REGIONAL SCIENCE, TECHNOLOGY AND INNOVATION POLICY
EVALUATION: APPROACHES AND METHODS

Drs. Jon Mikel Zabala Iturriagagoitia (INGENIO, CSIC-UPV)
Dr. Fernando Jiménez Sáez (INGENIO, CSIC-UPV)

Corresponding address: jonzait@ingenio.upv.es
Postal address: INGENIO (CSIC-UPV); Universidad Politécnica de Valencia; Camino de Vera s/n; Edificio 9B; 46022 Valencia (SPAIN)

Theme 2: Renewal of Innovation Policy
Ref: A 95

1 ABSTRACT

Policy evaluation is a key issue in the policy cycle, despite it has not been deserved a careful attention, specially with respect to ex-post policy evaluation in peripheral European regions such as Valencia. Some evaluation methodologies seem to be useful in some cases but not in others, where the availability of data and infrastructures to carry out evaluation, lags clearly behind. This rises the issue of how to compare the results in the design and implementation of regional S&T policies across regions in Europe.

Since evaluation, in any case, implies the measurement of some results and its comparison to a certain scale previously established, the research arena is opened to enquire on the design and implementation of comparable evaluation methodologies with relevant policy implications to be considered in the European Research Area.

Many methodologies, are inspired on the design and development of indicators used to capture direct results of policy implementation. Others are inspired on the consideration of additionality as the main result pursued with the application of a policy. Here, we propose to study the evaluation of regional S&T policies regarding the efficiency of such policies.

In this case, a study concerning the evaluation of the S&T policies implemented in the Spanish regions according to their efficiency levels will be a matter of interest. In order to achieve that goal the Data Envelopment Analysis (DEA) will be applied due to the advantages it offers in public studies.

2 INTRODUCTION

The consideration of the region as a geographical space with specific and different characteristics in terms of institutional, regulatory and legal frameworks to those of the nation (Cooke et al., 2000), have posed many concerns in the design,
implementation and evaluation of Regional Science, Technology and Innovation (STI) policies. This fact together with the current importance given to innovation and knowledge creation in the development of a geographical space must drive the design and application of ex-post evaluation methods to these policies.

Currently, the systemic approach to technological change is posing large attention on the complexity of relationships among agents. Evaluation methodologies, traditionally concerned with the accurate use of resources vs. results in terms of inputs, outputs and outcomes, are not ignorant of this new challenge posed by the design of policies aimed at promoting R&D and innovation activities under a systemic approach (Dietz, 2003; Luukkonen, 1998). This fact becomes specially remarkable within the limits of a geographical space with its own idiosyncrasy. Under this approach similar regional policies aimed at promoting the same goal in two different regions may yield rather different outputs and outcomes to the extent of being considered contradictory policies in terms of input vs. output evaluation. Similarly a nationwide policy instrument may lead to rather different outputs in different regions due to some extent to those special characteristics that differ from one region to another. In other words, we observe the design of STI policies under a systemic approach closely linked to an interactive innovation model view (Kline and Rosenberg, 1986) yet almost all evaluation methodologies still are anchored in approaches inspired by a linear-fashion view of innovation (Molas-Gallart and Davies, 2005) without any further consideration of the region’s systemic characteristics. As a consequence the evaluation results may lead to misleading conclusions and do not provide a solid ground for accurate policy assessments. Therefore the evaluation of regional STI policies (and instruments) as a discipline has to progress toward a systemic view where inputs, outputs, outcomes and (or) impacts are considered as well as agents and relationships all within the boundaries of the (regional) Innovation System.

The aim of this paper is to provide a picture of the evolution and current methodologies and practices used in the evaluation of instruments and tools addressed to the promotion of R&D and innovation activities at the regional level. Besides, we apply a currently very fashionable methodology based on an efficiency measure of input-output comparison to illustrate how these methodologies by themselves are just another linear-base method without any further consideration of the innovation system’s relationships and collaborations. Therefore we claim that their results have to be complemented with additional information on the System’s characteristics in order to offer policymakers a solid ground to evaluate a regional STI policy. We apply this methodology on a regional policy designed to promote firm incubator at the regional
level. In order to reach this goal we first will show a short description in the evolution of STI policy evaluation practices and their application in the EU, different EU member states and regions. Afterwards we will focus the attention on the advantages and drawbacks of both quantitative and qualitative evaluation methodologies in their application. Thirdly we describe how a methodology based on inputs and outputs measures to provide an efficiency ranking has to be combined with the region’s systemic characteristics to offer the evaluation of a regional STI policy instrument. Finally some conclusions are drawn.

