Evaluation of the effects of intervention programs
for unemployed: An overview of methods
Authors: Jaromir Vrbka, Georgeta Anca
Petre Luțan

EVALUATION OF THE EFFECTS OF INTERVENTION PROGRAMS FOR UNEMPLOYED: AN OVERVIEW OF METHODS

Jaromir Vrbka[1,a,*] and Georgeta Anca Petre Luțan[2,b]

[1]The Institute of Technology and Business in Ceske Budejovice, The School of Expertness and Valuation, Okruzni 517/10, 370 01 Ceske Budejovice, Czech Republic
[2]The Management Doctoral School, The Bucharest University of Economic Studies, 6 Piata Romana, 1st district, Bucharest,
010374 Romania
[a]vrbka@mail.vstecb.cz, [b]ancapetre24@gmail.com
*Corresponding author

Abstract: To assess the impact of policy instruments for the specific target group of unemployed jobseekers, various statistical and econometric methods are usually used, the most commonly used of which are unambiguously counterfactual methods. Their principle lies in creating a counterfactual situation, which estimates what would have happened in the absence of intervention. In quasi-experimental studies, this evaluation can be performed using a control group of non-treated individuals. In this paper, we provide an overview of the methods of creating this control group and consequent evaluation of the effect of the programs and we also analyze the use of methods in studies conducted in the EU Member States.

Keywords: counterfactual evaluation; quasi-experimental methods; intervention; unemployed jobseeker.

JEL Classification: C54, J08, J64

1. Introduction

Employment is one of the main objectives of the European Union (EU) policies. In EU countries, several intervention programs have been implemented to increase employment and reduce the number of unemployed people. The target group of these programs are unemployed jobseekers, most often young unemployed or long-term unemployed, or problem groups of low-skilled or low-educated jobseekers. The aim of these programs is to improve the placement of the target group on the labor market and to increase the sustainability in the employment they found. As these intervention programs are often (co)funded from EU sources, the EU is required to carry out rigorous evaluations of these intervention programs. The question is not only what has been achieved by using the funds, but also whether their use has caused a significant difference or what would have been achieved if they had not been used, in accordance with Elia et al. (2015). The budgetary problems of the EU Member States and the subsequent efforts to save budgets also increase the pressure on the need to know how effectively EU funds are spent. In this context, the question, therefore, arises of the real impact of policies, Potluka et al. (2016)

In recent years, the Commission's Directorate-General for Employment, Social Affairs and Inclusion (DG EMPL) has launched a number of initiatives to support the submission of impact

evaluations of European Social Fund interventions, Elia et al. (2015) This is probably the reason that the European Commission is also putting pressure on applying evaluation methods to empirically test the impact of EU cohesion policy, stated by Potluka et al. (2016). Impact evaluations are obligatory for the EU Member States as they are listed in EU Regulation no. 1303/2013: "*Ex post evaluations should be carried out to assess the effectiveness and efficiency of the ESI Funds and their impact on the overall objectives of the ESI Funds and the Union strategy for smart, sustainable and inclusive growth, taking into account the objectives set for that Union strategy. For each of the ESI Funds, the Commission should prepare a summary report summarizing the main conclusions of the ex-post evaluations.*" (EU, 2013) "*Evaluations shall be carried out in order to improve the quality of programming and implementation of programs, as well as to assess their effectiveness, efficiency, and impact. The impact of the programs shall be assessed in the light of the role of each ESI Fund in relation to the objectives of the Union strategy for smart, sustainable and inclusive growth, taking into account, as appropriate, the program's scale in relation to GDP and unemployment in the program area concerned.* (EU, 2013)

In order to evaluate whether the intervention has achieved the desired effect, has not achieved any effect, or even fails to achieve its objectives, it is necessary to measure and compare the change in development attributable to the intervention and without the intervention. This change is referred to as "impact" (which can be negative or positive). This term is used to measure the effectiveness of the intervention, Peterson et al. (2018).

