



**Paper prepared for the 122<sup>nd</sup> EAAE Seminar**  
**"EVIDENCE-BASED AGRICULTURAL AND RURAL POLICY MAKING:  
METHODOLOGICAL AND EMPIRICAL CHALLENGES OF POLICY  
EVALUATION"**

**Ancona, February 17-18, 2011**



**The use of figures in the evaluation of rural development  
policies: a quest for knowledge  
Counting, to tell and understand**

Le Roy A.<sup>1</sup> and Millot G.<sup>2</sup>

<sup>1</sup> Université Pierre Mendès France, CREG, Grenoble, France

<sup>2</sup> Ministère de l'Agriculture, de l'Alimentation, de la Pêche, de la Ruralité et de l'Aménagement  
du Territoire /SSP, CEP, Paris, France

alr@orange.fr



# **The use of figures in the evaluation of rural development policies: a quest for knowledge Counting, to tell and understand**

Le Roy A. and Millot G.

## *Abstract*

*Using figures seems to create rigour, objectivity, knowledge and it facilitates comparisons. Consequently, an evaluation without figures is hardly conceivable. Nonetheless, objectivity and precision can be just an impression given the fact that figures are constructions built on a modeled description of reality. The simplification of reality operated through a figure can hide subtle elements regarding the way public policies work. If figures can legitimately be used in evaluation, every kinds of figures and evaluations are not equivalent. Therefore, our main research question is what place for figures in evaluation? This contribution relates to research about policy evaluation, seen as a mean to produce knowledge useful for the understanding of policies and their implementation. Based on the analysis of the evaluations of rural development policies conducted by the French ministry of agriculture our goal is to increase practical and theoretical knowledge of those policies through well-designed evaluations.*

*Keywords: Data, evaluation, methods, rural development policies*

*JEL classification: R58 , Q18, H50.*

## **1. INTRODUCTION**

Figures give an impression of rigour and certainty about the points they defend. Those characteristics of figures are convenient for policy steering or knowledge production, but they can be an illusion. Despite their concrete aspect, figures are modeled representation of reality. The underlying theories can become obsolete or wrong because society changes or simply because scientific knowledge improves. Since policy evaluations strongly rely on figures, we can therefore wonder whether the results of the use of figures in evaluations are always sound. To address this issue, the following questions helped us in our thinking process. What is (are) the place(s) for figures in policy evaluation, considering this process as multiple aimed – monitoring, accountancy, learning, etc. –? What kinds of figures and data analysis do we need to answer a wide range of questions? What do we learn about policy steering from the various uses of figure we may encounter? We refer to analysis of the process of quantification of social, economic and environmental phenomena (Desrosière). Our approach is also inspired by the ideas of evidence-based policy (Laurent et al, Sanderson, Stame) and wishes to observe reality from different angles as promoted by Critical Realism (Lawson). It also refers to the Realistic Evaluation school (Tilley). Our thinking is situated at the joint between research and public decision. We do not tend to look systematically for strong evidences as is it sometimes defended in evidence-based policy making, but we focus on the quantification process and on the data

production process, regarding the future use of those data. From a more practical point of view, our thinking relies on the analysis of the practises of the evaluation of rural development policy at the Ministry of agriculture in France (MAAPRAT). Therefore it tends to increase practical and theoretical knowledge of those policies and of the systems they have effects on, i.e. rural areas and farms. Rural development policies face the challenges of being place based as well as European and if not integrated, they concern at least several sectors. Through this paper, we will present the uses of quantitative data in the evaluation of rural development policies (2) and analyze the role of quantification in the steering of enlightened policies (3).

## **2. GATHERING AND USE OF QUANTITATIVE DATA IN RURAL DEVELOPMENT POLICIES' EVALUATION**

When rural development policies analysis and evaluation refer to figures, which data sources are they based on (2.1.)? What kind of analysis and data processing can be implemented with those data (2.2.)? How do those figures insert themselves into an evaluative process in order to provide relevant answers to pertinent evaluative questions (2.3.)?

### ***2.1. Data sources***

We can identify three main types of data sources from those often used in rural development policies: (1) data from surveys conducted by public statistics offices, (2) administrative and financial monitoring data and (3) data from surveys conducted by evaluators.

*Data from surveys conducted by public statistics offices.* Data from those services are produced independently from implemented policies. First and foremost, they monitor the social, economic and environmental evolutions of the country and of the economic sectors. There are two ways to produce statistical data : surveys based on sampling – farm structure survey, farm accountancy data network – and exhaustive censuses – agricultural census, population census. A census is interesting because it provides an overview of all beneficiaries and non beneficiaries of a given policy. However, its frequency can make its results rapidly obsolete, especially for the evaluation of programs of short duration. Moreover, several policies will base their evaluation on a single census, therefore, it is impossible to schedule this census to match the schedule of every single policy. In addition, censuses are exhaustive from the point of view of people and businesses involved ; reason for which they are useful for the quality of sampling of subsequent surveys and speak to the knowledge of the population of potential beneficiaries of a policy. However, such studies cannot give as detailed information on all the thematics. Based on smaller samples and focused clearly on well-established topics, surveys can have more frequent updates and enable a more detailed description of observations. If the sample reflects correctly the overall population, it might be too small to offer accurate description for sub-populations such as « beneficiaries » and « non-beneficiaries ». Let us observe that a good number of surveys are steered at European level (FSS, FADN) and that Member States (MS) do not have

room to adjust their sampling criteria to their specific national situation. When used within the framework of an evaluation, data from censuses give snapshots of the situation before and after implementation of the policy, whereas, data from surveys contribute to explain the evolution observed inbetween.

