

Working Paper Series

Julie Pellegrin, Gelsomina Catalano

**WHAT LESSONS TO DRAW FROM MULTIPLE
REGIONAL CASE STUDIES:
BETWEEN COMPARABILITY AND SPECIFICITY**

WORKING PAPER N.02/2010

WHAT LESSONS TO DRAW FROM MULTIPLE REGIONAL CASE STUDIES: BETWEEN COMPARABILITY AND SPECIFICITY

Julie Pellegrin*, Gelsomina Catalano*

*CSIL – Centre for Industrial Studies (Milano)

This version: 12 January 2010

Abstract

This paper, presented at the Sixth European Conference on Evaluation of Cohesion Policy (Warsaw, 30 November-1 December 2009), addresses the specific case of regional case studies, i.e., case studies concerned with the effects of a SF programme implemented in a region. In the following, we draw on the concrete experience of the ex post evaluation of ERDF in 2000-06 (in particular Work package 4 “Structural Change and Globalisation” hereafter “WP4”) and other evaluations to review a set of principles that could help solve the dilemma between promoting specificity and making possible comparison and generalisation intrinsic to multiple case studies. The paper distinguishes three stages at which specific steps can be taken to ensure the final comparability and generalisation of findings: selection phase, implementation of the cases on the ground, and synthesis and generalisation. The intention is here to be deliberately concrete and useful, providing pragmatic solutions.

Keywords: Case study, Cohesion Policy

JEL codes: H72, O21, R58

1. Introduction

The realisation of case studies is a privileged methodological tool for evaluating Structural Funds (SF) interventions. As a matter of fact, case studies present two important advantages. First they are particularly appropriate for dealing with the complexity characterising SF interventions. Second, they enable to take into consideration the fundamentally different contexts in which SF interventions take place, i.e., Member States' regions characterised by very distinct socio-economic and historical features. Thus, multiple case studies make possible to carry out context-dependent analyses in the absence (or imperfect availability) of hard data.

Yet the recourse to multiple case studies is not without difficulties. In particular, the fundamental issue at stake has to do with how to synthesise the specific findings identified by the case studies and infer a coherent response to the evaluation questions. In other words: how to draw general conclusions out of specific cases?

Other difficulties (of an operational nature) materialise during the course of the realisation of case studies which are, in the last resort, related to the above issue. Indeed, this paper argues that, without neglecting techniques of synthesis, different steps should be taken at earlier stages of the evaluation process to ensure that the conditions are met for reaping the most of case studies. In particular, a carefully designed process of selection of the cases is a pre-condition for this. Also, many different devices and arrangements adopted during the development of the cases on the grounds are as – if not more – important in securing general findings that are both robust and useful.

2. Positioning of the case studies in the evaluation architecture

Multiple case studies can be adopted as a pivotal methodological instrument to carry out the evaluation of a Structural funds programme. It is worth insisting that even if the realisation of the case studies is at the core of the evaluation exercise, these case studies are positioned in a wider evaluation context. This means that the realisation of the evaluation study is not to be confounded with the realisation of the case studies; there are important steps taking place before and after the cases.

Before the case studies develop, a preliminary analysis must identify and / or justify the evaluation questions. An extended literature review is generally utilised in order to review the state of the art knowledge on the issue investigated on the basis of which working hypotheses can emerge. Even more ambitiously, a conceptual model can be developed on the basis of the literature review. This conceptual model is a logical system within which the working hypotheses are embedded. The literature review and conceptual model can possibly be complemented with a statistical analysis offering a first test of the working hypotheses before they are qualitatively addressed through the case studies. Overall, this first phase of the evaluation is crucial, not only because it identifies working hypotheses and evaluation questions addressed in the case studies, but also because it is a reference to get back to when drawing the conclusions of the evaluation study.

After the case studies have been realised, important stages take place that deal with the comparison of the cases, their synthesis and the generalisation of the findings (these issues will be addressed below). Overall, for the evaluation study to produce useful conclusions, the case studies must be correctly integrated in a wider theoretical and possibly statistical analysis.

