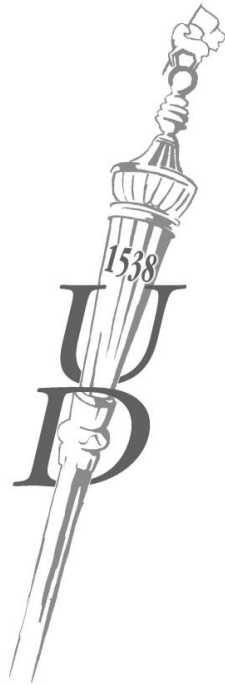


Thesis of Ph.D. dissertation

**INTERNATIONAL COMPARATIVE ASSESSMENT OF
SELECTED MID-TERM INSTITUTIONAL EVALUATION METHODS AND
REGIONAL PROGRAMMES IN THE EU-12**

Balázs Mezősi

Supervisor: Prof. Dr. János Nagy D.Sc.



UNIVERSITY OF DEBRECEN
KERPELY KÁLMÁN DOCTORAL SCHOOL
REGIONAL SCIENCES

Debrecen, 2015

TABLE OF CONTENTS

1	RESEARCH BACKGROUND AND PURPOSE	4
1.1	Relevance and rationale	4
1.2	Personal motivation	5
1.3	Research scope	5
1.4	Scientific relevance and purpose	7
2	RESEARCH METHOD	8
2.1	Hypotheses	8
2.2	Analyses	9
2.3	Research preparations	9
2.4	Research structure	9
2.5	Tasks, techniques and data sources	10
2.6	Research literature	11
3	RESEARCH CONCLUSIONS	12
3.1	Thesis 1 – Cohesion Policy results	13
3.2	Thesis 2 – ROP diversity and similarity	14
3.3	Thesis 3 – Method objectivity, and link to intervention objectives	16
3.4	Thesis 4 – Method simplicity and robustness	18
3.5	Thesis 5 – Method flexibility	19
4	NEW SCIENTIFIC RESULTS	21
5	PRACTICAL APPLICATION OF RESULTS	23
6	REFERENCES	25
7	PUBLICATIONS	28

1 RESEARCH BACKGROUND AND PURPOSE

1.1 Relevance and rationale

Evaluation of funding programmes is a natural expectation of all stakeholders (donors, acceptors, policy makers, broad public), driven by the absence of institutionalised feedback mechanisms, such as the market itself in market economy context.

Evaluations require methods and tools to provide answers to stakeholder questions. Among the many other existing classifications, these methods can be *objective* or subjective in terms of the outputs they provide. Evaluation methods that are *flexible* enough to yield objective, comparable and repeatable, precise outputs are the ones that provide unique opportunity to be used in cross-country cases or even at the entire Community level. Besides, methods are usually required to be *simple* to carry out and communicate and *robust*. These two expectations are seemingly contradictory, however, there are methods capable of balancing between the two. Actually, it is a logical request of evaluators, evaluation commissioners and above all, the final users of evaluations, the policy makers to have such methods and tools. Public interventions in general – and EU funded programmes in particular – have three objectives: effectiveness (and relevance), efficiency and accountability (MEANS, 2000). As a result, evaluations most often concentrate on these topics, as *evaluation themes*. In summary, policy makers and programmers need evaluations (and methods that back them up) as inputs to their activity, that on the one hand resolve the contradiction between broad use and simplicity, and also address each of the three above mentioned objectives.

In this research, I picked three mid-term evaluation methods that I have developed, tested and applied in several EU countries, and I aimed at presenting and analysing their application in international comparisons. These methods have been used in many EU mid-term programme evaluations in the 2007-13 EU programming period, out of which I selected five evaluations from three countries for my research to observe more closely – Bulgaria, Malta and Romania. In order to place the previously mentioned evaluation methods into context, an analysis of EU-12 Cohesion Policy intervention regional results and EU-12 regional (operational) programmes and their evaluations form part of my research. As a result, my research is an international comparative assessment of EU mid-term evaluation methodologies, backed up by a Cohesion Policy programme results and regional programme analysis, addressing five hypotheses.

My first hypothesis assumes that EU Cohesion Policy could contribute to the development of regions in the EU-12, however, it could not counteract the growing interregional disparities in these countries. My second hypothesis addresses the diversity and similarity of Regional Operational Programmes in the EU-12 in the 2007-13 period. The last three of my hypotheses are based on assumptions that the selected evaluation methods fulfil the requirements of objectivity, simplicity and robustness, flexibility and addressing evaluation themes driven by intervention objectives (relevance and effectiveness, efficiency and accountability).

I am aware that my topic selection is on the verge of regional development sciences and management sciences. However, the fact that the selected methodologies have been widely used throughout Europe in evaluations in regional development programme context, creating the foundations for future programming, makes them relevant driving factors of regional development. As such, they make appropriate subjects for scientific research within the scope of regional sciences.

1.2 Personal motivation

I have been evaluating EU programmes and policy documents since 2003. I had the privilege to be among the few experts who had the opportunity to take part in each and every evaluation stage (i.e. ex ante, mid-term and ex post) of the Structural Funds and Cohesion Funds co-financed Hungarian programmes in both past EU budgetary periods (i.e. 2004-06, 2007-13) and I have also been involved in the ex ante phase of the 2014-20 period. I carried out evaluations in six Central and Eastern European countries (Bulgaria, the Czech Republic, Estonia, Hungary, Romania and Slovakia) and Malta, commissioned by the national managing authorities or the respective Directorate General of the European Commission.

I have been project manager, and more importantly, a lead methodological expert and methodology developer in programme evaluation activities for the last twelve years. I had the opportunity to develop, challenge and test methodologies, tools and techniques to support evaluation tasks. Also, I have been member of the European Evaluation Society since 2010.

1.3 Research scope

The scope of my research is three well-acknowledged mid-term programme evaluation methodologies applied in - inter alia - five mid-term evaluations of National Strategic Reference

Frameworks (NSRFs) and Operational Programmes (OPs) in 2007-13 in three EU-12 countries: Bulgaria, Malta, and Romania, with an underlying outlook to EU-12 EU co-funded regional development programmes and regional development measures and their results.

