/AWU)
ATE	2 112	7	516	0.1	1 415
Average for a single beneficiary	9 017	36	16 892	1.4	11 646
Share of the ATE impact	23%	20%	3%	6%	12%
Total impact on beneficiaries	7 314 398	25 087	1 787 387	307	1 415
Aggregate national value	110 638 590	442 898	207 264 840	17 080	10 081
Share of the total impact	7%	6%	1%	2%	14%

The most marked estimated impact on the small farms at the 23% level is on the Subsidy and Tax Balance and on the increase in the area of the Utilized Agricultural Area (UAA) (20%) and on the increase in Productivity (12%) which is calculated as Net Turnover divided by full-time employees expressed in Annual Working Units (AWU). Although to a lesser extent, a positive effect at the 6% was also seen on Labour input. This is an important finding, because it proves that the measure also promotes employment in the countryside, at the same time as raising productivity. Due to the expansion of production the value of Intermediate Consumption costs also has increased by 3%. The impact on the increase in intermediate consumption is relatively low if compared to other indicators. This indicates a relatively high efficiency. The shares of the total impact on all small farms nationwide are somewhat lower, except the impact on Productivity at 14%. The shares for the Subsidy and Tax Balance and UAA stand at 7% and 6%, respectively. The positive effect on employment is less than 2% of the total employment in this group of farms. The share for the Intermediate Consumption is the lowest. The efficiency of the support provided by the measure can be estimated by the conditional gains in the indicators against EUR 1 000 of the public funding. The aggregate impact on the indicators at the small farm level is divided by the total cost of the measure (public funding) at EUR 46.4 million. The results of these calculations are provided in Table 4.

There are few studies that evaluate the economic effect of support for small farms in Latvia. Some research has been done in AREI. According to survey data, investment support for small farms has had the most pronounced positive impact on the value of produced and sold products, as well as on long-term investments (Veveris & Puzulis, 2018). A more extensive study is published on AREI website (AREI, 2017).

Table 4

Gains form the € one thousand of support

Indicator	Gains
Balance	€ 158
UAA	0.54 ha
Intermediate consumption	€ 39
Labour	0.0066 AWU

In this study, various aspects of small farms have been evaluated, including socio-economic ones. The direct impact of supported investments made by small farms has also been assessed using the quasi-experimental methods based on counterfactual, such as Propensity Score Matching and Generalized Propensity Score Matching. A few two-country studies evaluate the impact of RDP on small farms both in Latvia and Lithuania. Four economic indicators were calculated in collaboration by Latvian and Lithuanian researchers: gross farm income per work unit; standard output; gross investment and subsidies on investment per ha UAA. It concludes that changes in trends of subsidies on investment have consequences in income trends of small farms in Latvia and Lithuania (Veveris *et al*., 2019). The Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) is a method employed for the multi-criteria analysis by Volkov *et al*. (2019). They have established that from

2013 onwards, changing the principles of the Common Agricultural Policy Direct Payment schemes and payment rates with more support allocated for small and young farmers, the values of the composite indicator for social sustainability has increased in Lithuania. Through the example of small farms from Poland, Romania and Lithuania, it has been shown that financial support of small-scale farms is regarded by the owners as an important element of sustainable development, both from an economic perspective and for the fulfilment of social and environmental functions (Hupkova *et al*., 2023). Another study that has been conducted in Poland, uses Cluster Analyses and Principal Component Analysis (PCA). This study also confirms the importance of support. It was concluded that the development of small farms is particularly influenced by external factors (EU funding; national benefits), rather than internal (entrepreneurial) factors (Hornowski *et al*., 2020). Staniszewski and Borychowski (2020) use FADN data on European regions studying the impact of subsidies on farm efficiency. Their research confirms that the impact of subsidies on efficiency depends on the size of farms. Statistically significant, stimulating effects of subsidies were identified only in the group of the largest farms. Such results put into question the effectiveness of the CAP in stimulating the development of the European Model of Agriculture. Kryszak *et al*. (2021) also use FADN data in the analysis of the impact of subsidies under the CAP. Albeit they recognize the necessity to ensure adequate profitability for small farms, at the same time considering the challenges that agriculture will face in the future associated with the climate change and growing demand for food, support mechanisms for the largest farms should be revised. The authors consider their long term viability crucial for the global competitiveness of European agriculture.