In the case of WP4 a literature review on globalisation and structural change was used to elaborate a conceptual model composed of five working hypotheses about factors that determine the speed and direction of structural change: socio-economic change and human capital, regional specialisation, production system, innovation potential and internationalisation and relocation. A subsequent statistical analysis has confirmed the working hypotheses and identified four patterns of change characterising regions (locked in, slow adapter, fast adapters and shifters).

3. What is the basis for selecting cases, and how many?

When starting the phase of the evaluation study devoted to multiple regional case studies, the first concern has to do with the selection of the cases to be investigated. Which selection process to adopt, and in particular, which selection criteria both pertinent and realistic to use? The issue at stake is to pick up those regions that are representative of a category of regions sharing similar characteristics of interest with respect to the evaluation question.

Ideally, it is possible to resort to statistical data, relevant for the evaluation *problématique* and comparable across EU regions. A cluster analysis can be run in order to identify categories of regions sharing common characteristics relevant for the evaluation *problématique*. A multi-criteria analysis is then needed to select the regions within the identified categories that will eventually be the object of in depth case studies (see Box 1).

Box 1. Selection of regional case studies through a cluster analysis

First, a cluster analysis is intended to identify groups of regions sharing common characteristics in relation to the dimensions investigated as well as contextual realities faced by these regions. In principle, the identification of relevant indicators is done on the basis of a preliminary literature review scoping of the issues at stake, and possibly as a result, a specific conceptual model providing hypotheses in response to the evaluation question(s).

Such a cluster analysis needs to be complemented by a multi-criteria analysis based on “customised” qualitative sources of evidence in order to select the regions within the identified clusters that will eventually be the object of an in depth investigation.

The advantage of a cluster analysis is that it allows to derive natural groupings (or clusters) of observations and may be viewed as an exploratory data analysis technique. In other words, clustering methods may be intended largely for generating rather than testing hypotheses. In a cluster analysis, observations (in this case regions) are mutually replaceable with respect to the considered variables inside each group or cluster, even if the entities (regions) assigned to a group do not necessarily have all the same attributes. Within each cluster, entities are therefore similar. Compared to the descriptive analysis, which does not allow an easy handling of different variables simultaneously, a cluster analysis is useful to “classify” multivariate entities in few categories *a priori* not defined. It must be stressed, however, that, in a cluster analysis, much depends on the choice of the indicators selected to run the analysis.

There are several cluster analysis methodologies (for instance, hierarchical analysis or non-hierarchical analysis), and within each of these there are different methods. Furthermore, most methods allow a variety of distance measures for determining the similarity or dissimilarity between observations. The selection of the proper methodology is eventually based on the data characteristics.

Second, in order to select one region in each cluster, complementary and integrative criteria need to be considered through the realisation of a multi-criteria analysis. In the case of SF interventions, examples of criteria that can be utilised are the following:

- financial allocation;
- balance *between objectives*;
- balanced *territorial coverage* (e.g., between regions from the EU15 “old Member States” and the EU12 “new Member States”) ;
- other dimensions for typology of regions (e.g. urban/rural; core/periphery)
- *representation of the various types of management arrangements* associated with the delivery, conception and evaluation of Structural Fund programmes (e.g. central government versus devolved and autonomous regional administrations etc)

In the last resort, the choice of the criteria depends on the field of investigation of the evaluation. Overall, the analysis aims at:

- identifying EU regions where the evaluation *problématique* is particularly relevant;
- identifying EU regions where Cohesion Policy interventions were or could be relevant in the field investigated by the evaluation.

In fact, the realisation of a cluster analysis complemented with a multi-criteria analysis is not as objective as it seems and can rarely be carried out in such optimal conditions. First, it indeed needs to be stressed that the entire exercise depends to a high extent on the indicators eventually selected to run the analysis.

In addition, a number of difficulties related to the availability of indicators arise in the specific case of the evaluation of SF interventions. Fundamentally, the main obstacle has to do with the availability of statistics at the desired level of analysis. In general, comparable indicators available at NUTS 2 level offer a limited set of possibilities likely to be insufficient to cover the field of investigation of the evaluation. If the analysis needs to be cast at even lower level of analysis, e.g. NUTS 3 level as in the case of the evaluation of Objective 2 programmes over the 2000-06 programming period, the proposed path becomes next to impossible (beyond the selection process, the issue of the imperfect availability of indicators is also relevant for the development of the case study as quantitative indicators shall not represent an entirely adequate source of evidence – for a review of statistical data available at NUTS 3 level, see Annex).