*Monitoring data.* Monitoring data facilitate the steering of policies and provide means to verify that implementation follows the original scheduled program. From a financial point of view, data checks the rhythm of expenditures and from an administrative one, it makes sure of the eligibility of potential beneficiaries. Monitoring data are also steered by overall principles of rural development strategy, eg. the age and gender of beneficiaries help to secure the promotion of the equality of chances through the policy implementation. Monitoring data are to be collected during policy implementation but people involved in their production or gathering do not always see their importance for future evaluation. Other tasks to perform are often seen as having a higher degree of priority such as strategic programming or simply getting along with implementation per say (paying subsidies to beneficiaries, etc) and therefore monitoring may be partially overlooked or even omitted. Unfortunately, when evaluation comes, the missing data cannot be recovered. Also, monitoring data in the framework of one given policy can be relevant to other policies. For instance rural development policy evaluation, in France, asks for data from social welfare policy monitoring. Generally speaking, monitoring data provide an overall analysis of the strategy of the evaluated policy by highlighting the measures and objectives to which most money is dedicated, and thus revealing implicit or explicit strategies. However, monitoring data do not offer information about non beneficiaries and very few about potential beneficiaries whose demand for a subsidy was not accepted. Since monitoring data are mainly developed for the steering, they are not sufficient for the evaluation of policies, regardless of their usefulness in that matter.

*Data from survey conducted by evaluators.* During the course of an evaluation, in addition to statistical and monitoring data gathering, evaluators produce their own set of data through surveys, sending of questionnaires by post or email and interviews, face to face or on the phone. Those enquiries target mainly beneficiaries and non beneficiaries of measures as well as representatives of institutions involved in the policy implementation. Data thus produced are mostly qualitative and give more indepth descriptions and pertinent results than those obtained through other data analysis however sound. Unlike the previous methods, they are particularly accurate and relevant when it come to the analysis of sociological mechanisms involved in policy success or failure. If time and means are available, evaluators can even produce quantitative data using a large sample of interviewees. In that case, the statistical quality of the sample may sometimes be called into question since interviewees cannot be forced to send back questionnaires or to answer all questions.

If data production and gathering is a necessary step for evaluation, analysis and data processing must follow. This crucial second step gives the interpretation and meaning of the previously gathered raw data in order to answer evaluative questions.

## **2.2. Data processing and analysis**

The use of quantitative data within the framework of an evaluation is not the use of rough statistical data and cannot be restricted to the filling of tables of indicators. Much like in any political science studies, it involves a diversity of data processing and presentation methods and tools which are not specific to evaluation. Using data brings into question the time, space and topic in evaluative analyses.

*Diversity of data analysis.* The toolbox used in policy evaluation is basic descriptive statistics. Mere averages and standard deviations give interesting results when applied in comparisons between samples. The criteria defining the compared samples can be of various nature according to the issue addressed. They can be geographical in case of compensatory measures for less favoured areas (LFA) where a comparison between mountainous and non mountainous areas can be made. They can refer to the size of agricultural businesses in case of measurements of investments, whose effects may differ from small to large farms. They can also relate to the main production of farms; cost of setting up a young farmer on a crop oriented or on a cattle oriented farm is not identical. The results of those basic statistical tools can be presented in many different, rather visual, ways – tables, graphs... Still, a sound written commentary giving proper explanation the results and explanation of the simplifications entailed by the model used remains necessary.

The second pillar of the common agricultural policy (CAP) is a good example of a policy which requires geographical criteria analysis and cartography in its evaluation. Historically, measures in favour of LFA or Natura 2000 encompass geographical analyses. This need was enhanced by the recent integration of broader rural considerations into the second pillar, necessitating descriptions of social, economic and environmental diverse dynamics in rural areas. Cartography enables the identification of correlations between dynamics based on a geographical analysis. However, observation is no demonstration; the causal link underpinning the correlation, if any, still needs to be identified and proven. Requests for spatial analysis of the implementation of rural development policy led to the creation of a rural development monitoring system (Observatoire des politiques communautaires de développement rural), based on an agreement between the Ministry in charge of agriculture (MAAPRAT), the payment agency (ASP) and the National Institute for Research in Agronomics (INRA). This monitoring system includes a cartography tool.

*Ex ante* evaluations requires for a capacity to anticipate the effects of a given policy and the use of tools such as economic models, cost-effectiveness analysis, cost-benefit analysis or multi criteria analysis. Impacts of a policy are quantified, "value for money" is estimated, and a

great number of quantitative data relate to costs and expected effects (Le Roy, 2009). Whether they are used for *ex post* or *ex ante* evaluations, the interest of such evaluation methods relies on their capacity to draw an objective balance between pros and cons of a policy implementation (usually advantages and costs). However, the lack of data about some impacts may prevent them from being taken into account, therefore leading to biased results.

Statistical models can be used for the estimation of the effects of policies. Those tools were originally designed for medical research. They were tested in France to determine the true effects of agro-environmental measures (AEM). They offer a fine quantification of the effects really due to the evaluated policy but, to be used with good accuracy, they ask for great amounts of quantitative data and can face sampling issues. Those estimation tools can be useful for the evaluation of simple measures with a clear rationale but they are not suitable for wide programs and policies with many objectives and which may have to work in a great variety of contexts.

