Impact of EU subsidies on the of rural areas in Hungary

Fertő, Imre[1]; Bakucs, Zoltán[1]; Ágnes Varga[2]

1. Research Centre for Economic and Regional Studies, Institute of Economics, HAS, H-1112 Budapest, Budaörsi út 45, Kaposvar University, Corvinus University of Budapest, Hungary
2. Eötvös Loránd University

Correspondence: ferto.imre@krtk.mta.hu

Paper prepared for presentation at the 160th EAAE Seminar 'Rural Jobs and the CAP', Warsaw, Poland, December 1-2, 2016

160 EAAE Seminar
Rural Jobs and the CAP
Warsaw (Poland) December 1-2, 2016 The Staszic Palace

Impact of EU subsidies on the of rural areas in Hungary

Fertő, Imre[1*] ; Bakucs, Zoltán[2]; Ágnes Varga[3]

1. Research Centre for Economic and Regional Studies, Institute of Economics, HAS, H-1112 Budapest, Budaörsi út 45, Kaposvar University, Corvinus University of Budapest, Hungary, ferto.imre@krtk.mta.hu *
2. Research Centre for Economic and Regional Studies, Institute of Economics, HAS, H-1112 Budapest, Budaörsi út 45, Corvinus University of Budapest, Hungary, bakucs.zoltan@krtk.mta.hu
3. Eötvös Loránd University; varga.agi14@gmail.com

Abstract:

In recent years, we experienced a growing interest in the evaluation of EU co-funded programmes. The paper is a first attempt to analyse the impacts of such support on the wellbeing of Hungarian rural areas between 2002 and 2008, employing a two stages approach. In the first step, we construct a multi-dimensional RDI (Rural Development Index) measuring the overall level of regional development and quality of life in Hungarian small regions. In the second step we apply propensity score matching approach to evaluate the impact of the regional subsidies on the RDI. Estimations reveal four main findings. First, calculations suggest that concentration in the EU support grows with increasing amount of subsidies. Second, the convergence of support can be also observed. Third, we find considerable mobility in terms of the level of subsidies during analysed period. This indicates that there has been chance for poorly subsidised regions to improve their relative position and vice versa. Finally, our results imply that it is very difficult to identify any impacts of subsidies, because estimations are highly sensitive to the chosen indicators. The size of identified effects is rather small and its direction may equally be positive or negative. However, we can conclude that irrespective to the sign of estimated coefficients the size of impact of regional subsidies is

negligible. Consequently, further research is needed to explore impacts mechanisms of subsidies.

Key words: EU supports, impact evaluation, sub-region

1. Introduction

It is difficult to overestimate the role of Rural Development Policies (RDPs) in developed economies. 75 percent of the OECD countries' territory is classified as rural, and on average a quarter of the total population lives in these areas (OECD, 2006). In the past decades the global economy experienced an unprecedented growth of agricultural productivity – itself a laudable process, yet despite the lavish subsidies, leading to a fall in both agricultural employment and the weight of agriculture in national economies (at least when developed economies are considered). Whilst the agricultural output amounts to roughly 2 percent of OECD nations' GDP, the vast majority of rural land use is for agricultural purposes (e.g. 96 percent in the EU25, including forests). However, in the EU25 only 13 percent of rural labour is employed in agriculture (the OECD average is 10 percent producing a gross value added of only 6 percent even if only the output of rural areas is considered (OECD, 2006). Whilst the aims of EU Common Agricultural Policy (CAP) with respect to agricultural production were laid down in the 1958 Rome Treaty, and albeit with significant amendments, but it is applied up to present, the importance of rural development not directly connected to production was only recognized in the 70's. Thus the modern CAP, as developed in AGENDA 2000 shifted the support system towards an integrated rural development policy, creating the European Agricultural Model (Renting et al., 2009) with its primary aim to promote a viable and liveable rural environment rather than maximizing agricultural output (for further discussion see for example 'The new rural paradigm: policies and governance', OECD, 2006). It was a key revelation that besides production, a nation's agriculture contributes to the creation or preservation of a number of important values such as landscape, traditions-costumes, social structures and none-the-less environment protection. The most important pre-condition of the creation/preservation of the abovementioned values is the existence of sufficient active rural population. This highlights the importance of policies aimed to slow rural to urban migration, and reverse the constant increase of average rural inhabitants' age. The economic output of Hungarian rural areas is 50% less the national average and 3 times less than the predominantly urban output. For more details with respect to sectoral and regional differences in the EU and OECD countries see for example Bollman et al. (2005), Copus et al. (2006), or

Terluin et al. (2011). To sum up, besides economic and agricultural perspective, rural areas are also very important in terms of population, preserving the landscape tradition and non-the-less environment. In addition, NMS are more rural than OMS, and the income gap between rural and urban areas are more predominant in NMS than OMS. Consequently, the analysis of RDP is perhaps an even more relevant issue in these countries.