In practice, in the absence of perfect indicators, additional considerations of a qualitative nature and *sui generis* arrangements need to be mobilised. These considerations can even literally substitute for the cluster analysis. Thus, rather than a cluster analysis completed by a multi-criteria analysis, the selection process can be based on a multi-criteria analysis alone. This is illustrated in the WP4 case below.

In the specific case of WP4, the process was facilitated by the prior availability of a taxonomy of Objective 2 eligible areas developed by another Work package of the ex post evaluation of ERDF 2000-06. The following indicators were utilised:

- *A first indicator provides information the scale of assistance received in terms of share of population living in the assisted areas. It was thus distinguished between four groups*

of regions: regions with over 45% (I), 20%-45% (II), 5%-20% (III) and less than 5% (IV) of the population living in assisted areas.

- *Another indicator distinguishes between the urban vs. rural nature of the assisted areas within the region. Five categories are identified: “predominantly urban”; “intermediate rural, close to a town”; “intermediate rural, remote”; “predominantly rural, close to a town”; “predominantly rural, remote”. In order to assign each region to only one category, these have been labelled with the type of area where the (relative, in some cases) majority of Objective 2 population lives.*
- *A last indicator accounts for ERDF expenditure per person living in assisted areas.*

Additional criteria have been adopted:

- *Objective 2 Programmes operating in regions in which the share of population assisted is less than 5% have not been chosen because they are not likely to provide sufficient evidence when assessing the effects of Objective 2 assistance on regional performance.*
- *Qualitative considerations on the economic structure of the NUTS 2 region where the Objective 2 Programme operates. In this respect, of great relevance was the preparation of self contained fiches of information about each candidate regions. Each fiche contained a series of quantitative indicators and qualitative considerations presented in a comparable way.*
- *The nature of the competition challenge faced by each Programme*
- *Consideration for the country to which the selected region belongs (opted for having two Programmes in each of the three major countries in terms of regions covered by the Objective 2 - UK, France and Germany - and one in each of the other countries)*

It should be stressed that the task was greatly simplified by the request made by the ToR to address 12 out of a list of 24 proposed regions (and to have at least one Programme per country). In turn, the selection of 24 regions was (probably) driven by considerations related to the need to have a balanced set of regions selected as case studies at the scale of the whole ex post evaluation endeavour, i.e., throughout all the different Work Packages.

Other difficulties might arise in the selection process, which are not necessarily related to the realisation of cluster analyses. A specific difficulty arises in the case of the evaluation of Objective 2 programmes prior to 2007 when the zoning principle was at work. The difficulty had to do with the imperfect match between the area in which the intervention being evaluated applies and the area covered by indicators (see box 2 for an illustration in the case of WP4).

Box 2. Correspondence between regions and programmes in the case of WP4

In the WP4 case, it soon appeared that the Objective 2 programmes being studied through case studies were deployed in areas identifiable at NUTS 3 that corresponded only imperfectly to wider NUTS 2 regions. The eligible zones were not always necessarily incorporated in the category immediately above, and if / when they were, they sometimes were characterised by drastic departure from the general picture of the region available at NUTS 2. In this way, it was impossible to use available indicators at NUTS 2 as a proxy applying to the areas at lower geographical level without incurring the serious risk of being utterly misleading. The ToR required to select twelve Objective 2 Programmes 2000-2006 that will be the object of in-depth case studies. In order to do so, the corresponding NUTS 2 region (in a few occasions, the corresponding NUTS 1 region – see below) has been firstly assigned to the each of the twenty-two Programmes indicated in the ToR. There were different possibilities. In the majority of cases, the programme's eligible areas are contained within a NUTS 2 region. In these cases, the choice of the corresponding NUTS 2 region was straightforward. In a few cases, the eligible areas were scattered across different NUTS 2 regions. The "North of Netherland" programme, for example, operates in three NUTS 2 regions belonging to the NUTS 1 Noord-Nederland, so the latter has been selected as the corresponding region. In the case of Multi-regional programmes, such as "Västra", the corresponding region eventually selected is the one where the majority of eligible areas is situated. Thus, the NUTS 2 region chosen for the Västra programme was Norra Mellansverige, with almost 600,000 people living in Objective 2 areas. Finally, in the specific cases of the UK, German and Dutch regions, the corresponding regions are at NUTS 1 level due to the direct correspondence between this level and the respective Programmes.