### 3 STI Policy Evaluation to Review

Policymakers, scholars and even citizens might agree on the role of policy evaluation as the way to determine whether public resources have been accurately used to obtain a certain amount of outputs. If we pay attention on R&D and innovation activities, the definition of the role of STI policy evaluation becomes more controversial due to the special characteristics of R&D returns measurement (Gibbons and Georghiou, 1987). In fact STI policy evaluation is a relatively new practice compared to other fields where evaluation had become a common practice by the early 1980s (like for instance educational policy; Scriven, 1973). The European Commission has been involved in RTD policy evaluation from late 1970s (Guy and Arnold, 1998) and approaches, methodologies and structures have also been evolving accordingly. Institutionalisation of evaluation practices accelerated in the 1980s when a centralised evaluation unit was set up. During those dates were carried out the evaluation of EU major programmes (for instance ESPRIT in 1985, EUREKA from 1991 onward). Another milestone in the EU RTD policy evaluation history came with the launch of the SPEAR Programme (1989-1992) supporting research on evaluation methodologies and indicators. As a consequence we observe an increasing concern about the utilisation of research results, the measure of socio-economic benefits of R&D programmes and their results’ application to solve society needs. Several studies focused on efficacy of evaluation procedures (Chabbal, 1988; PREST, 1990; Krull et al., 1991). Nowadays the EU RTD policy evaluation has become a sounded practice and to some extent has contributed to the development of such practices among the EU member states.

In that respect STI policy evaluation as a practice has experienced a different pace in each country. Among EU member states with larger tradition in STI policy evaluation we can find the UK, Germany, France or Austria, where many of the current
evaluation practices and methodologies have been developed whereas other countries like Spain, Portugal or Greece have not been so active in the development of evaluation methods or even suitable structures for their own STI policies. This fact in the end has meant the adoption of practices and methodologies by these countries without any further concern on the appropriateness and adaptability to the specific case.

The evolution of STI policy evaluation at the regional level has been following a parallel history to that of the respective nation but during the last years this evolution has experienced a divergent path due to the emergence of new regional development theories. The concept of endogenous development has now given way to new concepts and ideas such as the “learning region” (Morgan, 1997; Landabaso, 2000). As a consequence, regions are increasingly important sources of innovation and economic growth. This new approach to the region devotes a larger attention to networks, clusters and regional innovation systems as key elements for their development. This increasing role of the region is rooted in what Storper named “untraded interdependencies” and that “take the form of conventions, informal rules and habits that coordinate economic actors under conditions of uncertainty” (Storper, 1997; p.5). Therefore current regional STI policies involve the consideration of different types of issues: the presence of intangible objectives, complexity of cause-effect relationships, systemic nature, embeddedness, dynamism and the devolution of power to the region (Díez, 2002). They have meant the consideration of new evaluation practices and methodologies since traditional ones do not seem to provide suitable guidance to policymakers in terms of a holistic evaluation result. There is also the tension between new approaches to STI policy evaluation considering these new issues into the methodological approach and the policymakers and evaluation clients' requirements of more traditional “ready-for-use” indicators.

4 STI Policy Evaluation Methodologies

The increasing consideration of evaluation as part of the policy cycle among policymakers has undoubtedly influenced the development of new evaluation approaches and methods. From the early application of quantitative methods based on an input-output comparison to a more qualitative approaches based on the case study methodology and the combination of both methods we can find a large collection of methodologies with the objective of providing outputs, outcomes and impact measurements. The challenges posed by the new regional development approaches
are also influencing on further developments of those methodologies in order to incorporate measures and indicators able to gauge and assess the impact of STI policies. Table 1 is a summary of those measures and the methodology in each case.