Despite its importance, the empirical literature with respect to the evaluation of rural development measures is rather poor. Most papers focus on the impact of agricultural policy on labour market or rural income distribution (e.g. Breustedt and Glauben, 2007; Elek et al., 2010; Esposti, 2007; Petrick and Zier, 2012; Pufahl and Weiss, 2009; Swinnen and Van Herck, 2010). A possible reason for the scarcity of relevant literature is that the policy evaluation or impact assessment of RDP is a rather complicated issue since complex notions are hard to quantify, whilst all relevant aspects of the impact should be included in a transparent and easy to handle fashion (from data point of view). There are two key issues here: first the problem of applying partial indicators (such as number of projects supported, area supported, change in employment, value of realized investments, and GDP change – see Michalek and Zarnekow 2012 for a critical review), and second, the issue of counterfactual situation , excluding the possibility of before – after comparison. Often employed naïve approaches for the impact evaluation of RDP such as simple case studies or partial indicators do not even attempt to create a counterfactual situation (Terluin and Roza, 2010). Generally, the most important drawback of partial measures is the lack of clear causality relations between partial measures and RDP (the problem to make distinction between impact of RDP and other exogenous factors). These issues may however be solved by the use of a complex Rural Development Indicator, RDI, originally proposed by Michalek and Zarnekow (2012) and counterfactual analysis. Contrary to Michalek (2012) who investigates only the impact of the SAPARD programmes in Poland and in Slovakia between 2002 and 2005 we focus on the period (2002-2008) covering all rural development policy measures. Thus we can assess the effects of the EU rural development policy in Hungary. The aim of the paper is to (1) apply the indicator for the 2 period, by providing a quantitative, ready-to-use tool for monitoring, impact assessment and agenda setting, and (2) to assess in a complex way the impact of RDP upon Hungarian rural regions by providing an overall picture of its effect upon the actual development of rural areas.

2. Methodology

The basic idea is simple: people do move (migrate) where their quality of life is better, thus by making a decision they implicitly weight the importance of regional characteristics that define the local 'quality of life'. These characteristics and weights will then be used to derive the RDI indicator.

More specifically, the empirical methodology consists of the following steps:

1. We summarize the data available for 3,164 administratively independent settlements into 174 small regions (a much deeper perspective than the 20 regions available under the NUTS-3 nomenclature), the subject of our analysis. Further, we employ principal component (PCA) and factor analysis to reduce to around 3-5 the number of available variables (around 130 depending on the time span employed). There are various procedures available for this task (see e.g. Afifi et al. 2004 as a practitioner's handbook), the natural way would be testing the data for the suitability of PCA (e.g. Kaiser-Meyer-Olkin measure and Bartlett's test of variable's independence), followed by rotation algorithms (e.g.Varimax), and finally, the use of a factor selection criteria (e.g. Kaiser considering only factors with Eigen values larger than 1).

2. Estimation of the migration function in order to derive the weights (βk in eq.1) needed for the complex RDI indicator. The baseline model of the migration function is:

$mp_{it}= \alpha_0+ \beta_k F_{ikt}+v_i+\varepsilon_{it}$, where (1)

α_0 – is the constant,

mp_{it} – net migration into region i, normalised by the total population of the region i,

F_{ikt} – value of factor k in region i, at time t – originating from step 1.,

ε_{it} – region specific residual, with the usual white noise properties.

Given the panel nature of data, and the strict underlying assumptions of panel models, a variety of models will be estimated using specification and diagnostic tests in order to select the 'best' model (see e.g. the handbook of Baltagi, 2008).

3. We may now estimate the RDI index takes the following form:

$RDI_i=h(\beta_k,Z_k^i)=\sum_k \beta_k * Z_k^i$, where (2)

RDI_i – Rural Development Index in region i,

Z_k^i – i region's k measurable characteristics,

β_k – weights for each k characteristic, specific for region mind i, and time t resulting from the estimation of the migration function (1).

Thus the RDI is a complex indicator based on regional characteristics of Z_k^i, weighted by the estimated coefficients of the migration function, β_k. Weights represent the 'relative social value' of regional characteristics Z_k^i which are heuristically used by those making a decision to stay or move from the region as measures for 'quality of life'. The estimation of regional characteristics is done by factor analysis techniques using all relevant variables (see data section for further details) available to describe the given region's social economic and environmental aspects.