It is worth stressing an additional consideration related to the selection process, which again seems to be endowed with little scientific relevance but which is nevertheless relevant. Qualified experts with "inside view" and an in depth knowledge of the socio-economic and cultural context of the region (not to mention linguistic skills) are an important guarantee of the quality of the case study. When such expertise is for operational or logistic reason unavailable, it is probably better not to try and carry out the case study. This is not to say that regions and cases must be selected on the basis of the availability of such expert knowledge, but that the absence of it is a sufficient reason not to select the case.

These different elements intervening in the selection process give a fair idea of the mix between objective / quantitative and subjective / qualitative criteria adopted in order to achieve a satisfactory sample of selected regions.

Finally, one last question related to the selection process has to do with the number of cases to select. If cluster analyses were the way to select the cases, this is a non-question since the eventual number of cases is determined by the number of categories identified in the cluster analysis. However, as noted above, a proper cluster analysis can rarely be conducted without introducing ad hoc arrangements. As a result, the number of cases is often determined by extra considerations. In order to have a chance of being representative of the large quantity of regions and to account for their diversity, it can be considered that 8 cases is a minimum, while managing more than 12 cases starts to pose serious methodological difficulties at the later synthesis stage.

4. Data collection and organisation

Having a set of case studies that are representative of wider categories is an important prerequisite making possible to draw general conclusions from the cases. Subsequently, different arrangements should be adopted during the realisation process of the cases, that are as – if not more – important to secure the quality (usefulness and robustness) of the overall conclusions.

Organising data to allow meaningful comparisons: the central role of the template

One first element necessary to carry on the case studies is the elaboration of a *template* providing a common format to all the case studies. The template is in fact a crucial factor on which depends the success of the evaluation exercise since it is the place where the dilemma between comparability and specificity can be best managed. As a matter of fact, the template serves to define a common structure making possible to *compare* the content of the cases. At the same time, it should also provide sufficient room for manoeuvre to the case experts for addressing the specificities of his/her case. A good template should therefore constitute a flexible framework for collecting and organising data which are by definition very specific and context-dependent in such a way that meaningful comparisons can be inferred later from the case studies taken together. There are different solutions in order to successfully strike the delicate balance between comparability and specificity. First, a set of working hypotheses can be proposed that country experts can “customise” by selecting the most pertinent in his/her case or by adapting them. Second, statistical data can be used to establish a common component in all the cases. Comparable statistics can be provided by the core team that are then complemented by statistics from local sources available to the country experts. Third, continuous interactions between the core team guaranteeing consistency across the cases, and country experts bringing forth the specificity of the context characterising his/her case should be ensured.

In the specific case of WP4, the following arrangements were adopted in order to reconcile the objectives of specificity and comparability:

- *A set of working hypotheses were elaborated beforehand (on the basis of a conceptual model itself resulting from an extensive literature review and a statistical analysis identifying preliminary and aggregate patterns characterising European regions). There were five of these hypotheses and the case experts were urged to select the most appropriate one(s) to be discussed in the case of the region considered. In this way, while all the case studies were related to a single conceptual model, they could focus on specific aspects of the conceptual model deemed to be particularly relevant.*
- *A core of comparable statistical indicators was provided to the case experts. The latter were invited to pick those statistics applying to his/her case and pertinent to the specific problématique developed in the case. Case experts were also expected to complement these to the extent possible by statistical data available from local sources (and therefore in principle non comparable). In addition, the case experts were expected to present the case’s statistics by placing them in the perspective of “comparator” available from other regions in the same country and to place the regional position and trend in the perspective of the EU average*