Cohesion policy programmes (NSRF, OP) are the primary, if not the only development resource in the EU-12 countries. Therefore, they provide excellent testbed and offer the most potential results to research programmes, evaluations and evaluation methods. As the main purpose of Cohesion Policy is decreasing regional development gaps, it is important to evaluate intervention *results*, how progress has been made toward this goal. From this perspective, it is the *regional operational programmes* that explicitly address to enhance regional development and minimize disparities, therefore, they provide the best subjects for my assessment.

The reason for the selection of *mid-term evaluation in 2007-13* is that mid-term evaluations carry strong formative, development policy aspect, and within that a sharp focus on the institutional aspect, i.e. relevance and rationale, efficiency and effectiveness. Second, mid-term evaluations cover the most areas by assessing the performance of the actual period retrospectively while they also formulate recommendations for the next programming period. Third, 2007-13 mid-term evaluations are the latest relevant (there are 2014-20 ex ante evaluations already available at present, and 2007-13 ex posts are under way, but their goal, scope and therefore their choice of method is different) approved evaluations available.

The selection of *institutional evaluation* is closely linked to the previous explanation. Mid-term evaluations in the 2007-13 period have a strong focus on efficiency (in compliance with Community guidance (EC, 2006)), and the evaluation of efficiency essentially assesses the performance of the institutional (management and implementation) system.

Methods are both implementing agents and building blocks of evaluations. Adequate and adequately used methods may provide sound evaluations. Therefore I chose to assess methods according to the principles and aspects identified by narrowing and distilling the findings of the scientific literature review and my own experience.

The reason for choosing *EU-12 countries* was practical: the accession of these countries in the EU implied the most changes, also in terms of Cohesion Policy. With Hungary being one of them, I had a special interest to use these countries as a basis for comparison on the application of methods. Also, I had personal work experience with their country context, programmes, evaluations and methods.

1.4 Scientific relevance and purpose

There have been methodology handbooks issued by both the EU Commission (DG Regio), member state central and local governments, professional organisations such as associations and other stakeholders on the practical means of programme evaluation. I have covered and referred to the most relevant of these works in my dissertation as theoretical and scientific bases supporting practical evaluation activities.

What is scarce, however, is new methods, practical feedback and guidance on using evaluation methods in diverse context. The feedback I am reflecting in this paper has evolved through 20+ programme evaluations in 7 countries over 10+ years, capable of acting as impartial and focused synthesis to enrich and complement the scientific background already available on evaluation methodology.

From a scientific perspective, the thesis serves the following purposes:

- 1. Directly: Analysing internationally validated and accepted EU programmes, mid-term evaluations and evaluation methods of the 2007-13 programming period in the EU-12**
 - i) Presenting the high level regional results of Cohesion Policy in EU-12;
 - ii) Providing a detailed description and analysis of the methodologies;
 - iii) Demonstrating their use through international examples;
 - iv) Synthesising the advantages and limitations of the application of methods.
- 2. Indirectly: Supporting evaluation activities foreseen for the 2014-20 programming period by formulating conclusions on existing methodologies and focuses**
 - i) Supporting EU Cohesion Policy related programme development and evaluation;
 - ii) Supporting non-EU programme planning;
 - iii) Facilitating evaluation capacity building on both programmers' and evaluators' side.

2 RESEARCH METHOD

2.1 Hypotheses

As a result of the scientific and practical literature review, coupled with my own international experience, I identified and pursued to challenge the following five hypotheses (H1-5) within the context of this research. Hypotheses 1-2 cover Cohesion Policy and regional development programmes in EU-12, while Hypotheses 3-5 deal with my main topic, method characteristics.

H1: EU Cohesion Policy could contribute to the development of regions in the EU-12, however, it could not counteract the growing regional disparities within these countries. ¹

H2: Regional Operational Programmes in the EU-12 countries in 2007-13 are diverse in terms of their approach and their relative weight in the NSRF despite their content (interventions) and their mid-term evaluations are very similar².

H3: The selected evaluation methodologies are capable of providing objective³ basis for evaluation and address at least one of the general intervention objectives of effectiveness (and relevance), efficiency and accountability⁴.

H4: The selected evaluation methodologies provide simple⁵ and robust answers to evaluation questions.

H5: The selected evaluation methodologies have the flexibility to be customised to the context of the evaluation.

¹ The negative assumption in the hypothesis is backed up by the Sixth Cohesion Report (EC, 2014, p.232)

² Both diversity (as a result of different economic structure, regional authority system, general approach of central administration to regional development, etc.) and similarity (reflected in EC and scientific publications) are assumptions based on previous research and work experience

³ For the notion of objectivity, I integrate and use the notions „factual” and „unbiased” (Scriven, 1972) and „that disregards expertise of biased persons” (Scriven, 1997) and „quantitative” (Terstiege, 2012). I do not consider „distance” as I agree with (Scriven 1972) that „Distance does not guarantee objectivity”

⁴ see MEANS, 2000

⁵ I pursue the notion of simpleness in the sense of „easy to understand” and „easy to communicate” to broad public

2.2 Analyses

My research is based on five hypotheses to be tested via two international comparative analyses: 1) a regional programme analysis and 2) a methodology analysis.

For the *regional programme analysis*, Cohesion Policy programmes in general and specifically Regional Operational Programmes (ROPs) in EU-12 countries in 2007-13 acted as vehicles to the testing of my first two hypothesis (H1-2) on the results of Cohesion Policy on development and regional disparities in the EU-12, and the diversity of programme approaches and relative weight of ROPs in their NSRFs despite the similar content.

For the *methodology analysis*, I have selected three evaluation methodologies that I have developed, tested and applied in many EU countries over the past 10 years. My aim was to conduct an analysis via comparing the experience drawn through the application of these methodologies in diverse country and programme contexts in order to test the last three hypotheses (H3-5) of the research.

2.3 Research preparations

The first step of the research was identifying a suitable selection of programmes and evaluations to conduct the international comparative methodology analysis. For the *regional programme analysis* (hypothesis 1-2) I used all EU-12 country ROPs and OPs with regional development measures. For the *methodology analysis*, the selection mostly included objective criteria (country, programme, evaluation). Out of the EU-12 countries having acceded the EU in 2004 and 2007, seven have been explored in detail. As a result, a Bulgarian and two Romanian programmes have been selected, complemented by two Maltese programmes (being similar in many aspects to CEE countries and therefore providing grounds for comparison) for the methodology analysis and testing hypothesis 3-5.