Three issues affect this quantitative data analysis in view of an evaluation. The first one is the geographical scale of analysis, the second one is the time frame necessary for the anticipated phenomena to happen and be observed and the last one is the thematic scope concerned.

*Spatial and temporal scales, concerned thematic scope.* Which is the most adequate geographical scale for analysis and data processing? Governance and implementation of rural development policies differs according to levels of observation and stakeholders. The sheer concept of rural development can be different from one region to another, from one MS to another or depending on the geographical level at which impacts are observed to another. Moreover, the analysis from the beneficiary point of view asks for microeconomic data and does not provide the same results as a macroeconomic analysis on a broader scale. The effect of an investment measure in the food industry may be positive for the beneficiary, but if the subsidised business supplants other local businesses, there can be little or no positive effects at local level.

The issue of temporal scale is quite similar. Significant amount of time might be necessary to observe and understand the changes induced by public policies. Moreover, data production need time, which is not always acceptable to politicians or administrations eagerly waiting for the evaluation results. Scientific and political schedules barely ever match. Simplified, over-simplistic production of data can ensue to meet the later's agenda. Neither political or evaluative timeframes are long enough to grant the time necessary for a policy to produce all its effects. Experts need to draw their conclusion from one program before preparing the next one, taking even more time away and triggering an early final evaluation. The introduction of *in itinere* evaluation in the EAFRD offers a partial solution to this issue. Adequate timing of an evaluation is a sensitive issue, setting temporal scales and horizons is not a neutral act.

The inventory of data sources to be used in the evaluation of rural development policies also raises the question of the thematic scope to be covered. This scope may be far larger than the sectoral scope of the evaluated policy. Measures aiming at diversifying agricultural households incomes toward craftsmanship or tourism may have negative impacts on craftsmen

and people who work in the tourism sector, since it may introduce market distortions. Finally, let us not forget that the European rural development policy includes also forestry and rural economy in general.

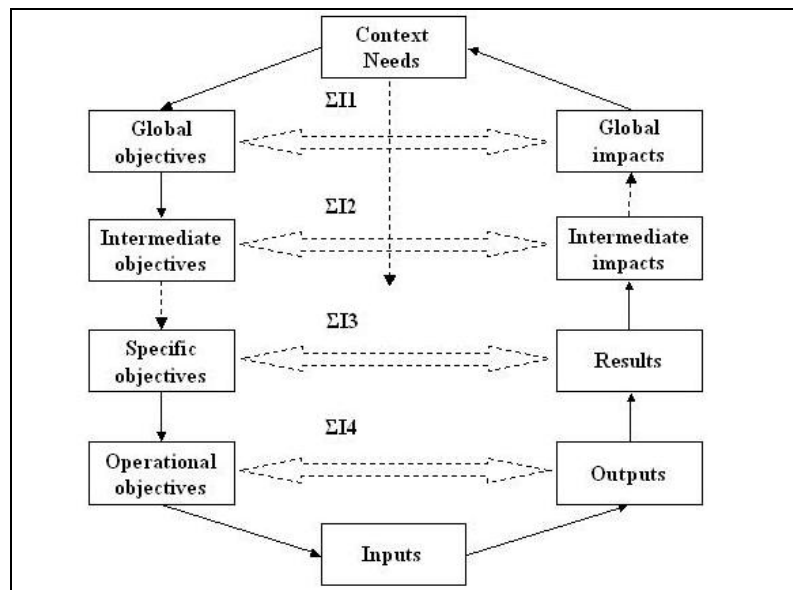
The previous analysis shows a great diversity of situations in terms of data production, gathering and processing. It highlights several sensitive issues that must be addressed and we may therefore wonder about the roles of figures in evaluation.

### 2.3. Figures in evaluations

The place and role of quantitative data in evaluation must be presented and explained, but it must be reminded that it is a mean to an end rather than an end unto itself.

*What are the places for quantitative data in evaluation process?* Considering that evaluation is mostly about studying the rationale behind a policy, (figure 1), figures are an illustration of the context, breaking down the several levels of data analysis, from microeconomic to macroeconomic into intelligible bits.

Figure 1: indicators systems ( $\Sigma I$ ) in the cycle of policy design and evaluation



Source: own elaboration

*Mostly about context?* Because of its European status, the second pillar of the CAP comes with a common monitoring and evaluation framework (CMEF). This framework sets an ensemble of evaluative questions and a table of indicators at different levels of the policy cycle. Indicators measure inputs, outputs, results and impacts of the policy, as well as the context in which the policy works and its baseline – its main trend of evolution. This framework also advises that a diagram be established to explain the rationale of the policy. This diagram not only creates or synthesizes knowledge on how the policy works, it is a part of the reference used



to judge the quality of the policy. It can be seen as a checklist of expected impacts (both positives and negatives) of the policy and is used to make sure that the effects of the policy – output, results and impacts – are consistent with its objectives (see figure 1). From this correspondence, at least theoretical, between objectives and effects of the policy, we clearly see that the indicators proposed for the monitoring and the evaluation of the second pillar of the CAP, are the same who describe the objectives of the policy, which can be “simply” considered as a desirable evolution of a social, economic and environmental context. Each level of objectives / effects, has its own indicator system ( $\Sigma I$ ), which measures the context of the policy at different levels, from the macroeconomic level to more detailed levels of description and analysis.