- *Indications were integrated directly in the template providing case experts with specific and clear suggestions on how to proceed. These indications were very mild, they actually were presented as “suggestions” to follow and options were open (the tone was carefully studied not to be commanding)*
- *A long list of possible sources of evidence was proposed to leave case experts the choice of the most appropriate references.*
- *Similarly, a set of methodologies were identified as being adequate to the evaluation study (survey, semi-directive interviews, mini case studies..)*
- *Generic interview guides (for both policy makers and beneficiaries – see below) were distributed.*
- *The very process of filling the template was organised so that case experts could adjust their work to the requirements of the core team. First, a series of milestones and intermediate deliverables were fixed. On these occasions, case experts were given the opportunity to receive feedback from the core team. Second, interactions between the core team and the case experts were expected to be continuous and iterative. Due to time pressure (and limited budget), the second way seemed to work better.*
- *A useful web platform was activated where the different case studies were progressively uploaded, together with all deliverables necessary for the compilation of the cases, and a case study “toolkit” comprising a methodological guidance note). This served to have the information circulate.*
- *The template was tested through the realisation of two case studies, and adjusted accordingly to the experience recorded.*

What’s in the template: narrative and evidence based analyses

Other important principles can be followed which do not deal directly with the issue of solving the specificity / comparability dilemma but which increase the general quality of single cases (and thus contribute indirectly to the dilemma)

First, the stress should be placed on the narrative character of the cases. Case studies are expected to “tell a story” both coherent and pertinent for the working hypotheses and / or the evaluation question(s). This story is a way to bring together different contextual features which can be considered to be details from the outside but which are in fact relevant explanatory factors. In this respect, a historical perspective might be particularly useful to fully account for the regional specificity (as testified in the case of WP4). In this way, the context-specificity of the cases is preserved and even furthered. A related issue is that the storyline should flow throughout the text, without coming against the interruptions of the different sections composing the template. In other words, the mechanical filling of the single sections, one by one and independently from the others is to be avoided. On the contrary, case experts are expected to take advantage of the different sections to let the story develops. This approach also accounts for the fact that the expected length is not to exceed a critical threshold (for example 50 pages).

Another important factor determining the quality of a case study – its usefulness – has to do with the quality of the evidence on the basis of which the case develops its analyses. A good case

study needs to be first of all strongly evidence-based. The raison d'être of case studies (and added value with respect to other methodological tools) is indeed to collect field evidence of primary source otherwise out of reach. In general, for logistical reasons (time and budget constraint), such evidence is prevalently of a qualitative nature but surely, quantitative data of primary source (e.g. survey – see below) represents a plus.

Two sources of field evidence are managing authorities implementing the programme considered and beneficiaries of the interventions examined. While the former are more easily accessible and therefore commonly included in fieldwork, the later should not be excluded. On the contrary, an important element to increase the usefulness of the findings is precisely to consider evidence concerning beneficiaries in order to indentify the effects of the interventions examined on those who are the primary targets. Such evidence – from local managing authorities and from beneficiaries – can be collected in various ways, for example through semi-directive interviews, focus groups, or more ambitiously through surveys. In addition, it is useful to enrich the text with different “exemplars” i.e., brief descriptions of projects and corresponding beneficiaries.

Quality control

Finally, one more principle contributing to the overall quality of a case study has to do with the very modality through which it develops. Perhaps more than for other methodologies, multiple case studies require that a strict quality control system be set up. Because of the nature of the evidence utilised, different checks of the accuracy of the sources utilised are useful. One first possible check can be realised by the country experts, going back to stakeholders to test the case's conclusions – although this is probably time consuming and not entirely reliable since the country experts may lack the necessary distance. Indeed, the involvement of more distant observers is probably a best guarantee. Since much depends on the interpretation of the country experts, a rich interaction between the latter and the core team is a good way to ensure that the main messages from the cases are correctly conveyed. Overall, each case needs to be reviewed by external experts from the core team and/or from outside. Interestingly, as noted above, such interactions are also useful to secure consistency (and hence comparability) between the different cases. Overall, different versions of the case studies are probably necessary to reach the desired level of quality and accuracy.