2.4 Research structure

The research set-up (see Figure on the top of the next page) summarises the frameworks of the research in terms of both the methodology analysis and the regional programme analysis and acts as a reference point throughout the research activities.

Figure 1 – Summary of research set up

Analysis	Focus	Hypothesis	Research aspect	Subject
International comparative regional programme analysis	All Regional Operational Programmes in the EU-12 in 2007-13	H1-2	Allocation approach Relative weight in NSRF Content (measures) Result analysis	1 All EU-12 regions 2 All EU-12 Cohesion policy programmes 3 All EU-12 ROPs and OPs with regional development measures (altogether 46 OPs + 5 programmes including regional development measures)
International comparative methodology analysis	Three methodologies (SPOT, PCA, LTA) used in EU-12 mid-term evaluation in 2007-13	H3-5	Objectivity Link to objective Simplicity, robustness Flexibility	1 NSRF mid-term Romania 2 SOP-IEC mid-term Romania 3 OPI mid-term Malta 4 OPII mid-term Malta 5 OPRD mid-term Bulgaria

Source: own construction

2.5 Tasks, techniques and data sources

I planned and executed the research as an iterative process which finally led to the findings, lessons learnt, conclusions and new scientific and practical, applicable results.

The primary means of testing the above hypotheses was desktop research and literature review. This activity covered the relevant regulations, policy documents, methodological papers, programmes and their evaluations and other related documentation. I also conducted interviews with relevant programme managers and programme evaluators in all countries researched to cross-check and challenge the hypotheses in the course of methodology development, the results of which have been used as input to this study. In the research, I used multiple data collection and analysis techniques. The primary objective was the systematic information collection, covering programming documents (NSRF, OP, Action Plans, measures, etc.), relevant (mid-term programme) evaluations, a sample of / extract from application files, strategic and monitoring reports, status reports on the implementation, annual reports, published, aggregated progress tables for validation of project level data set and other background strategic documents.

Data analysis included the examination of the availability, the relevance and the suitability of the indicators in order to map intervention logic, to take hold of and operationalise the results as well as to follow up development (baseline, milestone, actual indicator values). It also covered an analysis of the data collected from the implementation system and other databases, from statistical and research data and collected data, e.g.: project level data sets (UMIS, SMIS), chronology of calls request, the financial allocation by year by intervention or operation and the publicly available statistical data used from Eurostat and national statistical offices.

I also conducted personal interviews with the stakeholders, evaluators and policy officers directly or indirectly affected by the implementation or results of the programmes; exploring their opinion and their attitude to the results, documenting the collected information in detail.

2.6 Research literature

My research has been built on the foundations provided by the processing of literature, consisting of evaluation history (based inter alia on Patton, Legge, Horváth, Pálné, Faragó, Hirschmann, Rossi, Freeman, and Wright, Cassen, UN and OECD documentation) and evaluation theory (based inter alia on Bradley, Weiss, Huey-tsyh, Chen, Sanders, Worthen, OECD, UNDP and Tavistock), evaluation approach and classifications (based on Potter, Pollitt, Pawson and Tilley, Scriven, Weiss, Louw, Molle, Jobbágy, Lóránd) intervention logic and classification (based on Fitz Gerland, ESRI and EC documentation). This section is followed by a general overview of practical EU evaluation context, encompassing EU evaluation related EC regulations and methodology guidance (based on Scriven, Potter, MEANS, OECD, Tavistock, and EC), complemented by EU evaluation framework (EC and MEANS) and evaluation experience in the selected countries of Bulgaria, Romania and Malta (based on ACIS, UMIS, PPCD, SOP IEC MA).

In the research section, I present the methodology analysis: methodologies used in mid-term programme evaluations in the selected countries in 2007-13 and provide an international comparison and assessment on their application. (ACIS, PPCD, MRDPW, UMIS). Also in this section, I present a regional programme analysis on the Regional Operational Programmes of the EU-12 countries selected for the research, with special focus on, Bulgaria, Malta and Romania (EC, Eurostat, 2008 data sets)⁶.

⁶ The reason for using 2008 data is that reference data should comply with the 2007-13 period regional policy intervention values set in 2008.

3 RESEARCH CONCLUSIONS

This section contains the conclusions of my research conducted following my five hypotheses tested through two analyses. Section 3.1 contains the answers to the first hypothesis (H1) on the results of Cohesion Policy in the EU-12 regional development, Section 3.2 summarises my conclusions on my second hypothesis (H2) on EU-12 ROP diversity and uniformity in 2007-13 while Section 3.3-3.5 cover the answers to the last three hypotheses (H3-5), all focusing on evaluation methodology.

All methodologies described and analysed in this paper have been developed, tested, tailored and simplified through a minimum of 5-6 years in the course of iterative processes. All of them have been applied in several member states and have been accepted by local authorities and the EU Commission as methods capable of addressing and answering their evaluation questions. In conclusion, each of these methods provide a safe and cost-effective basis for evaluation, with regard to their potential advantages and limitations.

The table below summarises the main findings of my research on H3-5 through providing a brief answer to these hypotheses in relation to the methodology used. This table provides an overview of the aggregated results that will be further detailed in the next sub-sections.

Figure 2 – Summary table of conclusions by hypothesis (H3-5)

Applied methodology	International examples	H3: objective, linked to objectives?	Objectives linked	H4: simple and robust?	H5: flexibility?
Simple Progress Overview Tool (SPOT)	1 NSRF mid-term Romania 2 SOP-IEC mid-term Romania 3 OPI mid-term Malta 4 OPRD mid-term Bulgaria	yes	Effectiveness and efficiency	yes	yes
Post-contractual assessment (PCA)	1 OPI mid-term Malta 2 OPII mid-term Malta	yes	Efficiency (indirectly accountability)	partly	yes
Lead time analysis (LTA)	1 OPRD mid-term Bulgaria 2 OPII mid-term Malta	yes	Efficiency (indirectly accountability)	yes	yes

Source: own research conclusions

Having assessed ROPs of the EU-12 countries and the regional results of Cohesion Policy interventions and the content, application and result of the preselected prominent methods of the 2007-13 EU mid-term programme evaluations, my conclusions by hypothesis are as follows:

3.1 Thesis 1 – Cohesion Policy results

Hypothesis 1: EU Cohesion Policy could contribute to the development of regions in the EU-12, however, it could not counteract the growing regional disparities within these countries.

