
ADVANCED-EVAL WORKING PAPER SERIES
Technical Report TR1-4

Mid-term workshop proceedings

Current version: January 2008

**Proceedings: ADVANCED-EVAL Mid-term workshop review
(Brussels, 17 January 2008)**

VENUE AND DATE:

A Mid-term review workshop of the Project “ADVANCED-EVAL” financed under the EU 6th Framework Programme was held in Brussels (Hanse-Office) on 17 January 2008.

PARTICIPANTS:

Workshop participants (approximately 25 people) included representatives of the EC (DG-AGRI, DG-Research), European Investment Bank (Evaluation Unit), specialised Agencies in EU-member states responsible for evaluation of EU structural programmes (e.g. PARP Poland), project partners, external evaluators of RD programmes in EU member states, and research academia.

OBJECTIVES:

The overall objectives of the workshop were to:

- address the most important methodological problems linked to the evaluation of EU RD programmes;
- introduce new quantitative tools which allow a more accurate evaluation of the impact of RD programmes and were used in the framework of the project “advanced-eval”;
- discuss the applicability and significance of new evaluation techniques by comparing empirical results obtained on the basis of those techniques with other methods traditionally used in the evaluation of results and impacts of EU RD programmes;
- discuss the role of institutional and personal capacity building in the evaluation of RD programmes and sustainable rural development.

PRESENTATIONS:

The workshop plenary presentations and following discussions focused on four inter-dependent topics. These were:

1. The expectations of the European Commission concerning the evaluation of EU RD programmes;
2. The presentation and discussion of state-of-the-art current methodologies applied to evaluations of EU RD programmes in individual EU member states;
3. The introduction of advanced methodologies for *ex-ante*, mid-term and *ex-post* evaluations of RD programmes applied and developed in the framework of the “advanced-eval” project (based on mid-term and *ex-post* evaluations of SAPARD in Poland and Slovakia);
4. The empirical applicability of the proposed methodologies.

TOPIC 1: Evaluation of EU rural development policies: Expectations from the viewpoint of policy makers

Professor Christian Henning (CAU, Kiel) and Dr Hans-Joerg Lutzeyer (DG-Research) first welcomed the workshop participants.

Mr Guido Castellano (Evaluation Unit in DG-AGRI) then gave a presentation entitled “*Expectations of the evaluation of RD programmes from the view point of policy makers*”: this focused on the EU strategic approach to rural developments. Monitoring and ongoing evaluation based on the Common Monitoring and Evaluation Framework are expected to play an essential role within this framework (see: Rural Development Regulation, EC, 1698/2005). The basic purpose of the evaluations being carried out on RD programmes at both the Community and national/regional and field levels, are:

- the assessment of quality, efficiency, and effectiveness of measures;
- the identification and appraisal of needs for change;
- the assessment of strategic orientation for policies and programmes;
- a better programme focus on needs;
- compliance with legal requirements to evaluate; and
- accounting for public money.

The extensive scope of evaluations cover the transmission of inputs (financial resources) not only into outputs (usually measured in physical terms), but also into results (direct effects) and impacts (indirect effects measured at regional and community level). The role of the policy evaluation extends from what can be directly controlled (i.e. inputs => outputs) through what the programme influences directly (i.e. results), ending on what the programme influences only indirectly (i.e. programme impacts). Given this logic, the main stages in evaluation can be seen as:

- *structuring* (i.e. establishing intervention logic, defining key terms of evaluation questions and indicators);

- *observing* (identification of sources of information and defining the scope of case studies);
- *analysing* (i.e. carrying out quantitative and qualitative assessments and answering all evaluation questions); and
- *judging* (drafting conclusions and recommendations).

In terms of timing and duration, the *ex-ante* evaluation of a new programming period has to be carried out before the end of the previous programming period. Mid-term and *ex-post* evaluations are carried out during respective mid-term and post-programme application periods, and an ongoing evaluation is expected to be implemented during the whole duration of the programming and application stages (including former and new programming periods).

In his presentation, **Mr Castellano** stressed the role of the on-going evaluation which should involve not only a strategic orientation, but also an improvement of evaluation methodology and organisation. A continuous capacity building and exchange of best practice should be implemented as much as possible through the newly created *evaluation networks*. With regard to *ex-ante* evaluation, this should be based on common evaluation guidelines, and provide a timely contribution to the programming of the following application period. Especially important are the identification and appraisal of quantified programme targets; the definition of clear baselines; the identification of expected results; the identification of community value added, and the appraisal of procedures (financial and technical management).

The main features of mid-term and *ex-post* evaluations of current RD programmes are:

- focal points in the process of “ongoing evaluation”;
- mid-term evaluation in 2010 and *ex-post* evaluation in 2015;
- responsibility of Member States for carrying out evaluations;
- synthesis under responsibility of the Commission;
- work carried out by independent experts;
- evaluation at a programme level;

- applicability of common baseline and impact indicators;
- applicability of complementary programme-specific indicators.

Especially important is ensuring that indicators used in mid-term and *ex-post* evaluations of RD programmes adequately reflect the programme intervention logic, i.e. the hierarchy of indicators is consistent with the needs and hierarchy of programme objectives. The following types of indicators used for monitoring and evaluation can be distinguished: common baseline indicators, common monitoring indicators, common impact indicators and additional programme specific indicators. Mr Castellano stressed that all basic elements concerning the methodology of programme evaluation have recently been comprehensively described and are available in the *Handbook for the Common Monitoring and Evaluation Framework (EC, DG-AGRI December 2006)*.

TOPIC 2: Introduction to the Project ADVANCED-EVAL

In the “*Introduction to the EU-Project: Advanced-Eval*” **Professor Christian Henning (University of Kiel)** stressed that policy evaluation has become a major issue for many international organisations, including the European Commission. In particular, the EU has committed itself to obligatory *ex-ante*, mid-term and *ex-post* evaluations of rural development programmes.

Although the EU commission has set up a comprehensive evaluation framework, evaluation techniques applied to EU rural development programmes at present basically correspond to the simple measurement of programme progress indicators, while a comprehensive impact evaluation has not yet been applied. Nevertheless, it is well understood that, in contrast to pure monitoring, an effective programme evaluation has to correspond to impact evaluation, i.e. beyond the measurement of programme inputs and outputs, a quantitative relationship among programme inputs and impacts has to be empirically specified. A comprehensive quantitative policy evaluation improves efficient policy learning and policy planning. Moreover, efficient impact evaluation increases accountability of politicians and administration, and therefore improves efficiency of rural development programmes. However, at present, serious methodological obstacles to an efficient evaluation exist. In this context, ADVANCED-EVAL is a European research

project that began its activities in March 2006. Research in the Advanced-Eval project aims at:

1. Developing quantitative techniques that enable a non-biased *ex-post* evaluation of the impact of EU RD programmes, i.e. disentangling the programme effect from other exogenous factors and calculating a programme's substitution, replacement, deadweight loss and other indirect effects.
2. Deriving an aggregated indicator of regional development, the Rural Development Index (RDI), to be used for *ex-ante* and *ex-post* evaluations of RD programmes in rural areas.
3. Improving existing theories of rural development, in particular by investigating the role of social network relations in rural development and their influence on the impact of RD-programmes.
4. Developing quantitative techniques that enable a quantitative *ex-ante* evaluation of RD programmes.
5. Comparing evaluation results of quantitative advanced methods with those of qualitative methodologies.