4. Once the unbiased RDI is calculated, we are in position to actually analyse the impact of RDP's on sub-regions. Whilst in standard policy analysis settings, the sample-average treatment effects cannot be calculated because we only observe one of the two possible outcomes for each individual (or sub-region in our case), this issue is solved by the RDI allowing the creation of the counterfactual. Following the insights of impact analysis literature we can thus adopt the counterfactual framework developed by Rosenbaum and Rubin (1983). More specifically, sub-regions are selected into treatment and non-treatment groups that have similar potential outcomes (RDI scores).

4a. To solve the evaluator's classing problems the matching approach reproduces the treatment group among the non-treated by pairing each program participant with members of the non-treated group, controlling for observable characteristics. Estimating the treatment effects based on the Propensity Score Matching (PSM) requires two assumptions. The first is the Conditional Independence Assumption (CIA), which states that for a given set of covariates participation is independent of potential outcomes. A second condition is that the average treatment effect for the treated (ATT) is only defined within the region of common support. This assumption ensures that treatment observations have comparison observations "nearby" in the propensity score distribution. For more comprehensive discussion of the econometric theory behind this methodology we refer the reader to Imbens and Wooldridge (2009) and Guo and Fraser (2010). However, the PSM has several limitations. First, PSM requires extensive data sets on large samples of units, and even when those are available, a lack of common support between the treatment or enrolled group and the pool of nonparticipants may appear. Second, the assumption that no selection bias has occurred arising from unobserved characteristics is very strong, and most problematic, further, it cannot be tested.

We employ Propensity Score Matching (PSM) on the basis of observed covariates for both RDS and non-RDS. The method balances the observed covariates between the RDS group and non RDS sub-regions based on similarity of their predicted probabilities of being RDS sub-regions. The aim of PSM matching is to find a comparison group of RDS sub-regions from a sample of non-RDS sub-regions that is closest (in terms of observed characteristics) to the sample of RDS sub-regions.

4b. Having data on RDS and non RDS sub-regions over time can also help in accounting for some unobserved selection bias, by combining PSM and Difference-in-Differences estimator (conditional DID estimator). The conditional DID estimator (e.g. Smith and Todd, 2005) is highly applicable in case the outcome data on programme participants (i.e. RDS sub-regions) and nonparticipants (non-RDS sub-regions) is available both “before” and “after” periods (2007 and 2013, respectively). In our proposed study, the PSM-DID measures the impact of the RDS by using the differences in selected outcome indicator (ATE, or ATT) between RDS (D=1) and non RDS (D=0) in the before-after situations. The main advantage of the PSM-DID estimator is that it can relax the *unconfoundedness* assumption. The PSM-DID estimator also allows for quantile differences, that is assessing the effects of RDS at different points of the outcome variable’s (RDI scores) distributions. It means that we can compare individuals across both groups and time according to their quantile.

In the actual estimation propensity score matching may be used (e.g. *psmatch2*, available in STATA econometric package, see Leuven and Sianesi, 2003; Sianesi, 2004), Differences – in –Differences estimator, DIFF in STATA (Villa, 2011) or matching estimators for average treatment effects *teffects* match in STATA (Abadie et al., 2004; Caliendo and Kopeinig, 2008).

3. Data

In the first stage, applying approach by Michalek and Zarnekow (2012) we construct a multi-dimensional index measuring the overall level of regional development and quality of life in individual regions of Hungary. In the Regional Development Index (RDI), the development domains are represented by 132 partial socio- economic, environmental, infrastructural and administrative indicators/variables at NUTS4 level. The weights of these economic, social and environmental domains are derived empirically from an econometrically estimated, interregional migration function after selecting the “best” model from various alternative model specifications. The RDI was empirically applied to the regional development in individual rural areas of Hungary in the years 2002–2008. Due to its comprehensiveness, RDI is suitable for analysing the overall level of development of rural areas and also for evaluating the impacts of various structural programmes at a regional level. In standard policy analysis settings, the sample-average treatment effects cannot be calculated because we only observe one of the two possible outcomes for each individual (or sub-region in our case). Thus in second stage we employ a matching estimation technique to identify the treatment effects. Following the insights of impact analysis literature we adopt a counterfactual framework developed by Rosenbaum and Rubin (1983). More specifically, regions selected into treatment and non-treatment groups have potential outcomes (RDI scores).