The objective of EU Cohesion Policy is the moderation of regional inequalities. Though Cohesion Policy has justifiably supported development in the EU-12 countries in the 2004-2011 period (EC, 2014), it has failed at mitigating sufficiently the increasing interregional polarisation within the EU-12 countries according to my analysis.

Apart from the most widely acknowledged development indicator, GDP per capita, I based my assessment on two relevant items of regional development theory: the *Williamson hypothesis* describing a relation between development and homogeneity that resembles a 'U' upside down (Williamson, 1969), and the neoclassical theory based *beta convergence* phenomenon (Canova, 2004; Dvoroková, 2014) on the diminishing marginal returns on capital, claiming that underdeveloped regions grow at a rate higher than developed regions.

The diversity is obvious in trends and results across the EU-12. The beta convergence effect is underpinned by the relative development figures of the EU-12 since the accession times. At the regional level, however, interregional inequalities in countries have been disclosed through the provision of the minimum and maximum relative GDP/capita figures of the NUTS2 regions.

The changes in the 2004-2011 period in the countries of the EU-12 with more than one NUTS2 level territorial unit are indeed dramatic: there is an overall improvement effect assumedly partly deriving from the EU co-financed interventions - similarly to other previous study findings - however, the development gaps within countries has grown enormously between 2004 and 2011. In some cases, relative GDP/capita figures compared to EU-28 average figures have more than doubled between 2004 and 2011.

More developed countries suffered less from increasing gaps, while relatively underdeveloped ones has shown outstandingly growing gap figures. This phenomenon underpins the Williamson hypotheses, as in relatively more developed countries (the Czech Republic, Hungary, Slovenia), the development gap between regions did not increase in proportion with the huge leap of figures in the less developed countries (Bulgaria, Romania, Poland, Slovakia).

According to the last, accessible self-assessment on the performance of Cohesion Policy by the European Commission (EC, 2014), the Cohesion Policy contributed to the annual GDP growth of 0.5-1.0% in the EU-12 countries in the 2007-13 period, providing evidences on its

contribution to member state development. However, the report does not formulate straightforward statements on the changes of regional inequalities, though it is a fundamental role of the Policy.

As a conclusion, the attributability of both the beta convergence phenomenon and the Williamson hypothesis could have been identified in the topic of development of EU-12 countries and regions in the 2004-2011 period, indicating that Cohesion Policy has indeed supported development in these countries, but at the same time it has failed in sufficiently mitigating the increasing regional polarisation within the EU-12 countries.

3.2 Thesis 2 – ROP diversity and similarity

Hypothesis 2: Regional Operational Programmes in the EU-12 countries in 2007-13 are diverse in terms of their approach and their relative weight in the NSRF despite their content (interventions) and their mid-term evaluations are very similar.

The analysis identified three distinct approaches used in the selected NSRFs in the EU-12 in the 2007-13 EU budgetary period. Member states either chose to have a ROP for each region, or one for the entire country or to position ROP measures within another programme, the choice of which was logically dependent on factors such as size (i.e. territory), uniformity and national level regional development attitude of the country. 2007-13 approaches live on in the 2014-20 period as well, though there have been countries shifting from one approach to another (e.g. Hungary used to have 7 ROPs in the 2007-13 period while in 2014-20 it has one for the less developed regions together and one for the Central Hungarian Region (i.e. the compulsory minimum). Should there be no EC requirement to separate regional interventions based on the relative development indicators of a region, Hungary is likely to have only one ROP.

The analysis also disclosed a high concentration of ROP measures in the funding structure, with the four top countries accounting for 84% of such funds, while 8 countries share the remaining 16%. The relative weight of the ROP measures in the respective NSRFs shows a great spread, ranging from 5% to 34%. The reason for this lies partly in development approach as well: member states in 2007-13 and previous periods had to decide how to balance sectoral approach with territorial approach. Note that this has slightly changed in 2014-20 with the introduction of thematic objectives in CPR and ERDF regulations.

Another factor behind the allocations is the relative weight of topics that traditionally belong better to the territorial aspect, i.e. tourism, industrial parks, regional level accessibility,

transport, local social, educational and economic issues. These interventions are very similar in all EU-12 regional cases, with slight differences in emphasis, assumedly based on relative development status and institutional system and authority.

It is an obvious disadvantage at the choice of interventions that they target operational level objectives (means), but only indirectly pursue strategic level goals like extended regionalism (Horváth, 2015), more authority to regions or strengthening the second level of cities, i.e. (potential) regional centres to ease the polarised situation governed by the overwhelming power of the centre, with unmatched development differences compared to the countryside.

The ROP mid-term evaluations chosen for the research have shown clearly that the selected evaluations did not follow the focus areas and intensities pre-set by the European Commission (EC). The focus of the evaluations was more intense than the EC had outlined. Evaluation of rationale and impacts were moved to the back, while relevance, effectiveness and efficiency were all highlighted topics. In fact, three out of the five evaluation aspects were differently handled. I have to note, however, that these changes are not dramatic: there is only one step forward or backward on the four-level scale (from 0 to 3 stars, in intensity, see EC, 2006). What is more dramatic, is the narrowing of the scope of the evaluation from five to three topics, as two of them have been practically ignored in the course of the evaluation (in line with the expectations of the commissioners of the evaluations).

The concentration of topics showed high interest of the programme implementers in topics that are closely linked to *implementation* (effectiveness and efficiency) as opposed to *planning* (rationale, relevance). However, it is important to note that the general purpose of mid-term evaluations is the right mixture of the two: in the mid-term phase, there are already results at hand that enable the drawing of conclusions and initiating modifications to enhance performance in the currently ongoing programme. On the other hand, however, mid-term evaluations are commissioned in time periods, when programming of the next programme begins. As a result, conclusions drawn in mid-term are immediately integrated in the planning of the next programme.