Impact evaluations usually relate to a policy that has already been implemented. Prospective evaluations (ex-ante) address the question "Is this a good program?" by examining the links and consistency between global objectives, specific objectives and program measures; the existence and relevance of the outputs, results and indicators of program impact and the reliability of the level of quantification of objectives. Mid-term evaluations during the program examine whether the program remains relevant to local needs, the degree of effectiveness achieved in the monitoring indicators, the quality and relevance of these indicators, and the quality of program management. Finally, the prospective evaluations (ex-post) use the resulting monitored data to compare the expected objectives with the objectives actually achieved, to evaluate the long-term impact of the program, Gripaios et al. (2008)

Implementation of impact evaluations of EU cohesion policy allows us to examine, how one of the most important EU policies works. The studies about impact evaluations of policy programs are focused on the assessment of national policies and employ quantitative econometric methods to evaluate the impact of the European Social Fund employment interventions in the EU employment policy, Potluka et al. (2016). This study is therefore focused on the overview of the methods, usually and most frequently used for the impact evaluations of programs for unemployed jobseekers. The development of these evaluations has been rapid recently, but in fact realized evaluations were mainly based on qualitative research methods. Quantitative methods of impact evaluation, especially counterfactual impact evaluations (CIE) have been known and used for decades, but their application to the evaluation of public interventions within the EU cohesion policy is rare. These methods are not yet generally known, either by the managing authorities of EU country policy programs or by the evaluator, in accordance with Potluka et al. (2016). Therefore, we consider the main contribution of this study to provide an overview of methods, which, despite the rarity of the above-mentioned implementation of evaluations of policy programs in EU countries, the most commonly used and most suitable for the implementation of impact evaluation (Svabova et al., 2019b).

Evaluation of the effects of intervention programs
for unemployed: An overview of methods
*Authors: Jaromir Vrbka, Georgeta Anca*
Petre Luțan

The article is organized into the following chapters. The introduction shows the basic idea of evaluation of the impact of policy programs in EU countries. The review focuses on evaluations carried out in EU countries in recent years from the methods used point of view. The methodology aims at the base of the methods of carrying out impact evaluations of policy programs. The Results chapter presents the most frequently used methods of evaluations of intervention programs for unemployed carried out in EU countries. Conclusion and discussion summarize the base idea of the counterfactual situation methods and recalls the problems that usually arise in carrying out impact assessments.

2. Review

In a literature review, we focus on the methods that authors most frequently used in recent years to carry out evaluations of interventions for unemployed people in EU countries. In doing so, we use the Counterfactual Evaluation Archive (CEA) created by the Center for Research on Impact Evaluation (CRIE). This organization was established as a part of the Competence Centre on Microeconomic Evaluation and provides scientific expertise and methodological support on Counterfactual Impact Evaluation (CIE) to the Directorate General Employment, Social Affairs and Inclusion (DG EMPL) and Member States for the impact evaluations of interventions funded through instruments managed by DG EMPL, namely the European Social Fund (ESF) (CRIE, 2019). The CEA database contains studies by authors from EU countries, published as articles in journals or conferences and focused on (counterfactual) impact evaluations of labor market policy instruments.

According to CEA database, propensity score matching is the most frequently used method for CIE. The numbers of uses of the methods in the realized studies in the EU are in the following figure.