Data for the RDI calculations based on Central Statistical Office regional database provided by Databank of Centre for Economic and Regional Studies of Hungarian Academy of Sciences. WE employ 132 variables covering various fields of quality of life including demographics (15 variables), health services (9), business units (2), tourism and catering (9), retail sector (24) transport (7), community infrastructure (14), environment (4), culture (2), unemployment (4), education (16), social protection (17) personal income tax (3), number of houses (5), number of villages (1). In order to provide more comprehensiveness of dimensions of well-being we cannot take into account unequal number of indicators per dimensions. Data for the EU funding are based on Information Systems of National Regional Development. We use both value data of EU funds and number of projects funded by the EU.

4. Results

We present our results in two main steps. First, we provide an overview on the development of subsidies with special emphasis on their stability and dynamics. Second, we focus on the impacts of subsidies on rural well-being.

4.1. Development of regional subsidies

The descriptive statistics of the total (years 2002-2008) development subsidies, presented in Table 1. emphasise an uneven distribution of funds. The average value of support per sub-region amounts to HUF 2,2 billion, but there are sub-regions with no support at all (minimum value 0) whilst the maximum value of support per project was HUF 541 million. The uneven distribution is also reflected by the extremely high standard deviation. The picture is nuanced by the last two rows of Table 1. (per capita and per square km subsidy) where the inequality of distribution is less prominent.

Table 1. Descriptive statistics of subsidies

	N	mean	SD	Minimum	Maximum
support (mil. HUF)	1218	2253	18021	0	505647
project	1218	88	171	2	3686
support/project (mil. HUF)	1218	23	40	0	541
support/capita (thousands. HUF)	1218	29	39	0	661
support/km^2 (mil. HUF)	1218	4	34	0	963

Source: Own calculations

Table 2. presents the yearly averages support variables. Note the post EU accession (2004) non-monotonic increase of the average development funds. Somewhat surprisingly, the number of projects supported continuously decreases after 2004, resulting in a dynamic expansion of per subsidy per project averages. The support/km^2 increased five folds, whilst the support per capita roughly doubled between start and end period (an otherwise expected outcome – i.e. the distribution of funds is more likely to follow the sub-regions total population rather than area surface).

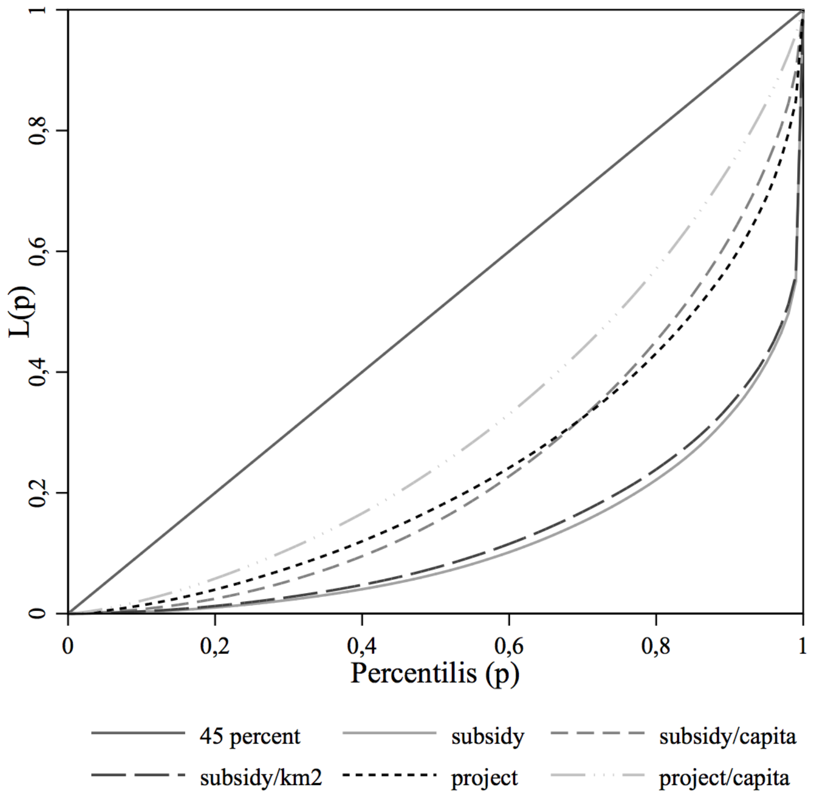
Table 2. Average values of subsidies and supported projects per sub-regions between 2002-2008 (million HUF and no.)

	support	No. of project	support/project	support/capita (thousands)	support/km^2
2002	1028	135	8	23	2.0
2003	997	110	9	22	1.8
2004	116	134	9	19	2.2
2005	2852	91	31	51	5.4
2006	177	58	30	30	3.4
2007	2328	38	61	7	4.4
2008	5668	46	124	49	10.6
total	15769	613	272	201	29.8