In the course of the implementation, it is understandable that implementers, politicians, and the broader public are more interested in current status and spending perspectives. There is a competition for progress among countries and within countries to absorb funds. The higher emphasis of efficiency in the analysis carries two messages in my interpretation: first, in the

countries that I selected there is a long track record of keeping an eye on accountability, and second, it was a clearly articulated need of these member states to reduce spending on the delivery of the funds (i.e. transaction costs) and improve the ratio of result / cost by increasing results, lowering costs or both.

However, the fact that efficiency caught up to effectiveness (claimed to be the single strong focus of mid-term evaluations by the EC) indicates that selected member states were much more keen to learn *how* they perform than *what* they perform. Losing the focus of achieving pre-set objectives and concentrating more on fine-tuning delivery is not entirely in line with the original intentions of the policy makers and the programmers.

The hypothesis has been justified, ROP measures in the EU-12 countries in 2007-13 are indeed diverse in terms of their approach and their relative weight in the NSRF despite their content and their mid-term evaluations is very similar.

3.3 Thesis 3 – Method objectivity, and link to intervention objectives

Hypothesis 3: The selected evaluation methodologies are capable of providing objective⁷ basis for evaluation and address at least one of the general intervention objectives of effectiveness (and relevance), efficiency and accountability⁸.

The SPOT, PCA and LTA methodologies detailed in this paper are capable of establishing an objective foundation and setting the scene for further evaluation. This evidence-based, qualitative basis provides opportunity for exploration of reasons behind which fall into the category of subjective and formative evaluation methods (e.g. consultations, surveys, interviews). Further investigation of reasons behind the facts may alter or underpin the initial finding based solely on figures. Using soft drilling techniques to interpret the objective results is not an option but a requirement to yield real benefits. The objective, quantitative approach preceding the subjective, qualitative techniques provides a logical framework for the evaluations, in line with the community guidance of the triangulation principle.

First, it is more logical and cost-efficient to start consultations having the facts and evidences at hand. Management of an evaluation is a multi-stakeholder process with lots of iterations.

⁷ For the notion of objectivity, I integrate and use the notions „factual” and „unbiased” (Scriven, 1972) and „that disregards expertise of biased persons” (Scriven, 1997) and „quantitative” (Terstiege, 2012). I do not consider „distance” as I agree with (Scriven 1972) that „Distance does not guarantee objectivity, it merely guarantees distance”

⁸ see MEANS, 2000

However, the number of iterations can be reduced drastically if there is a common, objective understanding among stakeholders, which can be established through the results of a quantitative analysis. Therefore, an evidence-based, objective approach helps to build trust and acceptance.

Second, consultations can be more targeted, more effective if there are certain focus points of the investigation, which can be established via quantitative, objective methods. This narrowing of the scope of further explanation not only facilitates the management of stakeholders, but also assists evaluators to concentrate their resources to areas which require a more in-depth analysis based on findings of the objective analysis.

Third, subjective elements, such as consultations may be used to challenge, cross-check and validate objective findings. In case of mismatch, the reasons must be observed on both sides (an inconsistency can obviously derive from interests or opinions, however, data clarity and input failures also lead to improper results). In such a case an iterative process may be initiated to come up with a final conclusion through a convergence of opinions and explanations with the involvement of the parties in dispute.

Being just a first step in evaluation, however, they are not capable of providing an exhaustive evaluation on the topic concerned. Evaluation needs to use objective, fact-based input as one important factor for the evaluation which has to be further investigated and triangulated with qualitative techniques.

All of the selected evaluation methodologies tackle more than one general intervention objectives, or, as they are translated to evaluation activities, the evaluation themes.

The methods have been developed and applied with the purpose to answer evaluation questions under the umbrella of evaluation themes. Therefore it might be obvious that they are linked to one of them, however, planning and implementation are rarely identical.

In all cases examined in this research, methodologies addressed more than one evaluation theme. Following their purpose, there is always at least one evaluation theme they directly target (Effectiveness and efficiency in the case of SPOT, and Efficiency at both PCA and LTA). This high concentration of efficiency might seem surprising first, however, bearing in mind that evaluations (particularly in the mid-term phase) focus on implementation, it is rather obvious.

The methodologies might indirectly address also effectiveness (SPOT) and accountability (PCA, LTA) related evaluation questions. As a result, these three methodologies provide a mixture that can be used to carry out objective, quantitative-based evaluations that tackle all relevant evaluation themes (intervention objectives), i.e. effectiveness (along with relevance), efficiency and accountability.

To sum it up, the answer for the hypothesis is that the three evaluation methods described in this paper are justifiably capable of providing objective, sound basis for evaluation activities and also address the relevant intervention objectives pre-set in MEANS, 2000.

3.4 Thesis 4 – Method simplicity and robustness

Hypothesis 4: The selected evaluation methodologies provide simple⁹ and robust answers to evaluation questions.

The evaluation methodologies selected in this paper can also provide simple and robust answers to the evaluation questions they are related to. However, in the case of the PCA methodology this hypothesis proved to be only partly valid.

All evaluation methods detailed previously use complex data through complex transformations to provide simple outputs. It is much more beneficial to have simple and sometimes complex (but not complicated) results which would require much resource for the evaluators to carry out, and additional resources on both the evaluators' side and the stakeholders' side to interpret the process, the data requirements and the results. On the other hand, these methods provide outputs (used as inputs for the evaluation), but not answers to the evaluation questions as already referred to in the previous section. They are essential building blocks of the evaluation, but their role is to create solid foundations or strong working hypotheses for further investigation.

All three methodologies (SPOT, PCA, LTA) assessed have been found user-friendly with potential results that are easy to communicate and work on with. All of them are fairly simple and powerful. They are not intended to and not designed to cover all evaluation aspects, but they all provide the opportunity of a sound evaluation basis directly proportional to their „costs” i.e. data requirements, preparation requirements and interpretation requirements.

⁹ I pursue the notion of simpleness in the sense of „easy to understand” and „easy to communicate” to broad public.

The explanation of the result that PCA only partly validates the hypothesis does not lie in simplicity, but robustness. PCA inevitably uses sampling prior to evaluation, and there is much room for selection bias or misinterpretation in case of this method. Therefore, interpretation of the results have to be accompanied with the assumptions and sampling methodologies used which limits simplicity of communication, though in their absence the results might be used as means of political or other agendas.

3.5 Thesis 5 – Method flexibility

Hypothesis 5: The selected evaluation methodologies have the flexibility to be customised to the context of the evaluation.