Accordingly, in addition to a project management unit the project is organised into five working groups whose work packages correspond to the main project targets mentioned above. According to the overall research plan the project has three phases:

- the first year of the project was used to set up theoretical and methodological conceptions and to begin data collection;
- the second year has been used to complete the data collection and estimate derived evaluation tools and models;
- in the third year, results of new quantitative tools will be compared to existing qualitative evaluation tools. In addition, derived new methodological and theoretical knowledge will be disseminated and discussed with scientists, politicians and practical evaluators.

To this end, several workshops and seminars will be organised. Overall, realised project progress corresponds to the project plan, where, to date, more than 20 working and technical papers have been produced. These papers, as well as further information, are available on the project webpage (<http://www.advanced-eval.eu/>).

TOPIC 3: Evaluation of SAPARD: Experiences in the Slovak Republic and Poland

In his presentation “*Evaluation of SAPARD: Experiences in the Slovak Republic*”, **Mr Gejza Blaas (Research Institute of Agricultural and Food Economics, Bratislava Slovakia)** focused on the description of the SAPARD programme that was implemented in Slovakia between 2000 and 2006. The total financial allocation for the programme amounted to € 102 million. Based on three priorities – improvement in the agricultural sector including the food industry, sustainable rural development, and development of human activities – the programme comprised nine measures, of which the most important were investment in agricultural enterprises and the improvement of food processing and marketing.

Originally, eligible projects were only able to be submitted in certain regions (those lagging behind), and targeted only a few commodity sectors (sheep, poultry, fruit and vegetables). Those limitations were lifted at a later date and eligibility criteria were extended to the cattle and pig sectors and to the whole territory of the state. Most applications were submitted within Measure 1, “Investment in agricultural enterprises”, and Measure 2, “Improvement of food processing”. Those two measures also attracted the highest total of approved commitments (1.9 and 2.7 billion Slovak Koruna respectively), and payments totalling 65% of all support funds disbursed to beneficiaries.

The majority of approved projects were submitted by large agricultural companies. The territorial distribution of approved projects and disbursed funds in absolute terms was biased in favour of administration districts in the South-West of the country (the most productive agricultural area), where the most viable farms have been located and the food industry concentrated. However, projects and funds associated with Measure 4b, “Rural infrastructure”, were spread much more evenly over the country.

There was some delay in the implementation of the programme, so the bulk of applications were submitted during 2003. That is why at the time of mid-term evaluation,

in October 2003, only a few completed projects were found ready for mid-term evaluation. The mid-term evaluation framework followed the scheme of common and measure specific evaluation questions and criteria. Since very few data on result and impact indicators were available, evaluations were based on information drawn from case studies, project applications, and general and sectoral statistics. In the follow-up to his presentation Mr Blaas focused on evaluators' experiences regarding the use of some evaluation criteria and indicators, which refer to Measure 1, "Investment in agricultural holdings".

Criterion: Improvement of beneficiaries' income

There are several possible indicators. While the evaluators opted for "gross profit" one could also use "gross margin" or "net value added". At the farm level (in the case of Slovakia - large mixed farms), the value of this indicator depends on many factors not directly linked to the supported enterprise, such as fluctuations of terms of trade, weather conditions, etc. It would be better to examine how the ratio between cost and sales revenue of the specific enterprise changed. We could also argue that the immediate effect of the supported investment could also be a deterioration of cost/revenue ratio because of higher depreciation costs.

Criterion: Increase in factor productivity

This can be measured in terms of physical units such as yields per cow, yields per hectare and so on. This is an obvious advantage because such information is usually available on the farm.

Criteria such as "complying with quality standards" or "complying with working, health and safety condition standards" may also be quite useful, but only if the appropriate information is available. As a rule, this must rely on a qualitative evaluation and this mostly requires an individual on-the-spot assessment.

Mr Blaas concluded that impact evaluation using a wide range of indicators only performs well if the necessary data are available. This stresses the importance of a functioning monitoring system. It should provide the necessary information, but not impose too large a burden on stakeholders.

In his presentation “*Evaluation of SAPARD: Experiences in Poland*”, **Professor Leszek Klank (Institute for Agricultural and Rural Development, Warsaw, Poland)** focused on the most important issues linked to the implementation of SAPARD in Poland. Among rural development projects implemented during the years 1991-2007 the most important were:

The World Bank Projects:

Agricultural Development Project	1991-1995	\$100 million
ASAL 300	1992-1995	\$300 million
RDP	2000-2004	\$120 million

EU Programmes:

PHARE	1989 to date	ca. € 2.5 billion
SAPARD	2000-2006	€ 1,068 billion (total)
of which		€ 0.680 billion (EU)
RDP/PROW	2007-2013	€ 16-17 billion

Furthermore, there were some other smaller RDP projects financed by UK, USA (USAID), Norway or Switzerland.

SAPARD was the only EU Rural Development Programme that went through all stages of programme assessment. It was the largest assistance grant for Polish agriculture and rural areas. Officially SAPARD started in 2000 but, in fact, the first applications were not submitted until the second half of 2002. The total value of the SAPARD programme was ca € 1.1 billion, of which the EU grant was ca € 0.7 billion (63.7%).

The main objectives of Poland’s SAPARD programme were as follows:

- Improvement in the competitiveness of Polish agriculture and agri-food sectors on both domestic and international markets;

- Adjustment of the agri-food sector to EU requirements, in respect of hygiene, sanitation and quality;
- Support to rural development by the development of an infrastructure and off-farm job creation in rural areas;

The SAPARD measures included:

- Improvement of agri-food and fish processing and marketing;
- Farm investment;
- Development and improvement of rural infrastructures;
- Diversification towards non-agricultural activity;
- Agri-environment and forestation improvement;
- Professional education;
- Technical assistance.

Although SAPARD was not very popular when it was first launched (political actions against EU accession) programme funds were eventually fully utilised. The utilisation of programme funds was as follows:

Improvement of agri-food and fish- processing and marketing

Some 1,342 eligible agri-food processing enterprises benefited from programme resources, ca 1.5 billion PLN allocated to this measure (€ 385 million¹)

Farm investment

A relatively small number of farms (1342) benefited from the SAPARD programme. The total amount of money spent on this measure was PLN 558 million or € 143 million (at the same time there were 1.95 million farms in Poland)

¹ Exchange rate: 1 euro/3.9 PLN

Development and improvement of rural infrastructure

The largest amount of money went to local governments (communes). There were 4,483 applications submitted by the communes for the amount of ca PLN 2 billion (€ 513 million). Almost 85% of eligible Polish communes participated in the SAPARD programme; however, as many as 57% were in the largest Polish region of Mazowsze.

Diversification towards non-agricultural activity

Only 4,071 applications in that measure were successfully submitted; the amount of money granted was PLN 330 million (about € 85 million).

Agri-environment and forestation improvement

The environmental programmes were not implemented due to formal reasons.

Professional education

Some 14 applications for the amount of PLN 44 million were granted.

Technical assistance

Only 12 applications were granted for the amount of PLN 5 million.

The *ex-ante* evaluation of SAPARD was an ongoing process that started with an initial seminar on the interpretation of draft regulations and the necessary preparation in June 1997. During 1999, after the setting up of the Department of Pre-Accession Assistance in the Ministry of Agriculture, PHARE assistance was given (seven days of consultancy) to help with discussions about the overall strategy, and designing the way in which consultations should be carried out. Twinning support for more intensive day-to-day advice became available in mid-1999. In a follow up 12 consultancy days were allocated in September 1999 to evaluate the first drafts of the proposals (Coherent National Agricultural and Rural Development Strategy).