Conclusions

1. According to the calculations, RDP investment support in small farms has a significant positive impact on the production subsidy and tax balance, agricultural land areas and labor productivity.
2. Although to a lesser extent, the measure also promotes employment in the supported farms. This has also caused a slight rise in intermediate consumption.
3. The impact of support at the national level is estimated at around 1-6% of the total amount of relevant indicators in the small sized farm group.
4. The analysis of other studies confirms that the support given to small farms improves both their economic performance and social sustainability. However, differences between small and other farms still remain significant.
5. Therefore, it is possible to assess that this type of support is effective and it is useful to make it available to an even larger number of small farms, because currently only part of the farms in the relevant group can use it due to limited funding.
6. Further studies are required for the approbation of the method for other rural support measures in order to more detailed estimation of its usability in assessing the impact of support.

References

Aitken, A. C. (1935). II. On Fitting Polynomials to Data with Weighted and Correlated Errors. *Proceedings of the Royal Society of Edinburgh*, 54, 12-16.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.

AREI. (2017). *Mazo un vidējo saimniecību attīstības iespējas un ieteicamie risinājumi LAP kontekstā* (Development opportunities for small and medium-sized farms and recommended solutions in the context of RDP). Public report. Retrieved February 12, 2024, from https://www.arei.lv/lv/2017-gads.

Battistin, E. & Rettore, E. (2008). Ineligibles and eligible non-participants as a double comparison group in regression-discontinuity designs. *Journal of Econometrics*, 142(2), 715-730. DOI: 10.1016/j.jeconom.2007.05.006.

Bloom, H.S. (2012). Modern regression discontinuity analysis. *Journal of Research on Educational Effectiveness*, 5(1), 43-82. DOI: 10.1080/19345747.2011.578707.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American statistical association*, 74(368), 829-836.

CSP. (2020). *2020. gada lauksaimniecības skaitīšanas rezultātu kopsavilkums* (Census of Agriculture 2020). Retrieved January 5, 2024, from https://stat.gov.lv/lv/statistikas-temas/noz/lauksaimn/publikacijas-un-infografikas/12232-2020-gada-lauksaimniecibas.

European Commission. (2021). *Evaluation of the impact of the Common Agricultural Policy on territorial development of rural areas*.{SWD(2021) 398 final}. EC working Document, 118 p. Retrieved March 26, 2024, from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021SC0394.

Gelman, A. & Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3), 447-456. DOI: 10.1080/07350015.2017.1366909.

Goldberger, A. S. (2008). Selection bias in evaluating treatment effects: Some formal illustrations. *Modelling and Evaluating Treatment Effects in Econometrics*, 21, 1-31. DOI: 10.1080/07350015.2017.1366909.

Hornowski, A., Parzonko, A., Kotyza, P., Kondraszuk, T., Bórawski, P., & Smutka, L. (2020). Factors determining the development of small farms in central and eastern Poland. *Sustainability*, 12(12), 5095. DOI: 10.3390/su12125095.

Huntington-Klein, N. (2021). *The effect: An introduction to research design and causality*. CRC Press.

Hupkova, D., Smędzik-Ambroży, K., Stępień, S., Borychowsk, M., & Tosovic-Stevanovic, A. (2023). Is the Common Agricultural Policy tailored to the needs of farmers? Opinions of agricultural producers from Poland, Romania and Lithuania. *Journal of Central European Agriculture*, 24(1), 291-302. DOI: 10.5513/JCEA01/24.1.3756

Kryszak, Ł., Guth, M., & Czyżewski, B. (2021). Determinants of farm profitability in the EU regions. Does farm size matter? *Agricultural Economics – Czech*, 67, 2021 (3): 90–100. DOI: 10.17221/415/2020-AGRICECON

Staniszewski, J. & Borychowski, M. (2020). The impact of the subsidies on efficiency of different sized farms. Case study of the Common Agricultural Policy of the European Union. *Agricultural Economics – Czech*, 66, 2020 (8), 373–380. DOI: 10.17221/151/2020-AGRICECON.

Thistlethwaite, D. L. & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational psychology*, 51(6), 309.

Tukey, J. W. (1977). *Exploratory data analysis.* Boston, USA: Addison-Wesley Publishing Company.

Veveris, A. & Puzulis, A. (2018). Small Agricultural Farms in Latvia and Baltic Sea Countries and their Possibilities. *Economic Science for Rural development*, 47, 351-358. DOI 10.22616/ESRD.2018.041.

Veveris, A., Šapolaitė, V., Giedrė Raišienė, A., & Bilan, Y. (2019). How Rural Development Programmes Serve for Viability of Small Farms? Case of Latvia and Lithuania. *AGRIS on-line Papers in Economics and Informatics*, 11(2), 103-113. DOI: 10.7160/aol.2019.110210.

Volkov, A., Balezentis, T., Morkunas, M., & Streimikiene, D. (2019). Who Benefits from CAP? The Way the Direct Payments System Impacts Socioeconomic Sustainability of Small Farms. *Sustainability* 2019, *11*(7), 2112. DOI: 10.3390/su11072112.