The methodologies assessed have been developed to act as potential bases for international benchmarking. This means that their parameters either originate directly from Community regulation (which is identical for all member states) or the parameters have the flexibility to find a common platform for comparison. This comparability of results makes them beneficial choices when selecting evaluation methodologies. The flexibility of the methods could have been assessed both in terms of planning (in the preparatory phase of the evaluation, exploring how a certain method could be used bearing in mind their limitations) and implementation (retaining the method during the evaluation but changing its use significantly as a result of changing environment, expectations or data availability).

All three methods (SPOT, PCA, LTA) assessed in this paper have shown flexibility and can be tailored to the context (in this sense meaning input and output requirements), though there is a minimal requirement to the application of the method to yield meaningful answers. This covers both data volume and structure (input) and the potential of the application, e.g. number of stages defined in the SPOT methodology (output). All methods have been tested in terms of their customisation potential prior to evaluation and have all proven themselves.

There have also been instances when the flexibility of the method have been tested in the course of the evaluation. For instance, the mid-term evaluation of the NSRF and the SOP IEC in Romania ran in parallel. As the NSRF is on a higher aggregation level, not all OPs could mutually comply with the data structure required for the SPOT method. Therefore, the NSRF SPOT method had to be altered after a first version with regard to making up a common platform that embraces and represents all OPs in a comparable way. The methods that are more likely to use a sample of projects as an input basis (PCA) are obviously more sensitive (or in

other ways, flexible) to the composition of the sample. Therefore, though flexibility is a potentially positive characteristic of an evaluation method, an inadequate choice of sample may bring false, misleading or manipulative results, which is a definite risk of flexibility. As a consequence, it is crucial to highlight that the more flexible the method, the more cautiousness and more advanced user is required to use it in a proper and unbiased way.

Each quantitative methodology assessed has the flexibility in terms of both scope (length, depth of assessment), and economics of scale (in line with the resources available for evaluation). Therefore they are capable of acting as the methodological baselines for small-scale evaluations and adequate fundamental methodologies – in more sophisticated form – for more complex evaluations. The application of the selected mid-term evaluation methodologies in diverse countries showed that methods could have been tailored to the needs of the evaluation (e.g. specific evaluation questions), data availability (fewer stages identified in the process), country context (newly accessing country with limited understanding of certain notions), programme (relevance and detail variable) prior to their use, i.e. in the proposal or preparatory phases of the evaluation. In the course of the assessment of these methodologies, I could identify several benefits and limitations of their application. The fact that method parameters have been tailored to existing data on several occasions during the evaluation is a clear indication of their easy adaptability.

4 NEW SCIENTIFIC RESULTS

The *regional programme analysis* (assessment of the EU-12 regions, regional programmes and ROPs) yielded the following scientific results:

1. **Cohesion Policy promoted country-level development, however, it could not counteract growing interregional differences in the EU-12 countries.** Future interventions co-funded by Cohesion Policy instruments should be programmed and implemented with a strong intention and view to decrease the regional polarisation within the EU-12 countries. This includes both the planning of the adequate volume of funds capable of creating a critical mass, the concentration of resources and building interventions on regionally differentiating, competitive abilities and further improving the institutional system responsible for the effective and efficient delivery of the funds.
2. **The diverse approaches to ROP measures in the EU-12 have brought very similar allocation results.** Despite the fact that EU-12 ROPs are very diverse (based on relative weight within NSRF, country ROP approach, and sectoral OP vs. regional OP trade-off), the structure and choice of interventions within is fairly similar. This carries the message that bottom-up planning prevails centralised (EC level or member state level) planning structures, as interventions align to existing capacities and future expectations but still remain development need driven.

The *methodology analysis*, carried out on the 2007-13 NSRF and OP mid-term evaluations in EU-12 member states yielded the following scientific results:

3. **Publishing own methodology developments and testing them in international context (SPOT, LTA and PCA).** This paper includes the introduction and analysis of selected methods, which have been developed and continuously tested by myself with the contribution of experts from all over the EU-12 region in the last 10 years. The publication and self-assessment and cross-cutting international comparison of these methods is a valuable add-on to the methodology toolset available to programme evaluators and programmers in the new period as well.
4. **Methods used in 2007-13 mid-term programme evaluations can be transferred to other intervention levels, fields and sectors including non-EU programme evaluation.** The tools described in this paper are flexible enough to be capable of enhancing programme implementation and evaluation through better monitoring.

5. **Selected methods in 2007-13 mid-term programme evaluations have not been assessed before from a scientific perspective, therefore advantages and limitations have not been collected, categorised and synthesised in view of the experience gained through their multiple application.** My dissertation should facilitate understanding of the use of these methods with both commencers of evaluations and evaluators, which may result in more focused and more efficient evaluation activities. Also, continued use of these methods in the 2014-20 period may provide comparable results across EU programming periods.

5 PRACTICAL APPLICATION OF RESULTS

The scientific results of my research are easily and promptly adoptable and applicable for both potential and practicing evaluators and programmers or commencers of programme evaluations.

Evaluators may benefit directly from the results of my research through:

- 1. Description and analysis of methodologies:** Methodologies are described in this paper so that even those starting their first evaluation activities may understand how to use them. The international examples make methods even more tangible and understandable and demonstrate their objectivity, simplicity, robustness and flexibility and linkage to general intervention objectives. By collecting and sharing these methodologies in an analytic way, through the presentation of their application and their results in multiple countries, evaluators may use them more widely or with more competence.
- 2. Multiple examples of methodology application:** All methodology descriptions are followed by at least two country and context examples. Thus, use and expectable outcomes of the use of a certain method may be assessed. The international examples may provide inspirations on how to scale, alter, or further develop these methods.
- 3. Summary of methodology features, advantages and drawbacks:** By testing and summarising advantages and limitations of methodologies, evaluators may more consciously choose the selected methods. By closing the gap between expectations and reality, evaluators might be more aware of what they can expect as results from the use of these methods.

From another perspective, *programmers, commencers of evaluations and policy officers* may also benefit from my research through:

- 4. Better programming:** By understanding better the dynamics and limitations of evaluation activity, programmers may adjust programming exercises so that future evaluation activities may be simpler, requiring less capacity and expense. By sharing a helicopter view on ROP approaches and practices across the EU-12, policy makers and programmers might be inspired to experiment with different strategies in terms of approach and the trade-off of sectoral and regional programmes.