5. Synthesis and generalisation

Bringing together the huge amount of evidence, data and information collected through the cases is probably the most challenging part of the evaluation exercise. Different elements should be kept in mind to help overcome the difficulties. For a start, it should be clear that the case studies are no single evaluation of the programmes considered. Rather, the evaluation judgement must take place on the basis of the different case studies analysed together. This implies, for example, that lessons learned are to be drawn on the basis of a horizontal reading of the different cases and not at the level of the single cases. Such horizontal reading of the cases – or their comparison – is not the end of the evaluation study, however. The patterns revealed through comparisons must be tested in other circumstances, outside the case studies: in this way only can findings of general reach be proposed.

Possible tools to obtain meaningful comparisons: synoptic tables

Such a horizontal reading of the cases is amply facilitated by the compilation of one – or several – synoptic tables. These tables constitute a methodological tool expected to facilitate the comparison between the cases. They present, for each region, a series of key features drawn from the case studies reports that present evidence relevant with respect to the working hypotheses / evaluation questions; they are synthesised so as to fit in the table(s)' cells. The tables cross these features with findings concerning the respective effects of Structural funds regional programmes. Comparing and crossing evidence is expected to identify patterns characterising sub-groups of cases. For example, under different circumstances, region X, Y and Z adopted similar policy (SF) strategies in response to a certain phenomenon (the topic object of the evaluation) yet obtained apparently different outcomes (different degrees of effectiveness). The hypothesis is that these circumstances play a role in determining SF effectiveness. This sort of propositions constitutes the basis on which rudiments of causal relations can in principle be inferred. Thus, these grids form the basis on which the formulation of an evaluation judgement can take place: links are established between different variables or factors and relations of causality are possibly revealed out of a complex assembly of elements of information with no apparent relation.

An alternative to the synoptic tables is the compilation of synthetic fiches for each region all structured around the same key features (the same as those for the table). The difference between the fiches and the table is just a question of presentation.

In the specific case of WP4, synoptic tables crossing the case studies with patterns of structural changes, policy responses and key findings have been produced in order to allow a horizontal reading of the cases. These tables constituted a methodological basis for drafting the main sections of the Final Report by providing field evidence. Table 1 provides an example of synoptic table used in the project.

Table 1. Analytical grid for establishing a typology of regional patterns and policy responses (WP4)

Structural change pattern	Working hypothesis	Policy mix: ERDF / national regional policy	Policy response: ERDF			ERDF measures	ERDF effectiveness	Policy learning
			Type 1	Type 2	Type 3			
	Region							
Locked-in	A							
	B							
	C							
Shifters	D							
...								

What generalisation is possible?

It has to be stressed that at this stage, the results obtained are valid only for the cases considered. In other words, the conclusions drawn from the above grids of (causation) relations are useful to identify the advantages / shortcomings / effectiveness of a specific intervention as opposed to another – and in a certain context. However, this is not enough to draw *general* conclusions on other interventions of the same type, in other contexts.

In order to bring the evaluation exercise to its concluding phase, it is therefore necessary to make sure that the results obtained can be *generalised*. The principle is simple: the identified relations need to be verified in other a priori similar situations (e.g., same interventions, similar context). A two legged approach is possible. This verification can be done through qualitative means, for example a review of documentary evidence concerning a certain number of regions, which are not those studied in the cases. A set of “policy effect checklists” can be prepared, systemising different types of policy response and policy effects in non-studied regions. Alternatively, a more participative approach (adopted in the WP4 case) consists in organising a workshop with participants coming from regions which were the object of the case studies and from other regions, the objective being to have them confront their experience and test the conclusions reached.

However, for the test to be accurate, it cannot be entirely random. A more systematic verification of the case studies conclusions can be done on the basis of a statistical analysis establishing a comprehensive taxonomy of regions. The qualitative test is organised by category of the taxonomy. It is worth noting that this taxonomy can or should be the same as that utilised at the selection stage (this of course is only possible when the required statistical indicators are available).

Precisely because of the difficulties to obtain such a taxonomy noted above, it is worth insisting that this later stage could be imperfect, incomplete or impossible. In this case, it is probably wiser to take the conclusions reached on the basis of the cases for good and acknowledge that the exercise cannot go beyond the identification of a set of explanatory factors valid in the specific cases considered. This is already an achievement given the complexity of the situations explored.