So to clarify the description of the current system of indicators of the CMFE, we may make a distinction between:

- monitoring indicators (inputs, outputs, results), which are linked with the direct effects of the policy. The causal link between the policy and the observed effect is easy to establish.
- and context indicators (context, baseline, impacts), which are linked with the indirect effects of the policy, or the modification of the context at a broader scale.

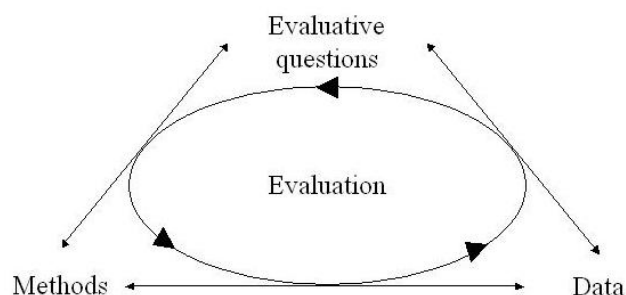
This concept of context is not taken into account enough in evaluation, or is too often disconnected from the policy itself. The tendency to focus on the policy and its mechanism may lead to overlook the context in which the policy was designed and on which it is supposed to have effects. This issue is even more crucial for the European rural development policy, since it is implemented in wide variety of national, regional and local contexts. Figure 1 is a simplification of reality, commonly used in evaluation and policy analysis. It supposes that the needs and objectives of the policy are formulated from the analysis of the initial context, and relies on the hopeful thinking that the final context will be consistent with the situation of society as anticipated during policy design. Therefore, it hides the fact that the policy is context-sensitive and in turns affects the context; context which translates differently at each level of policy objectives and effects in decreasing proportions from macroeconomic to microeconomic level. For example, the number of farms producing organic goods is both dependent of the measures to promote organic farming– as an output of those measures – and independent from them. Organic farming existed before its promotion by rural development policies. This evolution of the general context under the influence of policies has an impact also on statistical data production, even though it is theoretically considered as independent from policies. The modernization of French agriculture under the influence of the CAP ruled out hand-milking, and therefore the need for question about farms being equipped with milking machines in surveys. In evaluation, taking into account the role of context is essential, since evaluation tends to identify which proportion of the observed changes is due to the policy, and which is due to the natural evolution of the context.

*From microeconomic results to macroeconomic impacts.* Currently, figures can appear to be the core of policy evaluation practices. This distorted image of evaluation is mainly due to

the emphasis put on “notation” and the growing use of performance indicators and result indicators, coming from the sphere of private businesses. Performance and result indicators have a meaning and play a role in the sphere of public policies, but so do impact indicators. When evaluating a policy, we try to measure its overall impacts on society, its contribution to general interest. Due to their position as macroeconomic data, impacts indicators describe mainly the context of a policy, rather than they measure directly its effects. Moreover at macroeconomic level, several policies have an influence on the same impact indicator. It is therefore almost impossible to set a target level for those indicators. The number of jobs created thanks to rural development policies is one of those indicators under strong influences positive or negative from macroeconomic phenomena and other policies. The monitoring of such an indicator’s value tells us if the context evolves in the desired direction.

Context and policy are two parts of a dialectic system. The hegemony of indicators tends to separate the two. Focusing on indicators can lead to forget about the meaning of the policy itself. Consequently, transparency of evaluation and knowledge production through evaluation processes are diminished. Understanding the reasons for success or failure of a policy appears far more important to us. How a policy can be improved? Can success stories be adapted to other contexts? This is the kind of knowledge necessary and useful for future policy design.

Figure 2: evaluation: a process based on the triangle data, questions and methods



Source: own elaboration

*Figures: a tool for decision, but not an end.* In current evaluation practices, quantitative data are widely used and seem more important than evaluative questions. They are in fact part of a whole argumentation process, which tends to provide sound qualitative answers to those questions. The experience of the centre for studies and strategic foresight of the MAAPRAT in the domain of assistance to evaluation highlights three fundamental points, interrelated and on which evaluation relies: (1) evaluative questions, (2) quantitative data, (3) methods. Figure 2 presents this triangle.

As presented in sections 2.1. and 2.2. there are no single link between data – methods – questions. For each evaluation there are new sets of questions, new data and new methods. Economics plays an important but not exclusive role. Therefore, the use of data in evaluation implies to go constantly back and forth between data (rough or processed), methods and questions, in order to adjust all three together to produce the expected knowledge. An evaluative

question asks for some data processing so the first elements of answer can be enunciated. The results from this first analysis are often likely to raise new questions about the potential causes for the observed phenomena, among which the success, or failure, of the evaluated policy. Complementary data sets and analyses are then used to discriminate main from minor causes.

### **3. WHAT ROLE FOR FIGURES IN ENLIGHTENED DECISION MAKING?**

Evaluation is "a collective construction of practical judgements" (Perret, 2009), so it seems quite natural for this process to rely on heterogeneous information, quantitative and qualitative, in order to provide a socially legitimate judgement (Millot, 2010). Figures, which are used for evaluation, are embedded in such a process of collective construction and its limits (3.1.) and data production must take this embedment into account in order to contribute to policy evaluation and the subsequent knowledge improvement (3.2.).