A Ministry seminar was organised to debate the conclusions after more intensive consultation with individual members of the Department of Pre-Accession Assistance. The EU consultant's remaining time of some 18 days, plus approximately 10 days of the two main Polish consultants, was spent in commenting on further drafts and queries during the revision process. While the Agricultural Policy Unit (FAPA-SAEPR) had

provided some *ad hoc* analyses to *ex-ante* evaluation, the final *ex-ante* evaluation was carried out during December 1999.

Ex-ante evaluation took place over a period of three weeks (an independent team consisted of senior Polish consultants and an EU-based independent consultant who worked in association with the Institute for Rural and Agricultural Development, Polish Academy of Sciences). The *ex-ante* assessment coincided with the finalisation of the drafting of the Rural Development Plan (after consultations with experienced practitioners).

The availability of information had strictly limited the work of the drafting and evaluation team. Official information on rural society and agriculture had been assembled and improved using much of the recent work of researchers in this field, including that carried out under the auspices of the Institute for Rural and Agricultural Development, Polish Academy of Sciences. While a good description existed of the situation in agriculture and the food sector in some rural areas of Poland, both a suitable methodology that could show the programme's possible impact and a great part of necessary information were lacking. For the same reason many of the evaluation questions could not be answered. The situation was exacerbated by changes in the administrative structure of Poland, rapid changes in the agri-food sector as well as the informal status of peasant types of agriculture. In this situation some numerical estimates provided in *ex-ante* evaluations must therefore be treated with extreme caution. Some heroic assumptions had been used in the appraisal to avoid the error that no impacts occur if a number was not available.

The scope of SAPARD *ex-ante* evaluation was as follows:

1. Overall assessment of the SAPARD strategy;
2. Analysis of rural and agricultural problems and opportunities in the Polish SAPARD Programme;
3. Experience of rural and agricultural development and its relevance to the SAPARD Programme. Coherence of objectives of the SAPARD Programme in Poland;

4. Linkages between the measures proposed in the SAPARD Programme, national policy and other intervention;
5. Appropriateness of project selection criteria under the SAPARD Programme;
6. Appropriateness of rates of aid proposed under the SAPARD Programme;
7. Consistency of outputs and impacts of the SAPARD Programme;
8. Efficiency of implementation arrangements of the SAPARD Programme.

The analysis of the situation in rural areas was quite comprehensive and focused on the most important issues, such as rural unemployment, the poor educational level of the rural population, poor farming structure, inadequacy of rural infrastructure, low rural incomes. However, there was insufficient information regarding the availability of and access to capital in rural areas, mobility of the rural labour force, and functioning of rural institutions (the latter were expected to support the development of a market economy in rural areas).

Mid-term evaluation had been conducted between August 2002 and December 2003 and used mostly qualitative and formal methods for the SAPARD assessment. As the evaluation took place when SAPARD had just started there was not much to evaluate. An applied method also influenced assessment results. Two examples can illustrate this. The evaluators claimed that some 100,000 jobs would be created in non-rural activities in local economies, and farmers' and companies' income would go up by 20% due to SAPARD. However there is no formal justification of such statements. To the contrary, while some hundred thousands of newly unemployed in rural areas were recorded at that time, there is no evidence that any support from SAPARD had an impact in lowering this number.

Ex-post evaluation of SAPARD was conducted by Agrotec Polska, Agrotec Italy and IAFE Warszawa in November 2007. According to the first draft of the assessment report, the following aspects of the programme were examined:

- Analysis of the previous (existing) reports;
- Analysis of the degree and factors of the funds utilisation, relation and allocation;

- Assessment of the impact of selected measures on beneficiaries and their environment;
- Evaluation of the effectiveness and efficiency of the SAPARD programme (general approach);
- Assessment of the SAPARD monitoring quality;
- Assessment of the role of SAPARD on *aquis communautaire* implementation

Applied methodology included: data analysis, examination of strategic government documents, e.g. National Development Plan, National Cohesion Plan, etc, individual interviews, beneficiary survey (3,952 questionnaires according to measures and sub-measures), electronic survey (e-mail) for local communities (400 gmina), case studies for selected projects (48), expert panel, focus group interviews (MARD, ARMA, FAPA, Agricultural University, SGGW; Institute of Rural and Agricultural Development, and other qualitative techniques). Though the *ex-post* evaluation of SAPARD was much richer and deeper than two previous evaluations, the conclusions are as follows:

- The qualitative and legal (formal) approach dominated, which meant that only part of the EU requirements were fulfilled (CR1698/2005);
- Quantitative analysis was limited only to “before and after” comparisons of programme beneficiaries;
- Due to poor quantitative analysis, the impact of the SAPARD programme could not be separated from other development programmes or other exogenous factors;
- Comparisons of achieved results with *ex-ante* predictions showed large discrepancies;
- Only a limited role of SAPARD in rural development was found, contrary to expectations;
- Some research shows only a very weak or no correlation between the SAPARD programme and rural development;
- In some fields the competence of evaluators was insufficient (e.g. afforestation);

- Poor analytical methods of assessment were used; only effectiveness was examined but no programme efficiency;
- Farmers' lack of interest in the first years of SAPARD implementation was not explained.

TOPIC 4: Application of advanced methods to mid-term, and *ex-post* evaluations of RD programmes

Dr Jerzy Michalek (University of Kiel, Germany) presented a paper, "*Application of advanced methods to ex-post evaluation of RD programmes*". The main objectives of this part of the project were to:

- i) address methodological problems arising in the course of the evaluation of rural development programmes in the EU;
- ii) provide appropriate solutions to these problems by taking advantage of the most recent evaluation methodology;
- iii) apply modern evaluation methodologies for the assessment of the outcomes of RD programmes (i.e. SAPARD) in selected EU countries (i.e. Slovakia and Poland);
- iv) empirically estimate various programme effects as required in the EU evaluation framework (i.e. by providing quantitative answers to all relevant Common Evaluation Questions and Programme Specific Questions linked to the measure "investments in agricultural holdings"); and
- v) compare the estimated programme effects obtained by applying modern evaluation methodology with results which use traditional methods to *ex-post* evaluation of RD programmes in the EU.

From the methodological point of view, the basic problem faced by evaluators of RD programmes is: how to disentangle the effect of the programme from other (programme independent) factors. This is equivalent to answering the essential evaluation question: What would have happened to supported enterprises without an RD programme?

Clearly, a counterfactual performance of supported enterprises (i.e. non-supported situation) cannot be directly observed (fundamental evaluation problem, see Holland, 1986). For the same reason, a programme's effect can only be evaluated by means of formulating a counterfactual baseline and making comparisons between supported enterprises, possibly with identical ones that did not benefit from the programme. It is important to note that a considerable selection bias in evaluating the effects of RD programmes would occur if the mean outcome of those units which participated in RD programmes differed from the mean outcome of non-supported units used as controls, even in the absence of support.

In the majority of empirical evaluation studies of EU RD programmes, the effects of the programme are assessed on the basis of interviews with programme beneficiaries, without the formulation of a necessary baseline, i.e. without making comparisons of programme outcomes between supported and similar non-supported units. In a few cases, where such comparisons were attempted, this was done without any consideration for appropriate matching which led to a significant measurement error (due to a selection bias). While such comparisons continue to be standard techniques in both mid-term and *ex-post* evaluations of EU RD programmes, our project shows that these practices may be extremely erroneous. The main reason is that RD measures are either designed to target specific beneficiaries (e.g. *under-performed* communes, areas, etc.), or the specific programme eligibility criteria are formulated in such a way that programme participation is only feasible for the *best* economic units, e.g. Measure: "Investment on farms" in SAPARD (Poland and Slovakia). As a result, observed differences in average outcome indicators between programme participants and non-participants are subject to a considerable selection bias. Furthermore, "programme effects" calculated on the basis of data collected for programme participants only (using an "after vs. before" technique) or between programme participants and naively established control groups (e.g. all non-participants or a national average) cannot be attributed to the programme only, but also include other (i.e. programme independent) effects.