6. Concluding remarks

Overall, multiple regional case studies can be a powerful methodological tool to deal with, and even account for, the complexity of SF interventions in European regions, provided a few principles are respected. In order to make the most of the specific analyses they can provide while ensuring that the findings are comparable and can be generalised to other circumstances, two broad series of measures can be taken at different stages of the evaluation process. First, an appropriate set of cases (here regions) must be selected that are representative of wider categories (of regions). This is no sufficient condition but it is a necessary step making possible future meaningful generalisation. Second, different arrangements should be taken while implementing the data collection process. The elaboration of a template providing a common framework of analysis but at the same time leaving sufficient room for specific analyses is fundamental in this respect. Another important means consists in adopting a smooth communication and interactions framework between the different members of the evaluation team (in particular between the core team and

case experts, but also between case experts). As the information circulates, benchmarks are diffused allowing to model each specific case on a broad comparable pattern. Other important factors determining the quality of case study is their narrative aptitude –being able to bring together different apparently unconnected variables to account for a specific “story” at work in the region considered, and their being strongly evidence-based (mostly relying on primary qualitative evidence, but if possible also on primary quantitative data).

Once a set of regional case studies completed, a methodological tool like synoptic tables helps to compare the findings and formulate hypotheses on possible explanatory factors. Finally, the generalisation of findings must conclude the evaluation exercise. Without such a generalisation, the evaluation would be incomplete. It ideally lies on the basis of a statistical categorisation of the regions and consists in testing the relations identified in a case study region in other regions of the same category. It is worth stressing, however, that in the face of inevitably imperfect generalisation process, the set of explanatory factors and (causation) relations identified in each single case is already an interesting achievement in the field of SF evaluation.

Annex

Table A 1. Selected socio-economic indicators available at NUTS 3 level

INDICATOR	DESCRIPTION	YEARS	NUTS LEVEL
Gross Domestic Product per capita PPP	Gross Domestic Product per inhabitant	1999-2005	2,3
Rate of unemployment	Ratio between number of unemployed persons and number of people aged 15 and over	1999-2005	2,3
Long-term unemployment rate	Long-term unemployment (12 months or more) as % of total unemployment	1999-2005	3
R&D personnel in all sectors	R&D personnel in all sectors as % of total employment	1999-2005	3
Share of Agriculture employment on total	Employment in Agriculture / Total employment	1995-2006	2,3
Share of Industry employment on total	Employment in Industry / Total employment	1995-2006	2,3
Share of Services employment on total	Employment in Services / Total employment	1995-2006	2,3
Gross Value Added in Agriculture, hunting, forestry and fishing	GVA generated by Agriculture, hunting, forestry and fishing sector as % of total GVA	1995-2005	2,3
Gross Value Added in Industry	GVA generated by Industry sector as % of total GVA	1995-2005	2,3
Gross Value Added in Services	GVA generated by Services sector as % of total GVA	1999-2005	2,3

INDICATOR	DESCRIPTION	YEARS	NUTS LEVEL
Gross Value Added in Mining and quarrying; electricity, gas and water supply	GVA generated by Mining and quarrying; electricity, gas and water supply sectors as % of total GVA	1999-2005	2,3
Gross Value Added in Construction	GVA generated by Construction sector as % of total GVA	1999-2005	2,3
Gross Value Added in wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods, hotels and restaurants; transport, storage and communication	GVA generated by wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods, hotels and restaurants; transport, storage and communication sector as % of total GVA	1999-2005	2,3
Gross Value Added in Financial intermediation; real estate, renting and business activities	GVA generated by Financial intermediation; real estate, renting and business activities sector as % of total GVA	1999-2005	2,3
Gross Value Added in Public administration and defence, compulsory social security; education; health and social work; other community, social and personal service activities; private households with employed persons	GVA generated by Public administration and defence, compulsory social security; education; health and social work; other community, social and personal service activities; private households with employed persons sector as % of total GVA	1999-2005	2,3

Source: EUROSTAT