Source: Own calculations

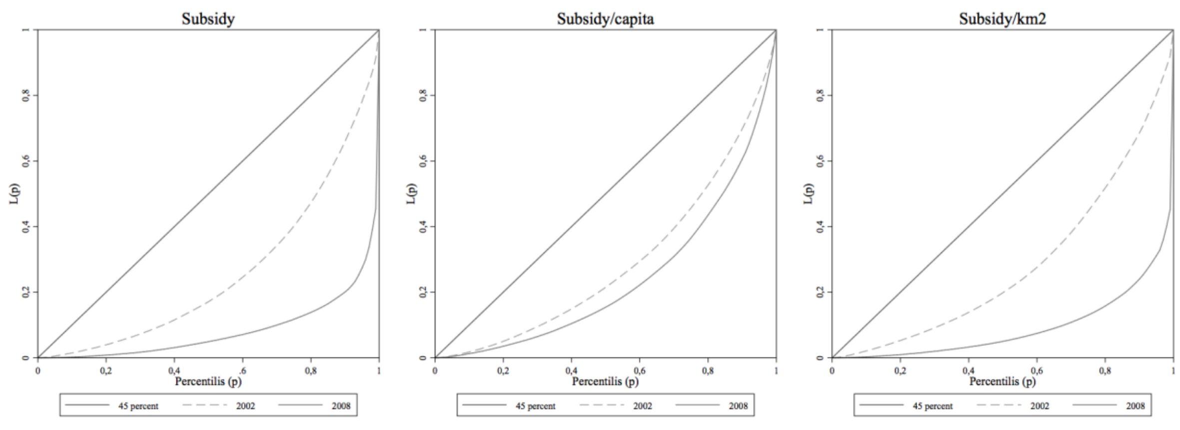
The Lorenz curves (Figures 1. and 2.) reinforce our prior beliefs with respect to increasing subsidy concentration. Figure 1. shows that the concentration of all subsidy indicators increased. The most prominent increase is recorded for total subsidies received and for the per square km support indicators, whilst the lowest for the per capita support. Figure 2. depicts the evolution of concentration for three support indicators between 2002 and 2008. The higher concentration ratio in 2008 is evident from the graph.

Figure 1. Lorenz curves of the sub-regional distribution of subsidies and project numbers



Source: Own calculations

Figure 2. Lorenz curves of the sub-regional distribution of subsidies in 2002 and 2008



Source: Own calculations

4.1.1. Stability of regional development subsidies

There are a number of tools available to assess the stability of regional development subsidies. Thus, at least two different types of stability should be distinguished. On one hand, the stability of subsidy distribution between periods, and on the other, the stability of the amount of support in a given sub-period between years. A natural question originating from the first is whether the subsidy between sub-regions is converging or diverging. The methodology of economic data convergence analysis evolved in the past decades from simple Galtonian regressions to panel unit root tests. Originally developed for the convergence analysis of economic development, the latter method has been widely applied to study inflation convergence (e.g. Lopez, Papell 2012) or even the assessment of trade specialisation (e.g Fertő 2006). In a bi-variate time series setting, the economic development of two countries converges (on long run), if the per capita GDP differentials are stationary. In a similar fashion, we may apply unit root techniques for the analysis of regional subsidy convergence. Moreover, with a large number of regions and a 7 years time span, we use robust panel unit root tests. Panel econometrics experienced significant advances in the past decade (see e.g. Baltagi 2008 for a detailed discussion). The literature recognises first and second generation unit root tests, depending whether individual or common unit root processes are assumed. Since it is difficult (or even impossible) to choose a 'best' approach, we employ a battery of unit root tests assuming both individual intercept (Table 3.) and individual intercept and trend (Table 4.) as deterministic specification.

Table 3. Panel unit root tests of development subsidies (individual intercept)

Method	Statistic	Prob.**
Null: unit root (assuming common unit root process)		
Levin, Lin & Chu t*	-30.128	0.0000
Null: unit root (assuming individual unit root process)		
Im, Pesaran and Shin W-stat	-9.071	0.0000
ADF - Fisher Chi-square	678.030	0.0000
PP - Fisher square	885.139	0.0000

Source: Own calculations

Without a trend (Table 3.) the results are clear: all tests soundly reject the unit root null in favour of stationarity alternative hypothesis. This suggests long-run convergence of

Hungarian sub-regions. Somewhat more mixed results are obtained if a trend is also considered (Table 4.), here the Im, Pesaran and Shin W does not reject the null, but the other 3 tests do.