#### ***3.1. The use of figures in knowledge production and political debate***

The diversity of vocabulary and uses of figures cannot hide that quantitative data must be the result of a collective production to be able to enlighten political debate and contribute to knowledge, even though this construction cannot be perfect.

*Figures: diversity of uses and vocabulary, but still a construction.* A figure does not have a single use or function. It is by definition, multidimensional. This should lead users to define precisely the function of the figures they use. This process enables the identification of the most accurate kind of data to serve the function it is expected to serve. Thus, within the framework of evaluation, we will be able to develop enlightened use of figures to produce knowledge about the mechanisms underlying rural dynamics and on which policies tend to act. On the opposite, data production frequently has to cope with various needs from users, some of which can be contradictory. The statistical service at the French Ministry of agriculture organizes users' committees before conducting surveys and censuses. Those committees include researchers, evaluators, managing authorities and statisticians in order to steer data production in a direction which will favour as many users as possible.

We may speak of data, variable, statistics, descriptor or indicator, all of them are a transformation of reality. Those data may be rough or refined, they all inform about the context, contribute to the quantification of effects or causes, proven or unproven causal links. They may also induce intuitions or reveal noticeable facts. Evaluation too has those functions.

Finally, whatever the word we use, a figure is a more or less refined construction which tends to complete several functions from an observation position. Observation is based on socio-economic analysis and can be defined as an ensemble of elements and quantitative or qualitative tools which enable the capacity of a society to produce knowledge (Le Roy, Offredi, 2010). This observation position can be implemented at several levels: policy mechanisms take place at microeconomic level and common general interest is defined at macroeconomic level.

Most of the difficulties in policy steering relies on the balance between effects on both levels. Statistics operates the same shift from microeconomic to macroeconomic level through its descriptive process. In any case, defining the point of view from which we look at reality is necessary to make sure that data convey the expected information.

We can therefore assert that quantification is a process, not a procedure, based on an agreement between people involved (Gadrey, 2006). Those conventions do not necessarily rely on any prior consensus. A strong idea, likely to guide us during a quantification process, then emerges: we shall not disconnect the design of indicators from the socio-eco-environmental system and its underlying theories, whose validity may be only temporary or strongly contextual. For example, farm yielding can be considered regarding used arable area (UAA) or workforce. Those information are different, none of them is superior to the other, they both answer the same evaluative questions through different criteria. Such a line of questions must respect the complexity of the issue and the diversity of concepts in terms of rural development. Therefore, we really need a collective data production, which would have a critical function towards the analysis of the existing categories. Geographically targeted policies highlight this difficulty. Areas with specific challenges such as Natura 2000 areas rarely fit the geographic subdivisions commonly used by statisticians. Therefore, an observation system based on a fine geographical grid is needed, data can always be aggregated into larger areas, consistent with the desired information. This requires flexible tools able to address the different needs of observation at all levels. This is clearly a quantification process for an enlightened steering of a visionary policy not merely the management of one, thereby serving policy making. Such a pattern expects us to change our mindset and question the aura benefiting figures. Criticism targetting figures like GDP, unemployment rate and inflation is commonly accepted nowadays. Those criticisms call into question the essence of public statistics and the needs they are supposed to address (Guibert, 2008). Figures come from an imperfect knowledge creating process.

*A imperfect construction asking for a wise use of data.* The evolution of institutional and decisional contexts (multiple levels of intervention, new forms of public policies) induces a multiplication of diversified demands for figure. Meanwhile, the developpement of policies based on projects and contracts favours the multiplication of indicators. This trend can be risky. Figures become simple packaging or marketing tools to drag rare finances, in a context of omnipresent competition and benchmarking, with high requirements for results. All of this happens at multiple level. However, statistical concepts, methods and tools, available at the national level may have very little relevance at smaller scales, and vice versa. The use of quantitative data must occur only in the framework of the theories underlying their production. The transposition of indicators from an aggregated level to a local one is risky because the models behind data production have their meaning only at their level of definition (Gadrey, 2003). The validity in time of an indicator faces the same transposition issue because of the evolutions of society, scientific concepts and policy objectives. The need for more sustainability

in the development of societies and rural areas is likely to call into question models based on an ageing consensus. Moreover, the way data is understood can change in time, its use may differ from its original one. In the 1950, GDP was originally introduced as a variable for Keynesian macroeconomic models. It measures only merchant and administrative productions, but its success induced a drift in its use to measure the economic performance of a region or well-being. However, this indicator was not built to that purpose and should not be used this way. In the end, criticisms toward GDP and its uses induced reflexions to create new indicators. Those would take into account some of the values that the market does not currently encompass (Gadrey and Jany-Catrice, 2007). In the end, this leads to new models and new quantification processes to measure richness or social and environmental well-being, knowing that "what counts is what is counted" (Viveret, 2002).

The activity of a farm can be explained through its used arable area, the size of its cattle or the number of people it employs. There are no systematic correlations between those data. Given the variety of agricultural contexts in Europe, those indicators produce different information and enlighten reality at various angles. Multiple points of view mean multiple policy options. Confronting those indicators reveals several agricultural models and contradictory objectives in policies. They test various hypotheses about the role we want for agriculture and rural areas in societies. As stated previously, the use of several indicators produces knowledge about the coherence of policies. Evaluation reports often highlight the contradiction between the will to modernize agriculture and the will to maintain job opportunities in agriculture. Therefore, we draw the conclusion that data must be a basis for a flexible array of indicators, which can enlighten the complexity of systems and not oversimplify them as a single indicator would do. Precise understanding of issues and the ways policies address them is at stake.