<table>
<thead>
<tr>
<th>METHODOLOGY</th>
<th>OUTPUT</th>
<th>OUTCOME</th>
<th>IMPACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Innovation Surveys</td>
<td>• New products and processes&lt;br&gt;• Increase in sales&lt;br&gt;• Increase in value added&lt;br&gt;• Patents counts&lt;br&gt;• IPRs</td>
<td>• Creation of new jobs&lt;br&gt;• Innovation capacity building</td>
<td>• Enhanced competitiveness&lt;br&gt;• Institutional and organisational efficiency&lt;br&gt;• Faster diffusion of innovation&lt;br&gt;• Employment</td>
</tr>
<tr>
<td>Micro methods</td>
<td>• Output and value added (collect baseline info for before-after comparisons)</td>
<td>• Sectoral productivity&lt;br&gt;• Industry Sectoral spillovers&lt;br&gt;• Additionality&lt;br&gt;• Leverage effects</td>
<td>• Firms’ competitiveness</td>
</tr>
<tr>
<td>Macro methods</td>
<td>• Output and value added</td>
<td>• Change in R&amp;D capital&lt;br&gt;• Human capital&lt;br&gt;• Social capital&lt;br&gt;• International R&amp;D spillovers</td>
<td>• Regional / Country productivity&lt;br&gt;• Employment&lt;br&gt;• Good governance&lt;br&gt;• Economic and social cohesion</td>
</tr>
<tr>
<td>Productivity Studies</td>
<td>• Output and value added</td>
<td>• Knowledge&lt;br&gt;• Geographical and international R&amp;D spillovers</td>
<td>• Regional / Country productivity&lt;br&gt;• Employment&lt;br&gt;• Economic and social cohesion</td>
</tr>
<tr>
<td>Control group approaches</td>
<td>• Output and value added (on supported and non-supported firms)</td>
<td>• Additionality&lt;br&gt;• Rate of return to R&amp;D firms</td>
<td>• Firms’ industrial competitiveness</td>
</tr>
<tr>
<td>Cost-Benefit Analysis</td>
<td>• Value added&lt;br&gt;• Benefit cost ratio&lt;br&gt;• Consumer surplus</td>
<td>• Health improvements&lt;br&gt;• Consumer protection&lt;br&gt;• Environmental sustainability</td>
<td>• Quality of life&lt;br&gt;• Standard of living</td>
</tr>
<tr>
<td>Expert panels / Peer review</td>
<td>• Publication counts&lt;br&gt;• Technological outputs</td>
<td>• Scientific and Technological capabilities</td>
<td>• R&amp;D performance</td>
</tr>
<tr>
<td>Field / Case Studies</td>
<td>• Detailed inputs and outputs</td>
<td>• Firms’ RTD capabilities&lt;br&gt;• On-the-job training&lt;br&gt;• Educational schemes</td>
<td>• Industrial competitiveness&lt;br&gt;• Quality of life&lt;br&gt;• Organisational efficiency</td>
</tr>
<tr>
<td>Network analysis</td>
<td>• Cooperation linkages</td>
<td>• Cooperation in clusters&lt;br&gt;• Social embeddedness</td>
<td>• Efficiency of institutional relationships</td>
</tr>
<tr>
<td>Foresight / Technology assessment</td>
<td>• Identification of generic technologies&lt;br&gt;• Date of implementation</td>
<td>• Technological capabilities</td>
<td>• Technological paradigm shifts</td>
</tr>
<tr>
<td>Benchmarking</td>
<td>• S&amp;T indicators&lt;br&gt;• Technological capabilities</td>
<td></td>
<td>• Industry competitiveness&lt;br&gt;• Good governance</td>
</tr>
</tbody>
</table>

Source: Tavistock Institute et al., 2003

This description of methodologies and measures arises a double question: on the one hand we need to know whether these methods do incorporate a systemic consideration of the programme, action or instrument being evaluated or still are based on measures designed just to contrast outputs, outcomes or impacts with the inputs provided by the responsible administration following therefore a linear consideration of the evaluated activity. On the other, to what extent do these methodologies and measures provide guidance with respect to current regional STI policy evaluation.
requirements? If we examine carefully the design of regional STI policies we can realise that their design (in some cases) is inspired in a coupling or interactive consideration of the innovation process. However, evaluation methodologies are mainly based on a linear conception of the innovation process. But we still can find the second lack of coherence in the evaluation of regional STI policies since their designs are introducing the new approaches to regional economic development but we observe the application of traditional evaluation methodologies which do not take into consideration the region as a peculiar geographical space.