In an appropriate evaluation framework the relevant baseline showing potential outcomes for programme participants in case of non-participation is therefore *crucial*. Yet, one of the difficulties commonly faced during the formulation of such a baseline is the problem of a perfect comparability (ideally, the same enterprises which participated in the programme should also be used for simulation of their performance without the

programme). As this is not possible, comparisons should be made in a manner which guarantees that both groups are almost identical (prior to the programme). This means that the distribution of characteristics (covariates) describing enterprises which are subject to programme support should be as similar as possible to the distribution of covariates characterising a group of enterprises used as controls (i.e. the statistical probability of receiving support should be similar for the matched supported and non-supported enterprises in each comparison group).

Propensity score matching (Rosenbaum and Rubin, 1983; Heckman, 1998; 2002; 2005; Dehejia and Wahba, 1998; Smith and Todd, 2003; Imbens, 2003; Lechner 2002; Zhao, 2005, etc.) was the technique used in our research to ensure comparability between supported units participating in a RD programme with non-supported controls. The method balances the observed covariates between the supported group and a control group based on the predicted probabilities of receiving support from the RD programme. Supported units are matched to non-supported units on the basis of the econometrically estimated propensity scores (derived from an estimated logit function). After the elimination from the analysis of those units which are outside of the region of common support (and, therefore, non-comparable), the exact matching and assessment of programme effects are performed. The aim of matching is to find the comparison group from a sample of non-supported units that was closest (in terms of observed characteristics) to the sample of programme participants (only enterprises within a region of common support are considered).

The individual farm data on supported and non-supported enterprises from the SAPARD programme in the years 2002-2006 were used to construct appropriate baselines and quantitatively estimate the net-effect of SAPARD at the micro-level in Poland and Slovakia. The disentangling of programme effects from other exogenously determined effects was done by making quantitative comparisons of outcomes between programme-supported enterprises and the appropriate baseline (simulation of effects in a “non-support” situation) prior to the programme, and after it by using the combination of Propensity Score Matching (Kernel matching) and a difference-in-differences (DID) methodology. The evaluation of SAPARD results at the micro-level was carried out on the basis of estimated Average Treatment on Treated (ATT) effects calculated for specific micro-economic results indicators (e.g. profits, productivity, employment, etc.). Furthermore, sensitivity analysis was performed in order to assess a possible impact of

unobservable characteristics on the magnitude of obtained results (Rosenbaum bounds methodology was applied).

The empirical results confirm a full applicability of modern evaluation techniques to the assessment of EU RD programme results at the micro-level. Furthermore, significant differences in outcomes of the SAPARD programme (at the micro-level) obtained in dependence on whether traditional (naive) or more advanced evaluation techniques were used, stress the critical importance of the methodology. For example, a strong growth of labour productivity and income reported in evaluation studies of the SAPARD programme that used traditional techniques (further interpreted as the effect of the SAPARD programme) was found almost negligible as soon as more advanced methods were applied (comparisons were made on the basis of similar control groups).

Results of this research have clear implications for policy makers. Firstly, more caution is required when drawing conclusions on the basis of evaluation studies that apply traditional techniques only. Secondly, the use of modern evaluation techniques should be seen as a necessary evaluation prerequisite. Thirdly, a wide diffusion of modern evaluation know-how among external programme evaluators should be sought.

TOPIC 5: Advanced methods of ex-ante evaluations of RD programmes

In a paper entitled “*Advanced Methods of Ex-ante Evaluation*” **Professor Christian Henning (University of Kiel)** focused on the derivation of an adequate model framework for an *ex-ante* evaluation of RD policies. While the contribution of adequate *ex-ante* policy methods which contribute to efficient policy planning and policy learning are quite obvious, there are additional methodological problems of quantitative *ex-ante* evaluation methods when compared to *ex-post* analysis. In particular, propensity score matching techniques cannot be applied since, by definition, no data on individuals that receive programme intervention exist. Accordingly, other advanced micro-econometric evaluation techniques successfully applied for *ex-post* evaluation cannot be applied prior to policy intervention. Moreover, the participation of individual decision units in envisaged policy programmes has to be estimated *ex-ante*.

In this context, two different approaches have been developed within WP 5. First, a modified matching technique is applied that allows the mimicking of individual choice

behaviour under the policy intervention, based on an observed variation of covariates seen before policy intervention. Thus, as long as the impact of policy programmes on covariates observable before the policy intervention can be estimated, modified matching techniques can be applied.

However, there might be policy programmes for which modified matching does not work. Moreover, modified matching does not include general equilibrium effects. Therefore, a second methodological approach applied for *ex-ante* evaluation in WP 5 corresponds to the derivation of a simulation model. The derived theoretical model framework corresponds to a combination of an agent-based model and an interregional computable general political economy equilibrium model (CGPE-ABM).

In particular, the CGPE-ABM model includes four modules. The core module is a static interregional computable general equilibrium model that, in contrast to standard CGEs, incorporates actor and commodity specific transaction costs as well as an endogenous local government sector. In addition, a firm's investment decisions depend on the agents' beliefs in future prices and technologies, where belief formation is modelled via an agent-based model: this includes firms' information and communication interactions with organised specific networks. Land market transactions are modelled by an agent-based model of bargaining processes among local buyers and sellers. Finally, network dynamics are modelled in a fifth module.

The derived theoretical model has been calibrated on the basis of empirical case studies of high and low performing rural communities in Poland and Slovakia. Based on the calibrated model, various policy simulations are conducted for high and low performing rural regions. Based on these simulation runs:

- (i) The impact of specific tangible and intangible factors in rural development is identified;
- (ii) The impact of specific tangible and intangible factors on the effectiveness of individual RD measures is quantified;
- (iii) New RD measures improving local network structures are identified;

- (iv) Evaluation results applying modified matching techniques as well as existing qualitative evaluation methods are compared with those derived from the CGPE-ABM approach.

TOPIC 6: Concept and measurement of Quality of Life (QOL). Application of a Rural Development Index to evaluation of RD and structural programmes

In a paper entitled “*Different Ways to Conceptualise Quality of Life in Rural Areas*” **Peter Kaufmann and Dr Sigrid Stagl** (SPRU: Science and Technology Policy Research, University of Sussex, UK) stressed that analysis of quality of life (QoL) from an economic viewpoint is connected with the utility concept. After classical economists implicitly assumed that utility could be measured cardinally, new welfare economics had settled on an ordinal utility concept focusing on relative ranking and revealed preferences from the 1930s onwards. Those who promoted new welfare economics did not believe that utility could be measured directly. The main advantage of indirect measurements of QoL is that they are usually well-grounded in economic theory. The downside of this is that its validity cannot be seriously tested empirically (Van Praag and Frijters 1999).

Welfare theory in traditional economics is based on the notion of utility and is measured by income. Rational individuals maximise their utility through an optimal combination of goods bought from their income. The maximisation hypothesis of utility has been widely criticised because it was seen to constitute the exception, not the norm of an individual’s choice (see the literature on ‘bounded rationality’ and ‘satisficing’). When trying to scale up from the individual to the societal level, the measurement of welfare experiences serious problems. Aggregation would work only if a nation could be perceived as one person (Sen 1979). Nowadays more and more laudable economists and psychologists ask for a radical reform of utility theory in economics (e.g. Frey and Stutzer 2002; Kahneman and Krueger 2006; Layard 2006) on the grounds that new insights from psychology and behavioural economics can enrich the analyses of economic behaviour and its measurement.