Table 4. Panel unit root tests of development subsidies (trend and individual intercept)

Method	Statistic	Prob.**
Null: unit root (assuming common unit root process)		
Levin, Lin & Chu t*	-30.969	0.0000
Breitung t-stat	-1.3688	0.0855
Null: unit root (assuming individual unit root process)		
Im, Pesaran and Shin W-stat	-0.7401	0.2296
ADF - Fisher Chi-square	410.699	0.0115
PP - Fisher Chi-square	709.659	0.0000

Source: Own calculations

Since our results are robust when other derived variables (per capita and per square km subsidies – not included here, but available upon request) we may cautiously conclude that regional development subsidies are converging in Hungary.

4.1.2. Dynamics of regional development subsidies

In order to complement the previous section, we take a closer look on the changes within the distribution of regional subsidy amounts. We first arrange the amounts of subsidy received into four quartiles, than estimate a Markov type transition matrix which reveals the probabilities of sub-regions moving between quartiles. Thus the probability of a sub-region remaining in the same quartile two consecutive ears ranges between 24.1 and 39. percent (on the diagonal). It is worth noting, the highest diagonal probability is displayed by sub-regions in the fourth quartile (i.e. the biggest beneficiaries of development funds). The probability to worsen a position shows a decreasing trend towards the first quartile. Contrary, the probability to improve from the first and second quartiles to the fourth one is below 20%. In the second and third quartile the probability to slip down a position is relatively high, 34 and 53 (24+29) %. Most importantly, Table 5. emphasises that beside convergence, there is significant mobility in the relative position of sub-regions in accessing development funds. Mobility is working both ways: a sub-region receiving higher subsidies in 2002 may fall amongst the regions receiving relative less support whilst an opposite scenario is equally likely. The analysis of mobility is especially important in the light of post accession increased amount of available funds.

Table 5. The Markov transition matrix of sub-regions by quartiles, years 2002-2008

	1. quartile	2. quartile	3. quartile	4. quartile
1. quartile	0.2820	0.2459	0.2820	0.1902
2. quartile	0.3410	0.2492	0.2328	0.1770
3. quartile	0.2376	0.2871	0.2409	0.2343
4. quartile	0.1447	0.2204	0.2401	0.3947

Source: Own calculations

4.2. Impact analysis of regional development subsidies

In line with the current literature, we analyse the impact of regional development subsidies by propensity score matching[1] (for a detailed discussion of the methodology see Guo, Fraser, 2010). The estimated propensity score is actually the probability of participation in a program (treatment), conditioned on control variables calculated for all sub-regions. A number of matching algorithms are available such as nearest neighbour, radius caliper, stratification matching and kernel matching (Abadie et al. 2004, Leuven, Sianesi 2009). Whilst asymptotically all matching procedures should result similar conclusions, small sample estimation may pose some problems. The following criteria were used to choose the appropriate matching algorithm: a) standardised bias, b) t-test and c) common significance and pseudo R^2.

Those sub-regions where the programme intensity was higher than 2/3 of the median were qualified as 'subsidised' for each indicator (i.e. subsidy per region; per capita; and per km^2). As before, we use all three subsidy indicators: total subsidy per sub-region, subsidy per capita and per square km. in a first step, a logit model (eq. 3) is estimated for all three subsidy indicators (thus the dependent variable changes).

$$Subsidy_{it}=\alpha_0+\alpha_1 RDI2002_{it}+\alpha_2 UNEMP2002_{it}+\alpha_3 UNEMP_{it}+v_i+\varepsilon_{it} \quad (3)$$

where Subsidy is dummy variable takes value one if a sub-region is identified as a subsidised one, and zero otherwise. $RDI2002_{it}$ is the 2002 level of rural development index and $UNEMP2002_{it}$ is the 2002 absolute value of unemployment - these variables control for the initial status of a given sub-region. In addition, the variable UNEMP captures the current

[1] We use psmatch2 STATA routin for the estimation.

level of unemployment in the sub-region. The results of the logit estimations are used to calculate the probability of participation (of being treated) of a given sub-region in the development projects. As discussed before, PSM methodology requires careful balancing of covariates, Tables 6 – 8 present 3 block of test results, of various matching procedures. Results emphasise the correct matching approach was used (e.g. where the mean values of covariates were significantly different in the unmatched sample, after matching the null of mean equality across treated and untreated sub-regions may generally not be rejected.