Coming back to the evolution of evaluation methodology we can trace the origins of regional policy evaluation across Europe by the 1990s when cost-benefit approach was the main evaluation methodology applied. Here the discussion centres on whether this approach can accurately calculate costs and benefits when valuing for instance intangible benefits (Georghiou, 1998). Afterwards, more qualitative and “political” evaluation methodologies were preferred (Hamblenton and Thomas, 1995). Many of these ideas about regional policy evaluation are already being transferred to evaluation practice where new focuses are being tried out for the evaluation of regional STI policies (Díez, 2002). Related approaches have been designed and could be grouped under the designation of participatory evaluation (Díez, 2001). They all provide an evaluation focus committed to the development of a change or improvement that is interactive, contextualised and directed to knowledge building. These techniques do not impose the design from outside; it takes shape through the collaboration of all stakeholders and their active participation in the evaluation process. This approach offers an interesting base for its application on regional STI policies since it makes possible to convert evaluation into an exercise contributing to achieving the goals of the so-called new regional policies. Kuhlmann stressed how the various interests and perceptions of the actions taking part in an action or policy instrument must be explicitly taken into account (Kuhlmann, 1998). Therefore it seems plausible the application of these techniques on regional STI policies.

However, nowadays the quantitative approach to evaluation has recovered a new boost. New methods based on efficiency measures provide a rather attractive result in terms of a ranking. This result is always very well received among politicians who always need convincing arguments to support or not the actions implemented by each policy. If we examine in detail the abovementioned methodologies we can realise that, without exception, all of them can be applied in some cases and cannot in others. Therefore no single evaluation method can serve as a methodological recipe applicable to each and every evaluation. In other words, there does not exist an ideal
methodological design and each situation, policy or instrument requires a unique specific evaluation design.

As a matter of fact we are facing the problem of applying this or that evaluation methodology, but new regional STI policies do not pay sufficient attention to designing mechanisms and structures that allow latter evaluation of these policies. Consequently, we should also consider whether the problem is just to find the right evaluation methodology to each policy level (regional, national…) and each policy context (inspired under a linear or an interactive innovation process model), or integrate the evaluation process into the policy cycle so every policy evaluation constrains are taken into considerations from the very beginning of the policy design.

5 Efficiency in the Evaluation of a Regional STI Policy

Policy measures might have multiple effects, being very difficult to show or evaluate their overall impact on innovative activities (Grupp and Mogee, 2004). In this regard, National Innovation Systems (NIS) were established as an object of planning, a comprehensive entity whose efficiency should be systematically evaluated (Miettinen, 2002). Due to the need of making IS more competitive, evaluation is considered as a central means of making the systems be efficient. As a matter of fact, the Science and Technology Council of Finland, does consider the evaluation of its NIS’ efficiency as a clue for the development of innovation in the country (Science and Technology Council of Finland, 1993).

When a national economy develops exceptionally quickly, there is a tendency to regard that country’s policy as an exemplary model that can be used as a “benchmark” (Diewert and Nakamura, 1999). When studying these successfully developed policies, it is a common practice to focus ones attention on the patterns of the recent past instead of the historical factors that might explain the success better than the policy pursued in last years. In addition to it, due to the existing differences among territories, not only the development of new policies but also their evaluation needs to focus more on the country/region to be developed rather than the one to be observed (Hadjimanolis and Dickson, 2001; Koh et al., 2005). All in all, these differences make evaluation oriented methodologies be more difficult in their application since involve the need to evaluate each policy one by one and not to make comparisons. Hence, the policy evaluation constitutes a key point in the policy cycle (Mytelka and Smith, 2002; Schwerin and Werker, 2003) in order to improve the impact and development of
Science and Technology in these countries/regions without benchmarking the policies implemented by those who succeeded in the past.