Economists also expanded on capital theoretical approaches. A further development of the SNA, the (green) net national product (NNP) comprises an economy’s consumption and net investments into manufactured capital, human capital, and natural capital.

Dasgupta (2004) supports a pluralistic conception of personal well-being based on the theoretical notion of socio-economic, civil, and political citizenship. Wealth available for current and future generations is measured with genuine investment to denote change in a society's wealth. This approach is still plagued by data gaps in existing monitoring systems.

Based on longstanding research in psychology on the measurement of subjective well-being, economists more recently also started to use direct measurement of QoL/well-being by using stated preferences from surveys. The main advantage of direct measurement is that it does not require value judgements by the analyst because the aggregation of the various dimensions of well-being is performed by the responder in arriving at his or her stated well-being. Thus, it holds the promise to tap directly into the well-being domain without much distortion. Although some theoretical and methodological challenges still remain to be solved, new questions can be posed that were not possible to address within a revealed preferences approach.

A further proposal along this line of research is to focus on ill-being and its determinants instead of well-being. Kahneman and Krueger (2006) suggest this both for theoretical reasons (because prospect theory tells us that potential losses influences people's decisions more than potential gains), and for political reasons (because it might be easier to agree on detrimental welfare effects than on welfare enhancing ones as there is an array of dimensions to be considered). This conceptual research is in its early stages, but might have potential for an important contribution to the literature.

A body of literature also developed on measures of rural disadvantage (e.g. Index of Rural Deprivation, Rural Discrimination), which raises additional aspects of QoL not addressed in more generic indices. They improve our understanding of exclusion and deprivation (e.g. mobility and opportunity deprivation), which may occur in rural areas when uneven distribution of public service provision combines with poor accessibility to services such as health, education and retail facilities. In a policy context, measures of rural disadvantage are used to facilitate the spatial targeting of resources to areas which are most in need, and may also be used to identify subgroups of the population who are most at risk of experiencing poverty.

A further strand of literature employs migration rates to measure absolute and relative QoL across regions. In this type of study, the decision of a household to migrate is treated as to maximise QoL (see J. Michalek's contribution in this report).

Amartya Sen's (1993, 1985) capability approach (CA) represents a framework covering a comprehensive conceptual space to assess social welfare by explicitly respecting not only the economic, but also the social, political and cultural dimensions of well-being. The main strength of the CA is its focus on outcomes and opportunities, as well as the broad coverage of potential well-being functions. The main weakness of the capability approach is that it is still new and underdeveloped both conceptually and empirically, mostly due to a lack of suitable data from existing monitoring systems. Still, this does not prevent the main international organisations from already applying it and learning from improvements as they go along; it is still judged to be better than working with previously applied approaches.

As nature is the very basis of our being, sustainability of resource use needs to be reflected in deliberations about QoL. The main challenge from an environmental point of view is first to establish a widely accepted conceptual basis, which is currently still somewhat fragmented, and then to establish an appropriate monitoring system. Economists largely resort to either welfare theory or capital theory (see also Dasgupta) to integrate the environment into their analyses. The former tries to incorporate environmental externalities into their framework, and the latter differentiates between human, manufactured, and natural capital stocks, and implements a monitoring system to ensure they are non-declining over time. For the latter approach, one needs to keep natural capital accounts (renewable and non-renewable resources, sinks, ecosystems), and identify specific critical natural capital with associated threshold levels which may not be surpassed by overuse.

Measurement is not easy because relationships between ecosystem services and human well-being are not linear. Changes in the area of high levels of ecosystem services might not have much influence on well-being, but when it comes close to thresholds, human well-being can be influenced over-proportionally, even as far as withdrawing the very basis of any well-being in extreme cases (desertification of land, floods, etc.). A range of proposals have been suggested so far, the most prominent being:

- the System of Integrated Environmental and Economic Accounting (a satellite system of the System of National Accounts), which also implements a critical capital view of asset accounting based on the notion of intergenerational equity;
- the Millennium Ecosystem Assessment, which recently put forward an elaborate proposal; and
- the indicator framework by the Commission for Sustainable Development of the United Nations, currently the most widely used approaches for national cross-comparisons.

Concerning weights derived by purely indirect methods, one can observe advancements in weighting and aggregation in recent years (for an overview, see Nardo *et al.* 2005; OECD 2002). There remain many reasons for being cautious. For example, analysts can manipulate weights according to their mental predispositions; experts' weights may not be transferable from one area to another; different methods of normalisations of scores can give different weights; and weights can become inconsistent when a large number of indicators are analysed (*ibid*).

To conclude, one can see interesting approaches and findings from new avenues of research. Although some of those are not yet fully developed, you should not be deterred from applying them and/or integrating insights into your own applications to enrich your own analyses and contribute to their further development. Basically, more output/impact oriented concepts are preferable, but quite hard to implement in the current state of affairs. Data gathering exercises are still more focused on inputs/resources, though there is some development. This puts limits on what an analyst can implement, especially if he or she uses sub-regional datasets. Still, we can undertake interesting analyses with existing datasets that go beyond what is usually done.

In the next presentation, “*Measuring of Quality of Life (QOL) and the application of Rural Development Index (RDI) for the evaluation of RD and structural programmes*”, **Dr Jerzy Michalek (University of Kiel, Germany)** reported on the methodology applied in the construction of a rural development index (RDI) used as an impact indicator in *ex-ante*, *ex-post* and on-going evaluations of EU RD and structural programmes at various regional levels (NUTS-2-NUTS-5).

The EC strategic guidelines for rural development in the EU emphasise the key role of RD policies and programmes in achieving important social objectives, e.g. improvement of the quality of life for populations living in rural areas, creation of employment opportunities and improvement of competitiveness of rural sectors, sustainable development of rural areas, etc. While RD programmes affect a large spectrum of rural development domains, evaluators of EU RD programmes are expected to answer common evaluation questions (CEQ), and derive their findings on the basis of a large number of partial indicators as a proxy for potential programme's effects (e.g. production, income, employment, competitiveness, environment, etc.). Though this procedure is currently widely applied, the evaluation practice shows that a rigorous assessment of the impact of RD programmes can be difficult and even misleading, particularly if only crude evaluation techniques are applied. The major criticisms of existing common evaluation systems and common indicators are concerned with (comp. Forstner and Plank, 2004):

- i) the relevance and appropriateness of particular indicators suggested by the EC;
- ii) the lack of a coherent evaluation framework linking programme inputs, outputs and outcomes;
- iii) difficulties in the formulation of appropriate baselines (counterfactuals) describing development of a particular indicator in a control group of non-supported regions;
- iv) the lack of prioritisation between many indicators leading to patchy and inconsistent reporting.

Given the above and taking into consideration that:

- i) in each rural region the number of potentially applicable RD measures under an RD programme can be very large (currently up to 37 measures can be implemented);
- ii) individual RD measures implemented under specific RD programmes may simultaneously affect the wide range of various rural development domains (e.g. economic, environmental, social, etc.); and

- iii) all partial indicators (computed for each individual measure) need to be calculated both for supported as well as non-supported regions (baseline);

it is understandable that the evaluation of RD programmes should seek to make use of a methodology enabling the aggregation of separate programme effects (expressed in terms of partial indicators) into a synthetic quantitative index to be used for an analysis of an overall impact of RD programmes/selected measures in a given rural area.