The debates of the last fifteen years raised by and around public statistics become more global. They induce "a global reflexion about economic, social and ecological balances of the planet" (Desrosières A., 2008). This is not the end of quantification, just the beginning of new figures, capable to contribute to the production of knowledge about the interaction between rural dynamics and rural policies. This is a major idea on which data production must rely.

### ***3.2. Producing quantitative data to serve an evaluative process and for a visionary policy***

Once the most dangerous traps of statistics are identified, we can now make proposals for the improvement of the content and methods of public statistics. This need for new statistical data increases with the evolution of demands and more generally with the evolution of the judgement values which counts – toward a better care for environment and a better social cohesion – This idea follows the hypothesis of J. Gadrey and F. Jany-Catrice (2003, 2007), the work of D. Méda (1999) and P. Viveret's report. The aim is not to subordinate data production to evaluation and the political sphere, but to discover and quantify in details new or unstudied social, environmental and economic phenomena. Therefore, evaluation and statistics share their

deontology: transparency, independence and impartiality. Ideas are needed about the best way to link data and knowledge production, through well defined point of view in evaluative processes.

Often, indicators are directly associated with evaluation. At least, they are seen as dependent of the political process. Indicators are supposed to be developed because a policy is implemented and will be evaluated. This is a mistake. Quantification is a process in itself. The example of organic farms can be here reminded. The presence of organic farming can be measured through indicators such as the number of organic farms without a policy to promote it.

*The different uses of figures according to the points of view of users and producers.* With quantitative measurement, every knowledge is more precise and more objective, but the importance given to figures and the way they are integrated in a reasoning is different from the point of view of researchers, citizens or politicians. Indeed, figures have a reputation of being a sound basis for decision making as long as politicians get useful empirical data to make their policies. Figures when used by researchers, look like they have the serious backing of genuine science. Moreover, those figures can help clarify blurry comparisons. However to be properly used, it is necessary to resist the temptation to aggregate partial data or to build simple or complex indicators to answer the needs of hurrying politicians, the needs of media who like small quantities of "meaningful" indicators (Jany-Catrice F. and Méda D., 2010).

Therefore, according to the function of figures and the point of view of the people who ask for them, causes for the use of figures and uses of figures are different. For researchers, the use of figures revolves around a theoretical debate leading to the choice of a dataset as a result of a quantified construction to test an hypothesis. In political spheres, figures aim at quantifying a phenomenon and highlighting an issue from which a debate will emerge about the solution to find. In both cases, the relation is a wayed one. A relation in both directions would trigger a dynamic process of knowledge creation. Without reciprocity, the partial debates on each side may be steril. From our analysis of the evaluation practices at the MAAPRAT, presented on figure 2, this is precisely what the use of figures in evaluations can bring: an interface between policy and research, toward more evidence-based policy making.

*Towards figures serving evaluation as a process producing knowledge.* As argued in previous paragraphs, behind data production is a diversity of potential uses. If none is superior to others, it is indeed important to find means to get evaluation, data production and policy analysis to work closely together. The task is not easy but should enable the production of data, collectively designed, useful for evaluation, the production of knowledge and the improvement policies. Thus, the quantification of environmental impacts in agricultural domain could lead to relations between concerned Ministries so common indicators and methods could be developed. That said, it is necessary to keep in mind that (1) information give power and (2) the word indicator can even be frightening. It may often be seen as an alarm data.

Clearly it is obvious to say that each proposed indicator must be set in time and space, but it must also be used with a clear reference to the way it was calculated, its definition and the information it gives. Then what recommendations can we make?

Two concrete questions raised during new data elaboration should be considered carefully. The first one is linked with the number of data produced and the second is related to the concrete way of production of those data. Too many data can be counter-productive at least for two reasons. First, there is a real competition, some may say unfair, with synthetic indicators (Jany-Catrice F. et Méda D., 2010). But above all, proposing a set of indicators enables free decision for people involved in judgements. Those are supposed to know what matters, but it is not always the case. A careful assistance to the proper use of indicators may often be necessary. Moreover, in terms of elaboration of indicators itself, it is necessary to find ways to include stakeholders in the debates so the results would be more legitimate.

This is not easy, but it is necessary. Such a process, organized around a relevant use of figures, collectively constructed, with a clearly assigned role (among all possibilities), will increase the usefulness of evaluations (Offredi, 2010). Such a change in practices would enable the creation of a sound operational knowledge for decision makers and for enlightening citizens' debate. Indeed the first aim of evaluation is operational knowledge, but scientific knowledge also benefits from evaluation. Evaluation is a great opportunity to create collectively some data to quantify identified mechanisms and strengthen the understanding we have of them. Such data are not just to introduce realism in evaluation, they must also convey meaningful information and provoke thinking about mechanisms underlying public policies.

### **3.3. Recommendations**

We argued that knowing the interaction between policies and systems on which they act is required for the improvement of policies. The close embedment of policies in their contexts makes policies themselves look like they are almost part of the context. It is hard to imagine our societies working without the omnipresence of public policies at many geographical levels. This is particularly true for agricultural policies. Therefore, it is possible to conceive policies as being an object for study by public statistics. This would imply to look more often at the links between policies and the changes observed in society. Do policies favor positive trend and prevent negative ones? Linking statistics and decision making does not mean that statistics will be subordinated to the political sphere but that statistical offices should be better used to provide assistance for policy steering. Statistics is indeed well situated to measure the effects of a policy and produce knowledge about its mechanisms.