- 5. Better evaluation orders:** Commencers of evaluators might target their Terms of References and evaluation themes and evaluation questions even better by acknowledging what popular and widely used methods are capable of, and what their limitations are. This may help them set their level of expectations and frameworks of evaluations (i.e. deadlines, capacities, data sources available) even better. ToRs may become more attached to the actual methodology guidance being aware of the finding that many of the mid-term evaluation in 2007-13 was not commenced and as a result, was not conducted entirely in line with the EC guidance.
- 6. Raising awareness on international benchmarking and comparison:** It is an open competition in the EU to allocate and spend funds on objectives supported by the EU. This is similar in concept to a zero sum game, therefore, it is important to use the opportunity to learn from each other and develop. International comparisons in terms of programmes, evaluations, methods, practices or mechanisms provide a potential that has not been entirely tapped yet.
- 7. More focus on results than on performance:** Those commissioning evaluations and writing the Terms of References of evaluations might realise, that it is one of the most important purposes of mid-term evaluations to assess progress in terms of indicators and financials as well. Having determined current status, however, it is important to go into details with efficiency, bearing in mind the order. In “summative view”: first, what we have achieved and second, how we have achieved it, while in a “formative view”: first, what to do to proceed and second, how to enhance steps forward.
- 8. More focus on regional development:** Programme designers should bear in mind that the primary objective of Cohesion Policy and its funds and programmes is to mitigate regional inequalities. All means (Cohesion Policy co-funded programmes) should serve this purpose and all corresponding actions should be aligned to this overarching guiding principle.

6 REFERENCES

1. Bradley J. (2005): CEE National Development Planning. In: Bradley, J. et al. (ed.): Integration, Growth and Cohesion in an Enlarged European Union. New York, Springer, pp. 275-311.
2. Bradley, J. – Petrakos, G. – Traistaru, I. (2005): Integration, Growth and Cohesion in an Enlarged European Union: An Overview. In: Bradley, J. et al. (ed.): Integration, Growth and Cohesion in an Enlarged European Union. New York, Springer, pp. 1-25.
3. Canova, F., (2004): Testing for convergence clubs in income per capita: a predictive density approach”, International Economic Review 45, p.49-77.
4. Dvorková, K. (2014): Sigma Versus Beta-convergence in EU28: do they lead to different results? in Mathematical Methods in Finance and Business Administration. Department of European Integration VSB – Technical University of Ostrava
5. ESRI (1993): The Community Support Framework – Evaluation and Recommendations for the 1994-1997 Framework , Final Report to the Department of Finance, April 1993
6. Faragó, L. (2004): A regionalizmus hajtóerői Magyarországon. Tér és Társadalom, 3.sz. p.1–23.
7. Hirschman, A.O. (1968): Development Projects Observed. Washington, D.C.: Brookings Institution.
8. Horváth, Gy. – Lóránd B. (2012): Decentralizáció és gazdasági fejlődés – az olasz példa Közgazdasági Szemle, LIX. évf., 2012. december (p.1273–1298)
9. Horváth, Gy. (2001/a): Európai regionális politika. Budapest-Pécs, Dialóg Campus Kiadó.
10. Horváth, Gy. (2001/b): Regionális támogatások az Európai Unióban. Budapest, Osiris Kiadó.
11. Horváth, Gy. (2004/a): Regionális egyenlőtlenségek Európában. Magyar Tudomány, 165. évf., 9. sz.
12. Horváth, Gy. (2004/b): Európai regionális politika. Budapest, Dialóg Campus Kiadó
13. Horváth, Gy. (2005/a): Az uniós tag Magyarország új kohéziós politikájának alapkérdései. In: Sorsfordító esztendő. A 42. közgazdász-vándorgyűlés előadásai. Budapest, TAS -11 Kft., pp. 408-416.
14. Horváth, Gy. (2015): Spaces and places in Central and Eastern Europe : historical trends and perspectives of regional development. Routledge
15. Huey-tyh, C. (2005): Practical Program Evaluation: Assessing and Improving Planning, Implementation, and Effectiveness. SAGE Publications
16. Legge, J. (1983): The Chinese Classics (Vol. III), pp. 33–34, The Shoo King: Canon of Shun, Taipei, Southern Materials Centre Inc.
17. Jobbágy, Valér (2014): Az uniós források hazai végrehajtásának és végrehajtó rendszerének értékelése, javaslatok az új programozási ciklusra. Doktori értekezés.
18. Loránd, Balázs (2009): Konvergencia és fejlesztéspolitika az Európai Unióban és Magyarországon A területi differenciáltságot befolyásoló hazai fejlesztéspolitika értékelése az I. Nemzeti Fejlesztési Terv pályázati forrásainak allokációja alapján. Doktori disszertáció.

19. Louw, J. (1999): Improving practice through evaluation. In D. Donald, A. Dawes & J. Louw (Eds.), Addressing childhood adversity (pp. 60-73). Cape Town: David Philip.
20. Molle, W. (2007): European Cohesion Policy. Routledge, London-New York.
21. Nemes Nagy, J. (1987): A regionális gazdasági fejlődés összehasonlító vizsgálata. Akadémia Kiadó, Bp.
22. Nemes Nagy, J. (2005): Regionális elemzési módszerek. MTA-ELTE.
23. OECD (1986): Methods and Procedures in Aid Evaluation: A Compendium of Donor Practice and Experience
24. Pálné K. I. (2009): A politika új színtere a régió (Századvég, Pécs-Budapest, 2009. p.242)
25. Patton, M. (1997): Utilization-focused evaluation, London: Sage Publications, Inc.
26. Patton, Michael Quinn Qualitative Research & Evaluation Methods: Integrating Theory and Practice, Fourth Edition, SAGE Publications Inc. 2015.
27. Pollitt, C. – Summa H. (1996): Performance audit and evaluation: Similar tools, different relationships? New Directions for Evaluation. Wiley Periodicals, Inc. and the American Evaluation Association 1996/71 p. 29–50
28. Potter, C. (2006). Program Evaluation. In M. Terre Blanche, K. Durrheim & D. Painter (Eds.), Research in practice: Applied methods for the social sciences (2nd ed.) (pp. 410-428). Cape Town: UCT Press.
29. Rossi, Freeman, and Wright (1979): Evaluation: A systematic Approach. Sage Publications Inc.
30. Sanders, F. - Worthen (2004): Program evaluation: Alternative approaches and practical guidelines, 3rd edition. Boston: Allyn and Bacon (Pearson).
31. Scriven, M. (1991): Evaluation Thesaurus. Sage Publications.
32. Scriven, M. (1997): Truth and Objectivity Evaluation, Chapter 33 In Eleanor Chelimsky & William R. Shadish: Evaluation for the 21st Century: A Handbook. ISBN: 9781483348896
33. Terstiege, Stefan (2012): Objective versus Subjective Performance Evaluations. University of Bonn, Bonn Graduate School of Economics
34. UN (1951): Measures for the Economic Development of Under-developed Countries, New York
35. Weiss, C. H. (1998): Evaluation: Methods for Studying Programs & Policies 2nd edition. Prentice Hall. ISBN 013 3097 25 0
36. Williamson, J. (1965): Regional inequality and the process of national development. Economic Development and Cultural Change 14, 3-45.