The proposed RDI in this research can be seen as a synthetic measure of quality of life in rural areas; this is expected to combine various economic, environmental and social aspects of rural development and link them within a theoretically consistent framework. A comprehensive description of various methodologies and problems linked to a derivation of a meaningful QoL index was provided in Kaufmann *et al.* (2007). The review of existing literature on a composite RDI shows that construction of such an index can be subject to numerous methodological problems, of which the most important are:

- i) selection of appropriate variables;
- ii) balancing between objective vs. subjective indicators;
- iii) weighting the factors according to their relative importance;
- iv) using unbiased aggregation techniques in the building of a composite welfare index; and
- v) making the index useful for policy purposes (e.g. programme evaluation).

Although many empirical trials were undertaken to construct such a regional development index, the empirical assignment of appropriate weights to individual RD domains appeared to be the most difficult task (in most cases, only qualitative and/or rather *ad hoc* methods were applied).

A theoretical foundation of the methodological approach to the construction of the Rural Development Index (as a synthetic measure of rural development) applied in this project, originates from the past 50 years of research which focused on the relationship between migration and quality of life (see Tiebout 1956; Liu 1974; Schultz 1982, Fuguitt 1985; Michalos 2003; Berger 2003; Douglas and Wall 1993, 2000; Greenwood 1997; etc.). Under this framework, an individual's decision to migrate can be derived from an

individual utility function with observed differences in various economic, social and environmental factors characterising the origin and destination regions and transaction costs linked to a migration decision as the main arguments. By using this approach, the weights of the most important economic, social, environmental and other development domains entering the RDI can be derived empirically from the econometrically estimated intra- and inter-regional migration function.

The empirical construction of the RDI was undertaken using regional data (separately for NUTS-2-NUTS-4 levels) in Poland and Slovakia (2002-2005) using the following steps:

- i) by defining relevant rural development domains to be taken into consideration prior to assessment of the overall impact of the RD programme;
- ii) collecting data on economic, environmental and social aspects of rural development in above areas (by regions NUTS-2-NUTS-4, whereby hundreds or thousand's of variables were collected);
- iii) translating the above variables into meaningful coefficients (e.g. per capita, per km², etc.), and performing standardisation of those coefficients for all regions and years considered;
- iv) grouping calculated coefficients into respective factors in order to reduce the dimension of the analysis (the factor analysis-principal components method was applied);
- v) deriving weights for each individual factor from an econometrically estimated migration function, after selecting the most appropriate model (i.e. econometric models included: panel estimate random effect net migration model, spatial error and spatial lag models, panel estimate nested error component regression net migration model, panel estimate random effect logistic regression inflow model, and panel estimate logistic regression inflow model with nested error structure).

For each relevant rural region a synthetic index RDI was computed as a weighted sum of factors representing various development domains, with respective weights derived from the selected inter- and intra-regional migration function.

A quantitative evaluation of RD programmes' impact was carried out in three-stages. During the first stage, an RDI (used later as a relevant outcome indicator) was calculated

at a relevant regional level (NUTS-4) and year (before and after the RD programme was implemented), using the methodology described above. At stage two, prior to implementation of an RD programme, the propensity score methodology was applied for construction of an appropriate baseline (control groups of regions), whereby relevant variables (grouped into factors representing different development domains) were used as covariates. At stage three the overall impact of the RD programme on RDI, and other relevant outcome indicators (e.g. employment, turnover, etc.), was analysed (before and after) by applying a combination of the difference-in-differences (DID) and Average Treatment Effects on Treated (ATT) methods.

The above methodology was successfully applied to the analysis of the impact of individual SAPARD measures in Poland and Slovakia at the NUTS-4 level (years 2002-2005). In both countries the impact of SAPARD measures on the synthetic RDI was very small.

In **Mr Michalek** opinion, the main advantages of the above approach to the evaluation of RD programmes using the RDI as an outcome indicator are as follows:

- Comprehensive consideration of various development domains (economic, social and environmental) in the form of a synthesised index is attractive from a policy point of view;
- The outcome measure is easily adjustable (easy inclusion of new indicators/variables);
- Weights of the index are derived empirically and can be statistically verified;
- The RDI accounts for all programme general equilibria effects, e.g. multiplier effect, replacement effects, substitution effects, etc.
- The methodology is powerful both at the aggregated level (e.g. NUTS-2) as well as commune level (NUTS-5) or even at the village level (depending on the availability of data);
- The methodology is applicable for an analysis of effects of large projects at all regional levels;

- An evaluation method using propensity score matching for the construction of a control group of regions and RDI as a potential programme outcome, is methodologically superior compared with traditional evaluation techniques suggested by the EC.

Further improvement of the above methodology under this project will include the application of an extended version of the matching methodology, i.e. a generalised propensity score, enabling the calculation of programme impacts in the case of a region's continuous programme participation (i.e. participation as a non-binary variable). The project will also try to check applicability of advanced evaluation methodologies to "old" member states (*inter alia* through verifying data availability and applying these techniques to the assessment of RD programmes implemented in selected regions (e.g. Schleswig-Holstein) of Germany.

TOPIC 7: Social networks and rural development

Dr Vincent Buskens and Professor Henk Flap (Utrecht University, Netherlands) presented the paper: "*Social Capital and Labour Market Performance in Rural Regions*". As we know from a large body of literature, mainly in sociology, social capital often has positive effects on labour market performance. People with suitable networks find jobs faster, they often have higher salaries, are promoted sooner, and perform better while doing their job. Moreover, trust among people is higher in settings in which people have better social capital. This also promotes the general well-functioning of certain countries or regions within countries.

Because social capital is so important for social well-being and performance, it is also important to know the level of social capital in rural areas, and to understand how social capital works in rural areas for *ex-ante* as well as *ex-post* evaluation of rural development programmes. Therefore, we adapted existing measurement instruments of social capital to the rural areas under study in our Advanced-Eval programme.

The measurement instruments are included in questionnaires that are currently held in low and high-performing regions in Poland and Slovakia. Sixty pilot questionnaires have already been satisfactorily completed, while the main survey, aiming at 600 questionnaires in both countries, is currently underway.

The measurement instruments include name generators, resource generators, and questions on social and civic participation. The name generators are questions that ask for the names of the respondent's contacts, e.g., household members, referees for a job, or people with whom the respondent talks about important issues. For these contacts we then record some main demographics, such as gender, age, and education. We also record the strength of the relationship, the length of time the relationship has existed, and how contacts are interlinked among each other. Together these data form measures for the quality and extensiveness of the respondent's social network.

In addition, the resource generators try to establish resources respondents can access through their social relationships. This is done by asking whether respondents know people that have specific jobs, or people that can help in case the respondent has a specific problem. Finally, we have some item batteries that measure the social activities and participation of the respondents as well as the respondent's opinion about the functioning of his or her neighbourhood in general.

Using these data we can describe differences in social capital between Poland and Slovakia, as well as between high and low-performing regions. In combination with performance measures, such as whether someone has a job or what someone's income is, we can also investigate whether social capital has a positive effect on the performance and well-being of respondents. In addition, because we have similar data in different countries and different regions, we can use multi-level regression analyses in which we explain performance and well-being in these regions using, among other control variables, social capital variables, and see whether social capital functions differently in, e.g., different countries.