First, it is necessary to have a more systematic collective gathering of data by linking more closely monitoring and administrative data and statistical data. For example, the questionnaires of the agricultural census of 2010 were already filled with administrative data about subsidies received by farmer from the CAP. This enables to save time during interviews with farmer. Moreover, in censuses and surveys, there is a need for greater information about

the status of farmers as beneficiaries or non beneficiaries regarding various measures. This information is useful afterward measuring opportunistic effects.

Public statistic services cannot anticipate the production of all the data necessary for policy evaluation. This implies that evaluators will always have to produce some data through interviews during evaluation. However, the quality of those data could be improved, with a better sampling of interviewees. Technical assistance from public statistics services could be helpful in that regard. Another solution to sampling bias could be to strengthen, afterwards and from a statistical point of view, the results from interviews conducted by evaluators.

Indicators highlight phenomena, but often in an indirect way. A phenomena can appear differently from a MS to another because of cultural differences. Thoses various appearances of a same phenomenon asks for different indicators. Commuting, a characteristic phenomenon of counter-urbanized areas, can be translated in terms of distance or time of commute. Therefore it is necessary to focus the line of questions on the phenomenon itself, rather than on the indicator to measure it, or even worse on the value taken by the indicator. To begin with, satistical data producers should enlighten evaluators on the meaning of the data they produce. They created those data so they know when the interpretation of the data becomes boarderline. Then, we need to continue the collective identification of the phenomena that really matters and that we should measure. For each phenomena, evaluative questions must be formulated and finally we must propose a set of indicators to highlight it at different angles. Focusing indicators on their meaning, that is to say on the phenomena they measure, will enable a better identification of a typology of rural areas in Europe, which is a sound basis for a relevant comparison between MS. It is also important to stabilize an array of sampling criteria in order to be able to conduct comparative analysis of rural dynamics and policies effects on those dynamics. This would imply a collective answer to the following question: which of the rural dynamics do matter?

In evaluation, it is often necessary to measure several phenomena, interacting together. Quantification often reduces too much the complexity of the measured systems or describes it as an ensemble of smaller simpler systems. Those practices do not reflect reality properly (Fouquet, 2010). Therefore, it is necessary to have "systems of several indicators" which do not overreduce the complexity of reality but keep its multidimensional aspect. Furthermore, it is also necessary to further develop the production and the gathering of data related to rural areas in general and not only in their agricultural dimension.

The second pillar of the CAP is a European policy, it includes partnership and subsidiarity at each geographical level. Therefore, the scale for production of the data is an important matter. The development of place-based approach of public policies asks for a representativity of samples at small scales so the data can be agregated on several perimeters, according to thematics dealt with, and so impacts be measured from local to global scale.

Policy evaluation produces knowledge and can therefore identify new phenomena. Statistics could enable detailed descriptions of those phenomena, but since they are new, data may be missing, or sampling may not be adequate to measure them accurately. It may be temporarily necessary to overweight some categories in the sampling process. Such situations



are great opportunities to create collective knowledge about farms, rural areas and the way they react to policies. Generally and in a more systematic way, the strengthening of an ongoing evaluation process, with a common work program between evaluators and statistical data producers will favour synergetic links between evaluation and statistics. Policies, which are implemented for several years, such as rural development ones, are likely to promote such an anticipated common work program. For example, it was decided to overweight the areas where territorial AEM were implemented in the sampling process of the Terruti-Luca survey, conducted by the statistics service of the MAAPRAT.

#### **4. CONCLUSION : FIGURES FOR VISIONNARY POLICIES**

A priori, figures induce rigour and facilitate comparisons in time and space, and an evaluation, which produces knowledge without the use of figures, is hard to imagine. From the use of quantitative data in the evaluation of rural development policies, we saw that the idea of objectively reliable figures can sometime be an illusion. The simplification of reality with a figure may hide much subtleties when it comes to understanding the way policies work. This can even lead to false results in evaluation, knowing that figures play entirely their role when they can be trusted. A quantitative data is the result of a construction process based on a conventional way to represent reality, shared among involved stakeholders.

Statistics produce photos, motionless pictures at distant moments in time, but evaluation is a process, which try to produce knowledge, similar to a movie. It tries to understand what happened between to images produced by statistical services. We can try to shorten the time between to pictures but we will not be able to understand better what happened in between. Therefore there is a need for statistical movies. Statistical data should observe phenomena when they happen, not only before and after they produced their effects. We need to be able to count, so we can tell and understand. It is necessary, to foster evaluation process with "relevant" figures, coherent with evaluative questions and capable to produce a refined knowledge. Thus, the contribution of evaluations to knowledge about a society, its policies and the way it all works together, enables us to change our point of view, which is often focus on the role of sciences in evaluation. The use of figures in an evaluation is intermediary between the use of figures by scientist and the use of figure by politicians. Evaluation is therefore an interface between those two spheres. Understanding and taking into account the needs of evaluation in data production is necessary. When quantitative data are considered, in their production and uses, as the result of shared conventions about multidimensional phenomena, the triad of evaluation between data, methods and questions can work properly. Then the quantification process becomes observation for a visionary policy, cleansed from the fault of a managing policy.