In this regard, it becomes a really interesting issue to determine the (in)efficiency of the policies (Hjerppe and Mäkelä, 1998; Glass et al., 2005), since it is commonly argued that the efficiency of an Innovation System determines its competitiveness (Science and Technology Council of Finland, 1993: 9-10).

The rhetoric of NIS served to integrate policy areas and find new basics for the traditional policy measures (Archibugui, 1999). However, it remains an open question as to how and to what extent the current policy making (and consequently evaluating) practices have been adapted to this new systemic language (Jääskeläinen, 2001: 217-219).

This is one of the main goals of this research. It will test whether the efficiency of the public (innovation) policies plays a relevant role in the behaviour of the (Innovation) system and helps to its understanding (Desmet et al., 2004). In this sense, it is aimed at evaluating public (Science and Technology) policies by means of the efficiency of the several instruments employed in time for its development. In this case and for Spanish regions, the creation of new university spin-offs will be the program to be evaluated (Isaksen, 1999).

In the case of Finland, the research on cluster programs (Pentikäinen, 2000) and regional collaborative projects based on EU structural funds (Kuitunen and Oksanen, 2001) has revealed an increasing collaboration among public players. However, it is hesitant about whether this collaboration has increased innovative activity or not, due to the increasing bureaucracy and the large number of actors involved, which have proven to be serious obstacles to the efficient implementation of these programs. The same way the study carried out in the ICT sector (Susiluoto, 2003) has been revealed as a relevant study in this sense, studying whether the ICT sector makes the efficiency of regions increase or not. In terms of such a system, the activities of diverse institutions are evaluated from the point of view of their contribution to national economic competition based on the development of high-technology products.

Mowery and Ziedonis (1998) show the difficulties of applying any new innovation policy intended to improve systemic performance. Despite the efficiency of the system (and consequently the one of the policy) is considered as a key factor in the performance, the lack of reliable data in this level of aggregation makes complex to assess policy makers in this regard.
Nasierowski and Arcelus (2003) prove the relation between the efficiency and the productivity of a country’s R&D effort, within the context of its National Innovation System. In this context, efficiency relates to the ability to transform R&D inputs into R&D outputs.

Generally speaking, the concept of efficiency is the Pareto optimum one. Hence, a Decision Making Unit (DMU) can be considered as efficient when it is not possible to reassign the existing resources in another way so that any other DMU improves without other getting worse.

Largely, any notion of efficiency relates a vector of inputs to a vector of outputs. There are two general approaches to measure efficiency: (1) parametric models, like Stochastic Frontier Analysis (SFA) (Coelli et al., 1998; Kumbhakar and Lovell, 2000), and (2) non-parametric models, like Data Envelopment Analysis (DEA) (Cooper et al., 1999). However, it is commonly argued that the DEA has some comparative advantages against the SFA when public sector activities have to be analysed (Martínez Cabrera, 2003).

The technical efficiency (TE) has an impact on productive processes. It can have both an input and an output view depending on the objective: maximization of outputs, or achievement of an already defined amount of outputs with the lower amount of inputs.

On the other hand the allocative efficiency implies getting the lower cost in the production of a given production level when the proportion of the used productive factors are changed according to their prices.

The evaluation of the efficiency requires the previous definition of a production function that features the productive process of the DMUs to be evaluated. Because of this, it is essential to know the technology used in the sector to the one the DMUs belong in order to define the input and output variables that may characterize this production function. Next figure illustrates exemplary the general idea of the efficiency concept.
The II' curve constitutes the so-called isoquant production frontier which represents the combination of needed productive factors to produce a unit of output, whereas points A, ..., E define scope and shape of the frontier. Point F stands above the frontier and illustrates an inefficient input/output combination. The technical efficiency of point F (Farrell, 1957) can be obtained by calculating \( \frac{0_{F'}}{0_F} \). Hence, TE has a range \( 0 \leq TE \leq 1 \). Point F refers to an inefficient observation because \( X_1 \) as well as \( X_2 \) can be reduced without any drop in output.