Figure 1: Numbers of studies on CIE according to the CEA database
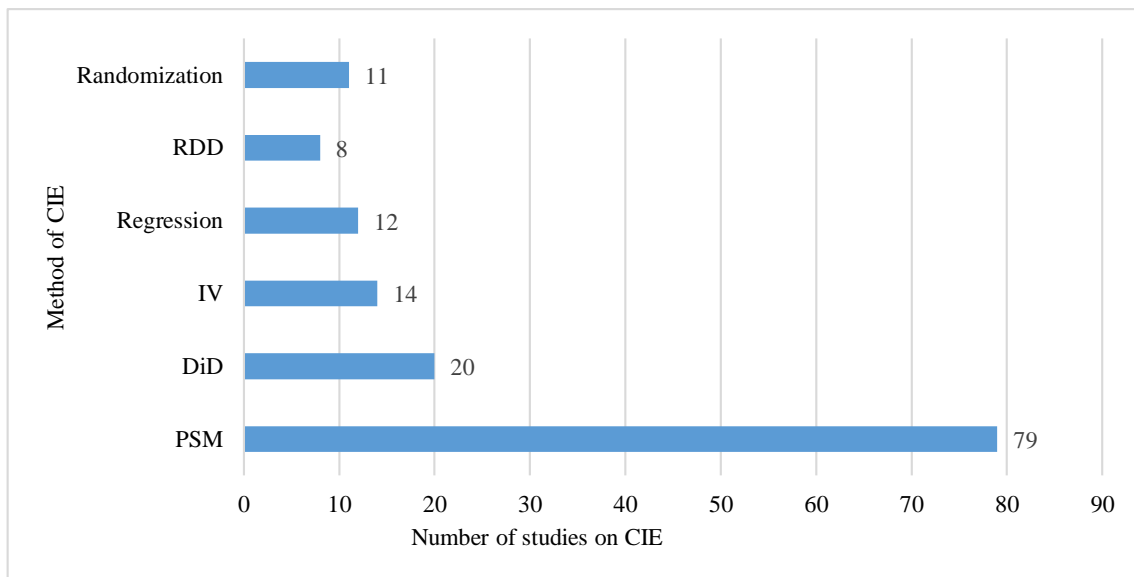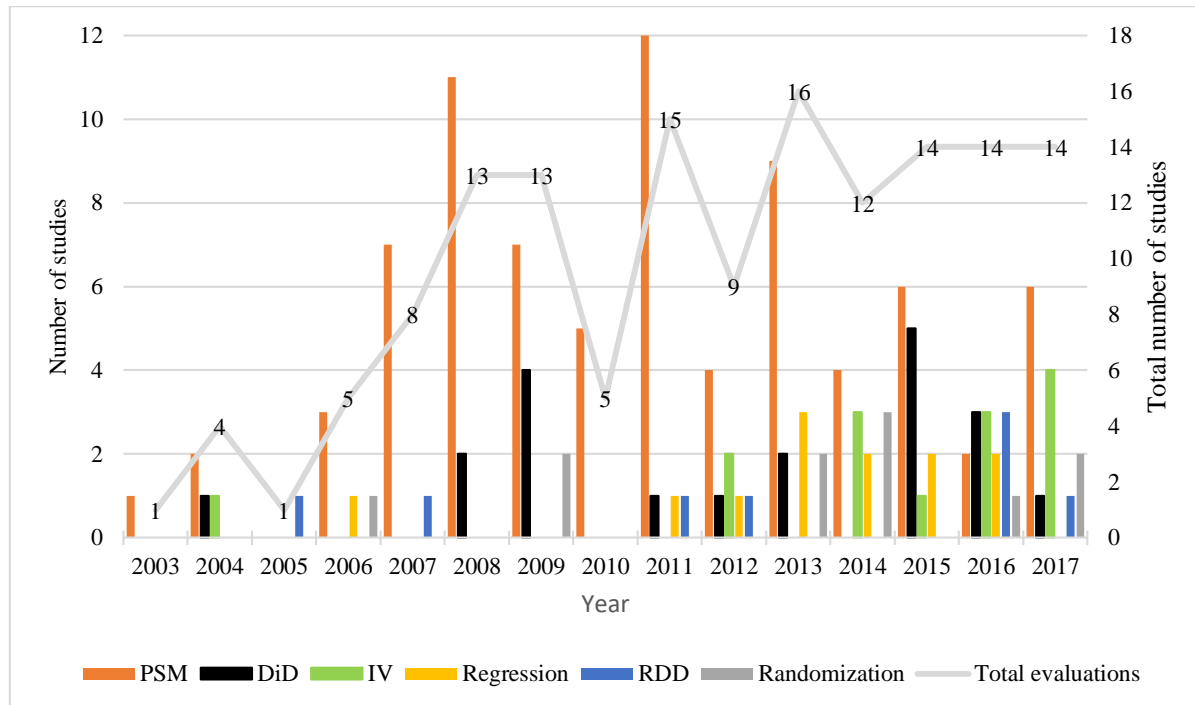

Source: own elaboration based on the data from CEA database

The total number of studies is 144. The number of studies on CIE of political programs for unemployed people increased mainly in the last years (the database contains the studies until December 2017). The following figure typified frequency distribution of the methods in the studies using CIE in the years 2003 – 2017 (PSM = propensity score matching, DiD = difference

in differences, IV = instrumental variables, RDD = regression discontinuity design). The total number of CIE studies over the years is also shown.

Figure 2: Numbers of uses of the method of CIE over the years 2003 - 2017

Source: own elaboration based on the data from CEA database

The number of studies has stabilized over recent years. Propensity score matching still leads in the use of methods. In recent years, all the above-mentioned methods of CIE have been used. However, it is evident that, given the EU requirements for carrying out these evaluations and the number of programs implemented under the programs for unemployed in each country, this number is still insufficient. Of course, analysts carry out also counterfactual evaluations of other interventions that are not always co-financed by the ESF, so it is clear that in fact the CIE methods are used more widely than shown in the previous figure. Nevertheless, we can say that there are not many of these evaluations of labour market policy programs and that evaluations of the results of interventions after their implementation is not yet common practice in EU countries. Rather, it is possible to say that various pilot programs are being implemented, in which countries acquaint themselves with counterfactual evaluation on selected instruments for unemployed jobseekers and learn to implement them.