## REFERENCES

- Berriet-Sollic M., Le Roy A., Trouvé A., (2009). Territorialiser la politique agricole pour plus de cohésion, *Revue d'Economie Rurale*, n°313-314 : pp.129-146.
- Bureau D., Mougeot M. Studer N., (2010). Mesurer la performance de la gestion publique à la lumière de l'analyse économique, *Revue Française des affaires sociales N°1-2* : pp.89-104
- Desrosières A. (2000). 2nd édition, *La politique des grands nombres, histoire de la raison statistiques*, Collection Science humaines et sociales, La découverte, Paris.
- Desrosières A., (2008). *L'argument statistique*, 2 tomes, Presses des Ecoles des mines, Paris.
- Epices et Tercia, (2008). *Evaluation ex post du plan de développement rural national* - questions transversales et synthèse générale.
- Fouquet A., (2010). L'usage des statistiques : de l'aide à la décision à l'évaluation des politiques publiques, *Revue Française des affaires sociales N°1-2* : pp.309-322.
- Gadrey J. (2006). Les conventions de richesse au cœur de la comptabilité nationale. Anciennes et nouvelles controverses, in Eymard-Duvernay F. (ed.) *L'économie des conventions, Méthodes et résultats*, Paris, La découverte.
- Gadrey J. et Jany-Catrice F., (2007). *Les nouveaux indicateurs de richesse*, Collection Repères, La découverte, Paris.
- Gadrey J et Jany-Catrice F. (2007). *Les nouveaux indicateurs de richesse*, Collection Repères, La découverte, Paris.
- Gadrey J. (2008). A la recherche de nouveaux indicateurs de bien être et de santé sociale, in *Les statistiques publiques en débat*, Regards sur l'actualité n°346, la documentation française, Paris, pp. 53-62.
- Guibert B., (2008), La production des statistiques publiques en France, in *Les statistiques publiques en débat*, Regards sur l'actualité n°346, la documentation française, Paris, pp.5-17.
- Jany-catrice F. et Méda D., (2010). Les conditions sociales de la production du rapport Stiglitz sur les indicateurs de performances économiques et de progrès social, in Degrave F (ed.) *Les transformations et innovations économiques et sociale en Europe : quelle sorties de crises ?* Regards interdisciplinaires.
- Jany-Catrice F et Zotti R., (2009). La santé sociale des territoires. Un indicateur de santé sociale pour les régions françaises, *Futuribles n°350*, mars.
- Labrousse A., (2010). Nouvelle économie du développement et essais cliniques randomisés : une mise en perspective, d'outils de preuve et de gouvernement », *Revue de la régulation n°7*, 1<sup>er</sup> semestre.
- Laurent Cl. et al., (2009). Pourquoi s'intéresser à la notion d'evidence-based-policy ?, *Revue du tiers monde n°200* : pp.853-873.
- Le Roy A et Offredi Cl., (2011). La quantification au service de l'observation sociale local : à quelles conditions ?, *Revue Française de Socio Economique*, premier semestre.
- Le Roy A. (2009). Favoriser le rapprochement des pratiques du calcul économique et de l'évaluation pour une meilleure compréhension de l'action publique, 9<sup>ème</sup> journée française de l'évaluation, SFE, Marseille 11-12 juin.
- Lawson, T. (1997). *Economics and Reality*. London: Routledge.
- Méda D, 2008, *Au-delà du PIB. Pour une autre mesure de la richesse*, Champs-Actuel, Paris.
- Millot G., (2010). L'évaluation des politiques publiques, aide au pilotage de l'action ministérielle, Centre d'étude et de prospective, *Analyse N° 19*, MAAP, Paris.
- Offredi C. (2004). Penser le développement durable et l'évaluation in *Dynamique de l'évaluation et du développement durable*, L'harmattan, Paris.
- Offredi C. (ed), (2010). *La notion d'Utilité sociale au défi de son identité dans l'évaluation des politiques publiques*, L'Harmattan, Paris.
- Perret, B., (2002). *Indicateurs sociaux, états de lieux et prospectives*, Rapport pour le CERC, Janvier.
- Perret B. (2008). Evaluation et statistiques : une coopération difficile, in *Les statistiques publiques en débat*, Regards sur l'actualité n°346, la documentation française, Paris, pp.45-52.
- Perret B. (2009). Indicateurs de richesse et évaluation in Offredi C (dir.), *La notion d'Utilité sociale au défi de son identité dans l'évaluation des politiques publiques*, L'Harmattan, Paris.
- Sanderson I. (2002). Evaluation, policy learning and evidence-based policy making, *Public Administration*, vol. 80, n°1 : pp. 1-22.

*Ancona - 122<sup>nd</sup> EAAE Seminar*  
*"Evidence-Based Agricultural and Rural Policy Making"*

- Tilley N., (2000). Realistic Evaluation: An Overview, Founding Conference of the Danish Evaluation Society.
- Trosa S., (1992). Le rôle de la méthode dans l'évaluation à travers l'expérience du Conseil Scientifique de l'Evaluation, in *Politique et management public*, Vol X, n°3 : pp.83-102.
- Truvé A., (2007). Le rôle des Régions européennes dans la redéfinition des politiques agricoles, Doctorat en sciences économiques, Université de Bourgogne, UFR de sciences économiques.
- Viveret P., (2003). *Reconsidérer la richesse*, le tour d'Aigues, Ed. de L'aube.