Table 6. Balancing tests of subsidies (common support: sub-region, biweight kernel) in subsidised and not subsidised sub-regions

		mean		% decrease		t-test	
Variable	Sample	treated	control	% bias	bias	t	p>t
RDI2002	unmatched	9.7e-05	0.0002	-4.9		-0.83	0.406
	matched	0.0001	0.0001	0.2	8.96	0.04	0.967
UNEMP2002	unmatched	363.83	154.8	22.3		3.49	0.000
	matched	175.9	147.03	3.1	86.2	1.01	0.311
$UNEMP_{it}$	unmatched	0.0066	0.0057	5.0		0.83	0.407
	matched	0.0045	0.0035	5.4	-9.0	1.32	0.186

Source: Own calculation

Table 7. Balncing tests of subsidies per capita (common support: sub-region, biweight kernel) in subsidised and not subsidised sub-regions

		mean		% decrease		t-test	
Variable	Sample	treated	control	% bias	bias	t	p>t
RDI2002	unmatched	-6.7e-05	0.0005	-23.3		-4.21	0.000
	matched	-4.9e-06	-3.3e-05	1.2	95.0	0.36	0.721
UNEMP2002	unmatched	296.48	260.12	3.5		0.60	0.550
	matched	274	262	1.1	67.5	0.25	0.804
$UNEMP_{it}$	unmatched	0.0073	0.0045	16.4		2.63	0.009
	matched	0.0062	0.0057	3.0	81.6	0.59	0.555

Source: Own calculation

Table 8. Similarity tests of subsidies per square kilometre (common support: sub-region, biweight kernel) in subsidised and not subsidised sub-regions

		mean		% decrease		t-test	
Variable	Sample	treated	control	% bias	bias	t	p>t
RDI2002	unmatched	0,0002	4,8e-05	6,7		1,9	0,276
	matched	0,0001	5,0e-05	4,3	35,9	1,1	0,311
UNEMP2002	unmatched	332,51	201,29	8,13		2,17	0,030
	matched	175,84	153,26	2,4	82,8	0,79	0,430
$UNEMP_{it}$	unmatched	0,0064	0,0061	1,7		0,29	0,770
	matched	0,0048	0,0046	1,0	45,3	0,22	0,829

Source: Own calculation

An important requisite of PSM methodology is to assessment whether the *common support* or *overlap* assumptions do hold (Caliendo, Kopeining, 2005). The test is based on the comparison of the distribution of estimated propensity scores in the treated and untreated samples. This may be done using graphical approaches (kernel density functions or histograms) or by applying parametric/non-parametric statistical tests. The result of Smirnov-Kolmogorov tests result suggest we may not reject the equal distribution of the two groups null hypothesis at 1% significance level.

Finally, and perhaps most importantly, we assess the ATT (Average Effect of Treatment on Treated) impact of development subsidies on sub-regions using two approaches (see Abadie et al. 2004 for a discussion of pros and cons). First a non-parametric Kernel matching (using bootstrapped z values) and second, nearest neighbour matching – allowing bias adjustment and heteroscedasticity robust variance estimation – is employed[2].

Table 9 presents our main results obtained with the abovementioned approaches. We reach the same – quite unfortunate – conclusion of extremely low, close to zero impact of subsidies on the sub-regions. The overall subsidy, and the per square km subsidy received seems to have a small positive impact (yet for the former this is significant only when bootstrap methods and 10% significance level is used. The per square km subsidy is significantly (small) positive with both methods. Contrary, when the per capita subsidy indicator is used, we obtain negative effects, regardless of estimation procedure.

[2] We apply STATA nnmatch program developed by Abadie et al 2004.

Table 9. Impact (ATT) of development subsidies

ATT	Coef.	SD	z	P>z
Subsidy*	0.0005	0.0003	1.67	0.095
Subsidy per capita	-0.0015	0.0003	-4.63	0.000
Subsidy per km^2*	0.00012	0.0003	3.69	0.000
SATT				
Subsidy	0.0004	0.0003	1.49	0.137
Subsidy per capita	-0.0013	0.0003	-4.07	0.000
Subsidy per km^2	0.0001	0.0003	3.52	0.000

Source: Own calculations; Note: *bootstraped z statistic (200 replications)

5. Conclusions

Estimations reveal four main findings. First, calculations suggest that EU subsidies concentrate where there have been previous EU subsidies. Second, some convergence of support can also be observed. Third, we find considerable variation in terms of the level of subsidies during the period analysed. This indicates that there has been a chance for poorly subsidised regions to improve their relative position or weaken their position further. Finally, our results imply that it is very difficult to identify any impacts of subsidies, because estimations are highly sensitive on the chosen parameters. The significance of identified effects is rather low and its direction can be both positive and negative. We conclude that, irrespective of estimated coefficients, the impact of regional subsidies is negligible. As a consequence, further research is needed to explore impacts mechanisms of subsidies.