3. Methodology

According to KPMG (2015), the assessment of the impact of the intervention is an analysis of cause-effect relationships (causality). The objective of evaluating an intervention is always to answer the question, whether participation in the program has the desired effect on the outcome variable. The evaluation of the programs, therefore, analyzes to what extent the changes in the result variable of the treated participants can be attributed to their participation in the intervention. Simply, this effect of the intervention can be obtained as the difference between the value of the outcome variable after the intervention and the value that this variable would have obtained in the absence of intervention (net difference). Thus, in this context,

*Evaluation of the effects of intervention programs
for unemployed: An overview of methods
Authors: Jaromir Vrbka, Georgeta Anca
Petre Luţan*

causality refers to the net increase or decrease in the outcome of the intervened units attributable to their participation in the intervention program, according to Cerulli (2015).

The strategy for the direct detection of the actual effectiveness of the program, in general, does not exist. There could be possible, for example, simply ask participants in the pilot program, how the participation in the program aimed for instance at improving their applicability on the labor market has really affected their employability. The answer to this question could, however, have less explanatory power than would be required: Although the respondents were quite truthfully, their voluntary participation in the program could be a reflection of their, say, awareness, effort or motivation, not only to take part in the program but also for example to employ. These respondents should, therefore, regardless of participation in the programs, have the effort to find a job. In other words, these respondents would not be a fully representative sample of the unemployed population for whom the program is intended.

Frondel and Schmidt (2005) stated as a result, the decision of policymakers whether or not it is appropriate to finance the program is usually a decision without any certainty. It is therefore highly advisable for policymakers to apply the results of the scientific evaluations of programs in their decisions. Consequently, the evidence and conclusions derived from accepted program evaluations should be the basis of any policy decisions, the setting of future parameters of intervention programs or their funding. For these reasons, each program should be evaluated, especially the publicly funded program, which was designed to create some changes. The purpose of this evaluation is to determine whether the program has achieved its intended impacts, and also to assess how the proposed program can be further improved, according to Onwuegbuzie (2017).

Knox et al. (2009) provided the term "rigorous evidence" of the intervention means quantification of the effects of the intervention based on evaluation designed such a way, that it can be concluded that the social intervention actually caused the effects observed. Such studies use primarily high-quality experimental or quasi-experimental designs.

A key concept of experimental studies (also called randomized controlled trials - RCT) is the randomization of individuals into a treated and control group. For individuals wishing to participate voluntarily in the program, the random mechanism decides whether they are truly "allowed" to join the program (these will form a "treated group") or whether they are "excluded" from the program (these will create a "control group" of non-treated individuals). Thus, all individuals in the control group were also applicants for the program but were not randomly selected to take part in it, Frondel and Schmidt (2005). Then, if the samples are large enough, random selection will completely balance all relevant observable and unobservable characteristics of the treated and control groups, thereby achieving comparability between them. Then, the average effect of the intervention can be ascertained by the difference in the average results of these randomly-formed groups of treated and non-treated individuals Such an evaluation of intervention then provides evidence that any differences in the outcomes of the two groups in the post-intervention period are due to the participation of the treated group in the intervention, Knox et al. (2009). Many authors call the method of randomized controlled experiments the "gold standard" of evaluation methods. However, in real-life applications, intervention is rarely assigned randomly. Therefore, it is much more common to apply a quasi-experimental approach rather than an experimental design evaluation.

The aim of a wide range of quasi-experimental approaches to mimic randomization (without its actual use) in impact evaluations is to build a comparison group of individuals who did not participate in the support program, stated by Arco-Tirado et al. (2018), so are not affected by political interference, Lenihan and Hart (2004). Thus, the quasi-experimental procedures identify a group of individuals that is as close as possible to the treated group in terms of basic

(pre-intervention) characteristics (also called covariates). This control group captures what the results would be if the program or policy were not implemented - the so-called counterfactual situation, White and Sabarwal, (2014). In literature, this situation is most often described by the words "what would have happened in the absence of intervention". Thus, the interventional effects can be measured as the difference between the treated individuals and their counterfactual individuals in the post-intervention period. (Arco-Tirado et al., 2018).