In conclusion, some final comments on the importance of these investigations for the project as a whole. First, for *ex-ante* evaluation, this research adds to the existing knowledge on the functioning of social capital and, with that, it provides a tool to predict the effect of specific rural development programmes depending on the social organisation in a region or country. Because social capital is difficult to measure, it will be unobserved in many areas in which rural development programmes are implemented. This implies that, for *ex-post* evaluation, social capital is part of the unobserved heterogeneity that might affect the evaluation results. By studying the importance of social capital in some rural areas in detail, we obtain insight into the extent to which

biases in *ex-post* evaluation might be related to variations in the effectiveness of development programmes as a result of differences in social capital.

The paper entitled “*Information networks, regional knowledge spillovers and regional economic growth: An agent-based approach*” by **Dr Volker Saggau (University of Kiel)**, focused on issues linked to the implementation of the combination of the agent-based model at the micro level, and the interregional computable general political economy equilibrium model (CGPE-ABM) at the macro level. One part of this software package is a module which has implemented information and innovation networks being the basis for the technological progress and development within a region. This module and its implications for regional economic growth with respect to network structures are explained in the following section.

Stimulated by new evolutionary economic approaches to innovation and technological progress, we analysed the impact of different information network typologies on regional knowledge spillovers, knowledge accumulation and technological progress.

Technically, a rather simple ACE-model is derived that particularly focuses on the role of information networks in the accumulation of knowledge by regional firms. New technological knowledge is exogenously generated in a leader region and randomly transmitted to regional firms. Within a region transmission of technological know-how occurs in information networks. Following central findings of existing ACE-literature on innovation (Dawid 2006), we assume that firms can only transmit information that they are able to process, where a firm’s capacity to process new technological knowledge depends on accumulated technological know-how. Therefore, given an exogenous rate of generation of new technological information in the technological leader regions, the speed of information accumulation within a region crucially depends on the speed of information transfer within the regional firm network, where the latter depends on network structures. Applying our simple agent-based model we simulate the impact of different network typologies on spatial diffusion of knowledge and regional technological progress. In particular, we simulate two different network types, i.e. small-world and free-scale networks, varying global network structures, which are clustering and centralising. The main results of our simulation studies are as follows:

- (i) Information network structures have a significant impact on both spatial information diffusion and regional technological progress.

- (ii) Information diffusion in networks is only imperfect, i.e. accumulated knowledge in regional networks corresponds only to a constant fraction of technological knowledge generated in leader regions.
- (iii) In particular, this fraction is c.p. higher for scale-free networks when compared to small-world networks. Moreover, this fraction increased for scale-free networks with the preferential attachment parameter and for small world networks with the α -parameter.
- (iv) In contrast to classical catch-up models, our network approach to spatial diffusion of technological knowledge implies that, except for extreme centralised or dense networks, catching-up does not occur. In contrast, depending on the concrete typology of regional information networks, a constant productivity gap to the technological leader is stabilised, or regions characterised by extremely clustered information networks are even increasingly falling behind.

With respect to policy evaluation of RD-measures, this example of this specific module of our CGPE-ABM already shows that measures focusing on improving information networks, such as, for example, Article 20a of COUNCIL REGULATION (EC) No 1698/2005: measures aiming at promoting knowledge improving human potential through:

- (i) vocational training and information actions, including the diffusion of scientific knowledge and innovative practises, for persons engaged in the agricultural, food and forestry sectors;
- (ii)..(v) have an impact on the development of a rural region. An empirical investigation of the existing network structures, such as in WP4, raises the possibility of evaluating the network structures by comparing them with best performing network structures with respect to knowledge spillovers and, implicitly, growth paths.

Although in the new ACE-literature on innovation the importance of established informal and formal inter-firm network structures for knowledge accumulation has clearly been recognised, it is still fair to conclude that there are hardly any existing studies which analyse the role of networks in regional knowledge spillovers. Our approach certainly contributes to the existing literature; however, to focus on the role of information

networks in knowledge spillovers, we applied a rather simple model neglecting other important factors determining innovation processes.

In particular, beyond knowledge spillovers, innovation involves other mechanisms, i.e. in-house R&D or learning-by-doing, which have been neglected in our approach so far. Moreover, our approach simply assumes that accumulated knowledge is the key factor of technical progress, although beyond knowledge technical progress often requires investment in new capital goods. This investment is often characterised by a fundamental uncertainty, i.e. firms have to form beliefs regarding future states of the world without knowing, *ex-ante*, the set of all possible future contingencies. Therefore, a firm's decision to invest in innovation can be better captured by ABM-approaches, assuming strong substantive and procedural uncertainty, than by dynamic optimisation models with Bayesian updating or even perfect foresight.

Social networks also have a significant impact on agents' belief formation, and thus on firms' investment decisions regarding innovation projects. However, these aspects of the impact of social networks on innovation have not been considered so far in this approach, but we do consider it with the other modules of WP5 and with the CGPE linkage to the GAMS environment.

The main objective of the presentation, "*Transaction cost representation in the market model (CGE)*", by **Professor Thomas Heckelei (University of Bonn, Germany)** was to show possibilities for modelling transaction costs of market access within a CGE model, differentiated by market levels. The required link to the Agent-Based Model, which calculates network dependent transaction costs for different market levels, requires specific consideration in this context. The general idea of linking an ABM and CGE model within the Advanced Eval project is to show the effects of transaction costs related to the access to different market levels. Transaction costs are modelled in a local network context by the ABM. In order to compute equilibrium prices and trade flows from the local to regional and international level and vice versa, a CGE model is employed. Apart from supply and demand behaviour at the different levels, the resulting equilibrium is also influenced by the relevant transaction costs associated with specific market destinations. In turn, transaction costs depend on the volume of trade flows. Consequently, a simultaneous solution to ABM and CGE needs to be found.

The simultaneous solution can be achieved by applying two different methodological approaches:

- 1 If the time to compute solutions to both models is rather short, a sequential calibration approach can be applied where a model's parameters are chosen such that the relevant model part matches the outcome of the previously solved model. The process is iterated until convergence of the relevant variables. In our context, the CGE model is calibrated to the transaction cost computed by the ABM, and then the ABM model is calibrated to the market variables from the solution of the CGE model. At the end of the process, market variables, as well as transaction costs, will essentially be the same in both models. This approach has the advantage that the representation of transaction costs in the CGE model only needs to provide a first order approximation to the transaction cost behaviour of the ABM.
- 2 In the situation where the computation to achieve model solutions is longer, it is desirable to calibrate the CGE model to a transaction cost surface, based on data from a series of ABM runs under the relevant range of market conditions. This allows the separate application of the CGE model for different market and policy scenarios, while approximately observing the endogenous change of transaction costs from the ABM model. In this case, a rather flexible functional form for the representation of transaction costs in the CGE model is required to provide an approximation over a larger domain.

A first order approximation of transaction costs for the iterative approach could be achieved using the classical Armington approach, distinguishing trade flows from the local market by destination. However, the typically applied homothetic functions (CET) would have to be changed to more general non-homothetic ones; this is because only those could be calibrated to zero trade flows without excluding a switch to positive trade flows in the model run. As corner solutions are generally a likely outcome of the considered transaction cost, this is an important model feature. If a model specification is desired (that calibrates to a wider transaction cost surface), flexible price transmission functions between markets are preferable. Here, a large range of different functional forms might be appropriate, and the choice should depend based on best fit criteria relative to the data generated by the ABM experiments.

Currently it seems likely that the ABM model will generate solutions with acceptable computational effort, thereby allowing for the iterative approach. Consequently, we will first apply the Armington approach in the context of a simple three market level model, with a very limited number of agents, to test the performance of the iterative process. Each of the agents represents a range of firms/households that supply/demand a certain good/factor/intermediate in a certain region. At the outset, neither migration of households nor investment of firms nor the policy sector was considered. Transaction costs arise on the good as well as on the labour market as search, control and monitoring costs.