References

Abadie, A., Drukker, D., Herr, J.L., Imbens, G.W. (2004) Implementing matching estimators for average treatment effects in Stata. *Stata Journal* 4, 290-311.

Afifi, A., Clark, V.A., May, S. (2004) Computer-Aided Multivariate Analysis, 4th Edition, Chapman & Hall/CRC.

Athey, S., Imbens, G. (2006) Identification and inference in nonlinear difference-in-differences models. *Econometrica* 74, 431-497.

Baltagi, B. H. (2008) Econometric Analysis of Panel Data. 4th edition. Wiley. New York.

Bollman, R., Terluin, I., Godeschalk, F., Post, J., (2005) Comparative analysis of leading and lagging rural regions in OECD countries in the 1980s and 1990s, 45th Congress of the European Regional Science Association (ERSA), Vrije Universiteit Amsterdam, pp. 23-27.

Breustedt, G., Glauben, T. (2007) Driving forces behind exiting from farming in Western Europe. *Journal of Agricultural Economics* 58, 115-127.

Caliendo, M., Kopeinig, S. (2008) Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys* 22, 31-72.

Copus, A., Hall, C., Barnes, A., Graham, D., Cook, P., Weingarten, P., Baum, S., Stange, H., Linder, C., Hill, A. (2006) Study on employment in rural areas (SERA). European Commission, DG Agri.

Elek, S., Fertő, I., Forgács, C. (2010) The possible effects of the CAP Reform on farm employment in Hungary. *Agricultural Economics Review* 11, 29-34.

Esposti, R. (2007) Regional growth and policies in the European Union: Does the Common Agricultural Policy have a counter-treatment effect? *American Journal of Agricultural Economics* 89, 116-134.

Guo, S., Fraser, M.W. (2010) Propensity score analysis: Statistical methods and applications. Sage Publications.

Imbens, G.M., Wooldridge, J.M. (2009) Recent developments in the econometrics of program evaluation. *Journal of Economic Literature* 47, 5-86.

Leuven, E., Sianesi, B. (2003) PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. Statistical Software Components.

Maddala, G.S., Kim, I.M. (1998) Unit roots, Cointegration and Structural Change. Cambridge University Press, Cambridge.

Michalek, J., (2012) Counterfactual impact evaluation of EU rural development programmes-Propensity Score Matching methodology applied to selected EU Member States. Volume 2: A regional approach. Institute for Prospective and Technological Studies, Joint Research Centre.

Michalek, J., Zarnekow, N. (2012) Application of the rural development index to analysis of rural regions in Poland and Slovakia. *Social indicators research* 105, 1-37.

OECD (2006) The new rural paradigm: Policies and Governance. Organisation for Economic Co-operation and Development, Paris.

Petrick, M., Zier, P. (2012) Common Agricultural Policy effects on dynamic labour use in agriculture. *Food policy* 37, 671-678.

Pufahl, A., Weiss, C.R. (2009) Evaluating the effects of farm programmes: results from propensity score matching. *European Review of Agricultural Economics* 36, 79-101.

Renting, H., Rossing, W., Groot, J., Van der Ploeg, J., Laurent, C., Perraud, D., Stobbelaar, D.J., Van Ittersum, M. (2009) Exploring multifunctional agriculture. A review of conceptual approaches and prospects for an integrative transitional framework. *Journal of Environmental management* 90, S112-S123.

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

Sianesi, B. (2004) An evaluation of the Swedish system of active labor market programs in the 1990s. *Review of Economics and Statistics* 86, 133-155.

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

Swinnen, J., Van Herck, K. (2010) Compensation payments in EU Agriculture. Trade Adjustment Costs in Developing Countries: Impacts, Determinants and Policy Responses, London: Centre For Economic Policy Research and Washington, DC: The World Bank, 361-381.

Terluin, I., Strijker, D., Munch, W., (2011) Economic dynamics in rural regions, in: Oskam, A., Meester, G., Silvis, H. (Eds.), EU policy for agriculture, food and rural areas. Wageningen Academic Publishers, Wageningen, pp. 317-332.

Terluin, I.J., Roza, P. (2010) Evaluation methods for rural development policy. LEI Wageningen UR.

Tocco, B., Davidova, S., Bailey, A., (2012) Commonalities and Differences in Labour Market Developments and Constraints in Different EU Regions.

Villa, J.M. (2011) DIFF: Stata module to perform Differences in Differences estimation. Statistical Software Components.