It has to be noted that the F and F' (DMU used for the comparison and measurement of the efficiency of F) have the same combination of productive factors. Farrell’s approach when measuring the technical efficiency consists of determining the result that the DMU being analyzed should achieve, if it was considered as another unit, that with the same proportion of productive factors but in a lower amount, achieves the same amount of outputs. Therefore, the consideration of the efficiency involves the comparison of a DMU not belonging to the production function frontier with another efficient DMU using the same proportion of factors, this is, a DMU located in the same vector. In some cases, this comparison will be made with hypothetical DMUs on the frontier, but that do not correspond to any real observed units, whilst in some other cases this comparison will be made between observed units.

As pointed out before, this research aims at evaluating the efficiency of public policies and their role in the behaviour of the IS by the efficiency of the program oriented to create university spin-offs (Rothaermel and Thursby, 2005). In order to achieve that goal, and due to the data needs of efficiency approaches, some variables have been selected and categorized according to their view in the process (input-output-outcome):
Table 2: Indicators to be used

<table>
<thead>
<tr>
<th>INPUTS</th>
<th>OUTPUTS</th>
<th>OUTCOMES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Budget (Euro) coming from the dependent organ (regional government, BIC…)</td>
<td>Nº of created spin-offs</td>
<td>Nº of employees due to the spin-offs</td>
</tr>
<tr>
<td>Budget (Euro) coming from other funds</td>
<td>Average life of the firms (spin-offs)</td>
<td>Nº of patents</td>
</tr>
<tr>
<td>Budget (%) coming from the dependent organ (regional government, BIC…)</td>
<td></td>
<td>Nº of subscribed contracts with the university</td>
</tr>
<tr>
<td>Budget (%) coming from other funds</td>
<td></td>
<td>Nº of institutions cooperating with</td>
</tr>
<tr>
<td>Nº of people employed in the program</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The data for the previous variables are currently being collected from all the Regional Business Innovation Centres (BICs) in registered in the National Association of Business Innovation Centres (ANCEs) in Spain since 1996.

From this analysis it will be possible to provide a new empirical base to deepen in the study and characterization of Innovation Systems, not only according to the data provided by general S&T statistics but also with a new approach oriented to determine the (in)efficient use of the resources in the development and implementation of innovation policies in the territories.

6 CONCLUSIONS AND RECOMMENDATIONS

Under the designation of policy evaluation we can find a huge amount of approaches and methodologies. Their evolution from the early 1980s could be considered tremendously fast compared to other disciplines but all suffer from an obsolete design considering the level of requirements posed by the new regional policies. More specifically we can find how regional STI policies have evolved to give room in their designs to the specific characteristics of the region as a geographical space unique and different to that of the nation. On the other hand we still can realise that evaluation methodologies are inspired by the linear innovation process model but many policies are designed under a systemic approach to the innovation process. Therefore we find two sources of lack of coherence between regional STI policy design and formulation and policy evaluation. These problems tend to be solved with the application of a “methodology mix” in the sense of apply an specific methodology from the pool with specific changes in its design in order to cope with the specific policy demands, just trying to adapt the methodology to the policy evaluation criteria. Therefore we can observe the application of quantitative and qualitative methods.
together to a policy in order to widen the scope of the evaluation according to the new demands posed by regional policies.

This is the case of new evaluations carried out under a quantitative approach based on the consideration of inputs provided by the policy and the outputs, outcomes and (or) impacts measured after the policy application. This approach is based in an efficiency measure which determines the most efficient use of inputs regarding the level of outputs (outcomes or impacts). This approach is obviously based on a linear consideration of the innovation process and, taken it solely, could render misleading results. We propose to combine its use with the analysis of the region under a systemic approach in order to put into a context the results obtained with the previous efficiency ranking.

However, there is no evaluation methodology that can be applied in every situation. Each policy needs its specific evaluation methodology design. Furthermore, the policy design process has to include the ex-post evaluation process as part of the policy cycle in order to really integrate this phase and design the most suitable methodology to apply to the policy.

**REFERENCES**