Since the control group is designed so that the non-treated individuals are as similar as possible to the treated individuals, the difference in results between the treated group and control group can be ascribed to the program or policy. (White and Sabarwal, 2014) In the experimental approach, this follows directly from the construction of the method. In the quasi-experimental approach, it is ensured by matching individuals based on their observable characteristics. These types of intervention policy impact evaluations are called Counterfactual Impact Evaluations (CIE) and are the most commonly used evaluation methods. CIE methods are very data-intensive. (Plzakova, 2016)

The primary disadvantage of quasi-experimental methods of CIE is that they cannot exclude the possibility of bias, which can cause deviations in the evaluation of the impact of the intervention. We can imagine this bias, the so-called "selection bias" as a factor of motivation, talent or self-selection, which plays a strong role for the treated individuals to decide to participate in the intervention and also in their subsequent job search, so that in their values of result variables. In quasi-experimental approaches, if we do not take into account the presence of such a motivation factor, it may cause that the better outcome of the intervention's participants will be wrongly attributed to the effect of the intervention on them. Armstrong and Patnaik (2009) thus, the evaluation of intervention as a whole may be biased and the impact of the intervention overestimated.

3.1 Matching

In the quasi-experimental methods, the comparison group is most often set up by matching individuals with respect to their observed pre-intervention characteristics. (White and Sabarwal, 2014) Perfect matching would be created if each individual in the treated group was matched with an individual in the comparison group who is identical in all relevant observable characteristics such as age, education, religion, profession, property, etc. However, such an ideal situation is very difficult to achieve in reality. Finding the right match for each program participant therefore usually includes the most accurate estimation of variables or determinants that explain the individual's decision to join the program. (White and Sabarwal, 2014)

Some authors also use a number of various matching methods when developing the CIE to find out which one achieves a better balance on covariates for treated and control units before calculating the effect of the intervention. Once both groups are fully comparable based on their covariates, the impact of the program is quantified as the difference between the observed results in the group of treated individuals and the results estimated from the control group. (Arco-Tirado et al., 2018)

From the methods of matching individuals, in many cases, the method of exact matching is considered ideal. However, this method is very data-intensive in such sense, that small samples may have a very weak exact match of the values of covariates of treated and unsupported covariates, resulting in small matched samples. We consider the propensity score matching to be the most commonly used method in counterfactual impact evaluations, also in the field of evaluations of interventions focused on unemployed jobseekers.

Evaluation of the effects of intervention programs
for unemployed: An overview of methods
Authors: Jaromir Vrbka, Georgeta Anca
Petre Luțan

3.2 Methods of counterfactual impact evaluation

Cerulli (2015) divides (quasi-experimental) methods of counterfactual evaluation into two basic groups according to the selection mechanism of intervention allocation. It states that policy programs do not usually select individuals to be treated by a random selection. Instead, individuals are selected for the program either on a self-selection basis or by the Agency on the basis of a pre-determined selection mechanism. This selection mechanism can be based on observables or unobservables characteristics of individuals. (Cerulli, 2015)

In the case of selection of individuals on the basis of observable characteristics, the analyst knows exactly and can measure the factors that determine the individual's choice for the program or its selection by the agency. Methods of CIE based on selection based on observable characteristics include: (Cerulli, 2015)
- regression adjustment method;
- before-after comparison;
- matching method, most commonly used:
    - exact matching;
    - propensity score matching;
- statistical reweighting.

Based on the selection of participants to the intervention according to their unobservable characteristics, the most commonly used methods of CIE are: (Cerulli, 2015)
- difference in differences;
- regression discontinuity design;
- instrumental variables method.

4. Results

In describing the methods of counterfactual impact evaluation, we will focus on the methods most commonly used when evaluating programs for unemployed jobseekers. The effect of the intervention is quantified as the difference in the result variables in the treated and non-treated unemployed group. Both groups are monitored for a specified impact period, for example 6 months, 12 months or even 24 months after the end of the intervention. The evaluation focused on the results of treated individuals in the labour market during the impact period. The result variables can be, for example:
- time (months or days) to the first obtained job,
- duration (months or days) of the first job,
- the number of months from the impact period spent on employment.

4.1 Counterfactual as self-estimated by program participants

This method relies on the self-assessment of a hypothetical counterfactual by program participants. In practice, this approach works quite simply: participants are required to provide information about their value of the outcome variable after participating in the program, and subsequently to estimate their value of the outcome variable, provided they do not complete the program. The difference between these two values is then equal to the individual effect of the intervention on the beneficiary (Mueller and Gaus, 2015). According to Onwuegbuzi (2017), this form of counterfactual analysis can be used to analyze not only whether the program had an effect (or not), but also why it was. The disadvantage of this method is the bias that the self-assessment brings to the results. Program participants may overestimate or underestimate the actual impact of the program and the situation in which they would not participate in the program. The motivation factor plays a major role in the results of the self-evaluation method.