Official documents

1. ACIS (2010): A Formative Evaluation of Structural Instruments in Romania, Final Report. July 2010. Ministry of Public Finance Authority for Coordination of Structural Instruments (ACIS)
2. EC (2006): Working Document No. 1: Indicative Guidelines on Evaluation Methods: Ex-Ante Evaluation European Commission. Directorate General Regional Policy
3. EC (2014): Sixth report on economic, social and territorial cohesion. 2014. July, EC DG Regio (adopted by the European Commission on 23rd July 2014)
4. Eurostat (2008): statistical database. <http://ec.europa.eu/eurostat>
5. MEANS (2000): Evaluating socio-economic programmes: Glossary of 300 concepts and technical terms Volume 6.. ISBN 92-828-6626-2 European Commission Luxembourg: Office for Official Publications of the European Communities.
6. OECD (2002): Glossary of Key Terms in Evaluation and Results Based Management. 2002. Organisation for Economic Co-operation and Development.
7. OECD (2008): DAC Principles for Evaluation of Development Assistance - Development Assistance Committee, OECD, 1991, 2008.
8. OPRD MA (2011): Mid-term evaluation of Operational Programme Regional Development (OPRD) 2007-13. Ministry of regional development and public works, Bulgaria
9. PPCD (2010a): Mid-term evaluation of OPI Malta. Planning & Priorities Coordination Department
10. PPCD (2010b): Mid-term evaluation of OPII Malta. Planning & Priorities Coordination Department
11. PPCD (2011): Extract from project implementation database cut-off date: 30.09.2010
12. SOP IEC MA (2010): A Formative Evaluation of Sectoral Operational Programme Improved Economic Competitiveness (SOP IEC) in Romania, Final Report. September 2010. Ministry of Public Finance Managing Authority (MA) for SOP IEC
13. Tavistock (2003): Sourcebook. Evaluating Socio Economic Development. Methods & Techniques. The Guide. Tavistock Institute
14. UMIS (2010): Unified Monitoring Information System (database cut-off date: 31.12.2010)
15. UNDP (2006): 'The Evaluation Policy of UNDP', UNDP Executive Board Document DP/2005/28, May 2006. <http://www.undp.org/eo/documents/Evaluation-Policy.pdf>.
16. UNDP (2009): Handbook on Planning, Monitoring and Evaluating For Development Results. United Nations Development Programme, NY.

7 PUBLICATIONS



UNIVERSITY OF DEBRECEN
UNIVERSITY AND NATIONAL LIBRARY



Candidate: Balázs Mezősi
Neptun ID: IR1R7L
Doctoral School: Kerpely Kálmán Doctoral School (Regional Sciences)

Registry number: DEENK/217/2015.PL
Subject: Ph.D. List of Publications

List of publications related to the dissertation

Foreign language international book(s) (2)

1. Várnai, É., Matolcsy, J., **Mezősi, B.**: EU funds in Central and Eastern Europe: Progress report 2007-08. [s.n.], [s.l.], 57 p., 2009.
2. Várnai, É., Matolcsy, J., **Mezősi, B.**: EU funds in Central and Eastern Europe. [s.n.], [s.l.], 57 p., 2008.

Hungarian scientific article(s) in Hungarian journal(s) (2)

3. **Mezősi B.**: Sok kicsi sokra megy.
KPMG Energ. Évkv. 2015, 84-85, 2015. ISSN: 2060-6818.
4. **Mezősi B.**, Árva L., Várhelyi T.: A kormányzati politika felelőssége a tőkebevonás terén.
Valóság. 48 (2), 1-10, 2005. ISSN: 0324-7228.

Foreign language scientific article(s) in Hungarian journal(s) (3)

5. **Mezősi, B.**: A practical means of setting the scene for evaluating aggregated absorption risk in EU co-funded programmes related to project failure.
Agrártud. Közl. 59, 63-66, 2014. ISSN: 1587-1282.
6. **Mezősi, B.**: A simple but powerful tool for the assessment of financial progress in EU co-funded grant programmes - Simple Progress Overview Tool (SPOT).
Agrártud. Közl. 59, 67-72, 2014. ISSN: 1587-1282.



Address: 1 Egyetem tér, Debrecen 4032, Hungary Postal address: Pf. 39. Debrecen 4010, Hungary
Tel.: +36 52 410 443 Fax: +36 52 512 900/63847 E-mail: publikaciok@lib.unideb.hu Web: www.lib.unideb.hu



7. **Mezősi, B.:** Application of lead time analysis as a method in EU co-funded mid-term programme evaluations in 2007-13: the case of Bulgaria and Malta.
Agrártud. Közl. 61, 9-14, 2014. ISSN: 1587-1282.

Foreign language conference proceeding(s) (1)

8. **Mezősi, B.,** Matolcsy, J.: A means for assessment of reasons behind progress in EU Structural Funds co-funded programmes: Simple Progress Overview Tool.
In: The 9th European Evaluation Society International Conference : Abstract Book. European Evaluation Society, [Prague], 170, 2010.

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of Web of Science, Scopus and Journal Citation Report (Impact Factor) databases.

12 October, 2015