4.2 Before-after comparison

A commonly used evaluation strategy by creating a credible counterfactual situation is the "before-after comparison" method by comparing the intervention individuals with themselves in the pre-intervention period. Thus, in this "before and after" method, the comparison group is the group of participants themselves, before the program is implemented. On the basis of long-term data, the overall effects of the program are evaluated on the basis of changes in the outcome variables between the pre-intervention period and the post-intervention period by a simple difference in the mean values of these variables. (Frondel and Schmidt, 2005) For example, in the evaluation of the energy-saving programs for families, when comparing the energy consumption of a family before and after the intervention, we cannot exclude, for example, the weather factor (rainy or freezing days) that affects their energy consumption.

4.3 Counterfactual evaluation using exact matching

Exact matching determines the most general possible model of program results in terms of observable individual data. Each treated individual for whom at least one match in the control group can be found is used to estimate the effects of the intervention by comparing the sample averages. (Frondel and Schmidt, 2005) Matching is carried out according to all or selected covariates, the exact match of the treated and non-treated individual is sought, for example in the age, education, gender, length of previous unemployment, etc. The disadvantage of this method of matching for carrying out the evaluation is a so-called dimensionality problem. Exact matching is possible only if covariates have a very low dimensionality, for example only 3 values. However, if we have a small sample and/or covariates are large in size, or worse, they are continuous variables, the exact matching method is impossible. In addition, Lee (2016) states that the dimensionality problem can result in very few observations in matched samples.

4.4 Counterfactual evaluation using Propensity score matching

To address the dimensionality problem of exact matching, Rosenbaum and Rubin (1983) suggested that individuals could be matched based on only a single value - the propensity score. Using a single propensity score number that synthesizes all values of observable characteristics of individuals, the individuals from the treated and control groups are then matched and the effect of the intervention can then be calculated as an average of the differences in their results. (Cerulli, 2015) By matching the propensity score, which reflects the likelihood of participating in the intervention, we can achieve consistent estimates of the effects of the intervention in a way similar to the exact matching based on all covariates (CRIE, 2019). As reported by Arco-Tirado et al. (2018), the propensity score is a conditional probability that an individual with certain observed baseline characteristics will participate in the intervention. The model of this probability is most often created on the basis of logistic regression, but the method of binomial CART trees is also used.

After estimating the propensity score of participation in an intervention, for each treated individual, one or more non-treated individuals with the closest propensity score are assigned (Svabova et al., 2019a). The assignment takes place in different ways, for example:

- nearest neighbour matching,
- radius matching,
- caliper matching,
- kernel matching.

The disadvantage of this method is its high data intensity. The evaluator must have truly large samples of both treated and non-treated jobseekers to be able to do the matching, ensuring

Evaluation of the effects of intervention programs
for unemployed: An overview of methods
Authors: Jaromir Vrbka, Georgeta Anca
Petre Luţan

the best comparability of the paired matched. If the matched samples are as similar as possible, the difference in the result variables can be attributed to the intervention.

4.5 Difference in differences

The difference in differences (DiD) method has become a very popular method for estimating the impact of political interventions. This method is used for several subpopulations, some of which are subject to political intervention and others not, with results being measured in each group before and after the political intervention (Trivellato, 2011). DiD method estimates the effects of intervention policy by comparing between individuals and between periods (Wang, 2016).

The difference evaluation design is appropriate if the result variable is a repeatable event (for example, wages) observable before and after the program. (Bertrand et al., 2002) Thus, the minimum set of information always consists of a pair of observations, before and after the intervention, both for the treated and for the non-treated group. (Trivellato, 2011) The basic assumption is that the time trend in the control group is an adequate representative of the time trend that would occur in the treated group in the absence of political intervention. (Athey and Imbens, 2002) In other words, we assume that without intervention, treated and non-treated groups would follow the same trend. However, the development of the treated group will change with the intervention. This change is quantified as the effect of the intervention.

However, this assumption of the same trend is also the weakness of this method. Any change in the development of the treated group towards the non-treated group will be attributed as the effect of the intervention, regardless of the possible development of the environment. This may overestimate the effect of the intervention. This potential error of the DiD method is usually addressed by exact matching of individuals. (Frondel a Schmidt, 2005)

4.6 Method of instrumental variables

If entry to the intervention program is influenced not only by the observable characteristics of individuals but also by their unobservable characteristics (e.g. motivation or skills), estimates of the effects of intervention that we obtain from counterfactual impact evaluation methods (propensity score matching, difference in differences, but also regression model and statistical reweighting) will be biased (overestimated or underestimated). (Cerulli, 2015) The assignment of an individual to intervention becomes correlated with outcome variables.

The solution to this problem is the method of instrumental variables or its generalization, the two-stage least squares method (CRIE, 2019). The average effect of the intervention on the treated individuals can then be quantified if it is possible to identify and measure a variable (instrument) that affects the selection of individuals into the intervention program but does not affect the values of the result variable or other unobserved variables that determine it. (Abadie et al., 2002)

CRIE (2019) proposes as potential instrumental variables, for example, the distance of the place of residence of a participant in the program for unemployed from the place of work or from the relevant labour office.

4.7 Other methods

Other methods for assessing the impact of policy programs include regression analysis, regression adjustment, cost-benefit analysis, shift-share analysis, selection and assistance modeling, panel analysis. Of course, each method has its advantages but also disadvantages.

However, the fact remains that the most used methods in general still remain matching based on the propensity score, instrumental variables, or difference in differences.

5. Discussion

The first very important aspect of quality evaluation of the interventions implemented is the existence and availability of quality data. Evaluation of the effects of the policy and establishing the appropriate counterfactual situation requires reliable data. Unfortunately, there are major problems in this regard, in accordance with Gripaios et al. (2008). The analyst must have access to databases on intervention participants, data on all eligible intervention participants (i.e. those who have not finally participated), the intervention data itself, as well as the results of participants and non-participants after the intervention. However, unless it is a prospective evaluation of future programs where policymakers realize the importance of recording quality and correct data, in most cases of retrospective evaluation, the analyst has access only to general databases of the entire population. In this situation, as Trivellato (2011) mentions, if an analyst has to use universal databases to evaluate the effects of an intervention, he usually faces two problems.

The first problem is the accessibility to the data needed for evaluation and the resulting sample size used for evaluation. The target group of the policy instrument under evaluation is often a small segment of the entire population that the analyst must filter from available data. Thus, if a large database of, for example, country's citizens is used as the data source, the part of the database relevant to the impact evaluation of the instrument is usually small. Often, serious inaccuracy problems arise because the samples are too small to detect the effect of the program and/or potentially distort the results. (Trivellato, 2011)

The second problem is data quality. Result variables, as well as other variables, needed to check eligibility of intervention for individuals, other variables characterizing individuals used in the evaluation may not be fully available, data are often missing or are incorrectly recorded.

In order to avoid widespread problems of data access, quality, abundance and accuracy, it would be best to take into account the data analyst requirements for policy evaluation prior to its implementation.

6. Conclusion

The European Commission's recommendations for policymaking in recent years include the implementation of ex-post evaluations of the realized programs. Each year from the country's budget and the European Social Fund large amounts of money are spent to finance interventions, including interventions for unemployed jobseekers. Therefore, the question arises whether these resources are spent efficiently and whether the programs implemented bring the desired effect: for the unemployed in the form of better or faster employability of the participants and the retention of their employment. However, the effect of an intervention on an individual or on the whole treated group can only be assessed against a situation in which the intervention would not exist. Such a situation, in reality, cannot be achieved. An individual can only be in one of two situations at a given time: either he has participated in the intervention or not. Therefore, it is necessary to model the given counterfactual situation using available data on non-treated unemployed individuals. The use of counterfactual evaluation methods makes it possible to find out what would have happened if the participation in the intervention had not been involved. Consequently, we quantify the overall effect of the intervention as the difference in the outcome variables of the treated group versus their counterfactual non-treated

Evaluation of the effects of intervention programs
for unemployed: An overview of methods
Authors: Jaromir Vrbka, Georgeta Anca
Petre Luțan

group. The choice of the appropriate method of assessing the effects of the intervention and the associated choice of the appropriate method of matching intervention participants with control counterfactual non-participants is in the hands of the analyst and depends mainly on the nature of the data at his disposal. For example, if he knows both pre-intervention and post-intervention data for both groups, treated and non-treated unemployed jobseekers, he can use the difference-in-differences method. If he only knows the data of the treated group in the pre-intervention and post-intervention periods, the only suitable option may be the before-after comparison method. It is therefore very important to monitor the necessary data for carrying out program evaluations. Ideally, this monitoring of appropriate data is already proposed when designing an intervention program.

References

Abadie, A., Angrist, J. & Imbens, G. (2002). Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings. *Econometrica,* 70(1), 91–117.

Arco-Tirado, J., Fernández-Martín, F., Ramos-García, A., Littvay, L., Villoria, J. & Naranjo, J. (2018). A counterfactual impact evaluation of a bilingual program on students' grade point average at a Spanish university. *Evaluation and Program Planning,* 68, 81–89.

Armstrong, J.S. & Patnaik, S. (2009). Using quasi-experimental data to develop empirical generalizations for persuasive. Advertising. *Journal of Advertising Research,* 49(2), 170–175.

Armstrong, J.S. & Patnaik, S. (2009). Using quasi-experimental data to develop empirical generalizations for persuasive advertising. *Journal of Advertising Research,* 49(2), 170–175.

Bertrand, M., Duflo, E. & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics,* 119(1), 249–275.

CEA: The counterfactual evaluation archive. (2019). The Centre for Research on Impact Evaluation. https://crie.jrc.ec.europa.eu/CIE_database/cieDatabase.php

Cerulli, G. (2015). *Econometric Evaluation of Socio-Economic Programs*. Berlin, Heidelberg: Springer-Verlag.

CRIE. (2019). Centre for research on impact evaluation. https://crie.jrc.ec.europa.eu/

Elia, L., Santangelo, G. & Schnepf, S.V. (2015). Synthesis report on the 'Pilot projects to carry out ESF related counterfactual impact evaluations'; European Union, 2015. EUR27622;

EU. (2013): Regulation (EU) No. 1303/2013 of the European Parliament and of the Council of 17 December 2013. Available at: https://eur-lex.europa.eu/legal-content/SK/TXT/HTML/?uri=CELEX:32013R1303&from=EN

Frondel, M. & Schmidt, C.M. (2005). Evaluating environmental programs: The perspective of modern evaluation research. *Ecological Economics,* 55(4), 515–526.

Gripaios, P., Bishop, P., Hart, T. & Mcvittie, E. (2008). Analysing the impact of objective funding in Europe: A review. *Environment and Planning C: Government and Policy,* 26(3), 499–524.

Knox, V., Hill, C.J. & Berlin, G. (2018). Can evidence-based policy ameliorate the nation's social problems? *The ANNALS of the American Academy of Political and Social Science,* 678(1), 166–179.

Lee, M-J. (2016). *Matching, regression discontinuity, difference in differences, and beyond,* 1st ed. Oxford, U.K.: Oxford university press.

Lenihan, H. & Hart, M. (2004). The use of counterfactual scenarios as a means to assess policy deadweight: An Irish case study. *Environment and Planning C: Government and Policy*, 22(6), 817–839.

Mueller, C.E. & Gaus, H. (2014). Assessing the performance of the "Counterfactual as self-estimated by program participants." *American Journal of Evaluation,* 36(1), 7–24.

Onwuegbuzie, A.J. & Hitchcock, J.H. (2017). A meta-framework for conducting mixed methods impact evaluations: Implications for altering practice and the teaching of evaluation. *Studies in Educational Evaluation,* 53, 55–68.

Peterson, I., Maron, M., Moillanen, A., Bekessy, S. & Gordon, A. (2018). A quantitative framework for evaluating the impact of biodiversity offset policies. *Biological Conservation,* 224, 162–169.

Plzakova, L. (2016). Public funding and its impact on tourism development. *Proceedings of the 3rd International Multidisciplinary scientific conference on Social science and Arts - SGEM 2016, Book 2: Political science, Finance, Economics and Tourism:* 537-544, Bulgaria.

Potluka, O., Brůha, J., Spacek, M. & Vrbova, L. (2016). Counterfactual impact evaluation on EU cohesion policy interventions in training in companies. *Ekonomický časopis,* 64, 575–595.

Rosenbaum, P.R. & Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika,* 70(1), 41.

Svabova, L., Durica, M. & Kliestik, T. (2019a). Modelling the costs of unemployment for young graduates in Slovakia: A counterfactual approach. *Politická Ekonomie,* 67(5), 552–566.

Svabova, L., Durica, M., Kramarova, K., Valaskova, K., Janoskova, K. (2019b). Employability and sustainability of young graduates in the Slovak labour market: Counterfactual Approach. *Sustainability*, 11(16), 4462.

Trivellato, U. (2011). Fifteen years of labour market regulations and policies in Italy: What have we learned from their evaluation? *Statistica* 71(2), 167–187.

White, H. & Sabarwal, S. (2014). *Quasi-experimental Design and Methods, Methodological Briefs: Impact Evaluation 8*. Florence, Italy: UNICEF Office of Research.