Professor Eric Linhart (University of Kiel, Germany) presented his paper “*Elite networks and local government performance in rural regions*”. This part of the project deals with the question, what impact network building has on political and economical outputs. Leading questions are: How can we observe network building? Under which conditions does network building have positive effects? And how can we measure the influence of networks on outputs? To find answers to these questions we use an exchange theoretical approach, and distinguish between exchange networks within political boards or committees, and between influence networks between political actors and interest groups. In the following subsections, we describe our conceptualisation with respect to these two kinds of networks. As we are at the beginning of this part of the project, we can only present the conceptualisation, not analytical results, at this point of time.

Exchange networks

Package building and logrolling are well-known methods to overcome gridlock situations in political boards. The logic behind these concepts is as follows: in a number of bargaining issues, only a minority likes to change a common status quo and install a new policy. The majority prefers to block the incentives of this minority. However, although the actors of the minority group are smaller in size, their utility gain with regard to a new policy exceeds the other players utility loss, so that, all in all, a new policy is the better alternative with respect to the aggregated utility of all actors. If these minority groups vary in their composition from issue to issue, they are able to trade votes and thus form a majority with respect to the whole negotiation system. Table 1 illustrates this phenomenon; it is a simple example with two issues and three actors.

Table 1: Incentives for vote trading

Issue I	A1	A2	A3
u(sq)	0	0	0
u(new policy)	5	-1	-1

Issue II	A1	A2	A3
u(sq)	0	0	0
u(new policy)	-1	4	-1

It is easy to see that, in both issues, there is a majority preferring the maintenance of the status quo. The majorities consist of the players A2 and A3 in issue I, and of the players A1 and A3 in issue II. Assume now that the players A1 and A2 are able to build an exchange network and support each other in the installation of a new policy, they can overrule player A3 and overcome the status quo. Both actors are then able to raise their utility from 0 to 4 (A1) or 3 (A2), respectively. The sum of all players' utility values increases from 0 to 5.

It is notable, however, that exchange networks can also have negative effects in terms of collective utility. Just replace both values -1 for actor A3's evaluation of the new policies by -8 and we see that the existence of an exchange network between A1 and A3 reduces the utility sum from 0 to -9.

The rest of the story is told within one sentence: Knowing the actors' preferences and the outputs of their decisions, we can identify if exchange networks did play a role or not by comparing the output with the computed outputs with and without exchange. And we can say if vote trading had (or would have had) positive or negative effects for the collective of players.

Influence networks

We are further interested in the question of whether informal influence networks between interest groups and political actors exist, via which interest groups can exert pressure on the deciding actors. If this was the case, interest groups would have something like informal power, although there is no formal voting power. To discover if there are such

influence networks which give informal power to interest groups, we can again compute two (or more) simulations of an exchange model, based on different assumptions concerning power distribution. When we compare the models' results with the real world output, we can draw conclusions as to which was the most realistic assumption and how powerful interest groups are. Here again, we can compare the summed up utility values of the various models, and gain impressions as to whether political actors were only interested in maximising their own utility, or if the preferences of citizens and interest groups also played a role.

GENERAL DISCUSSION AND CONCLUSIONS

A general discussion was opened by **Dr Peter Wehrheim** (DG-AGRI). Dr Wehrheim congratulated project members for their very successful efforts in establishing a link between advanced evaluation theory and its empirical application (Poland and Slovakia). Considering significant budgetary outlays for EU RD programmes in the programming period 2007-2013 (approximately € 78 billion) all efforts aimed at the improvement of evaluation techniques leading to more effective and efficient programmes seem to be well justified.

An important issue remains the reduction in the complexity of RD programmes, and a better understanding of complex causalities between financial inputs and impacts of RD programmes. Although quantitative evaluation analyses using more advanced techniques are necessary and very promising, there were many reasons why these techniques had not yet found wider applications. The most important reason was insufficient knowledge among external evaluators of RD programmes regarding quantitative evaluation techniques and modelling. **Mr Wehrheim** stressed that training provided to external evaluators and managers of RD programmes has to be seen as a crucial element of a capacity building process, and the University of Kiel is in a very good position to play a leading role in this respect.

However, the diffusion of presented evaluation techniques will probably also be strongly linked to the extent to which those techniques will be made "user friendly". While resources of public administration dealing with evaluation of RD programmes shrink continuously, the demand for evaluation techniques enabling the provision of

quantitative and qualitative answers to all CEQ moves in the opposite direction, i.e. it grows. In this respect there is a need to extend the old traditional sectoral approach to rural economy by incorporating into it new elements, especially those focused on institutional framework and institutional capacity building.

The following open discussion focused on topics linked to the empirical application of proposed advanced methodologies, and, in particular, underlined the necessity of producing relevant background training materials within this project (e.g. self-contained manuals), which would enable external evaluators to make use of the outputs of this project. The discussion also showed that a strong reliance on traditional *ex-ante* and *ex-post* evaluation methodologies (after vs. before) without making appropriate links to baseline (i.e. comparisons with an adequate control groups), can be very misleading and can result in a poor quality outcome that cannot be defended either from a methodological or empirical perspective. Though the general responsibility for programme evaluation is confined to national authorities, the Commission will play an active role in establishing the common framework for a continuous improvement in the quality of evaluations.

References:

Berger, Mark C., Blomquist, Glenn C. and Sabirianova, Klara P. "Compensating differentials in emerging labor and housing markets: estimates of quality of life in Russian cities." Institute for the Study of Labour, Discussion Paper Series, IZA DP 900, 2003.

Dasgupta, P. Human Well-Being and the Natural Environment Oxford: Oxford University Press, 2004.

Dehejia, Rajeev H. and Sadek Wahba, "Propensity Score-Matching Methods For Nonexperimental Causal Studies," *The Review of Economics and Statistics*, February 2002, 84 (1), 151–161.

Douglas, Stratford and Wall, Howard J. "Voting with your feet' and the quality of life index: A simple non-parametric approach applied to Canada", *Economics letters*, 1993, 42, (2-3), pp. 229-236.

Douglas, Stratford M. and Wall, Howard J. "The revealed cost of unemployment", Review Federal Reserve Bank of St. Louis, 2000, 82(2), pp. 1-10.

Forstner, B. and Plankl, R. "Mid-term evaluation of single farm assistance under the Regulation (EC) No. 1257/1999: methodological problems, data shortcomings and possible solutions". 2004, FAL, Braunschweig.

Frey, B. S. and Stutzer, A. Happiness and Economics: How the Economy and Institutions Affect Well-Being, Princeton and Oxford: Princeton University Press, 2002.

Fuguitt, Glenn V. "The nonmetropolitan turnaround" *Annals of the Review of Sociology*, 1985, 11, pp. 259–280

Greenwood, Michael J. "Internal migration in developed countries," in Mark R. Rosenzweig and O. Stark, eds., *Handbook of Population and Family Economics*. : Elsevier Science, 1997, pp. 647-720.

Heckman, James, Hidehiko Ichimura, Jeffrey Smith, and Petra Todd, "Characterizing Selection Bias Using Experimental Data," *Econometrica*, September 1998, 66 (5), 1017–1098.

Heckman, James J., Characterizing selection bias using experimental data, Cambridge, MA, 1997.

Heckman, James J and Edward Vytlacil, “Structural Equations, Treatment Effects, and Econometric Policy Evaluation,” *Econometrica*, 05 2005, 73 (3), 669–738.

Holland, Paul. W. “Statistics and Causal inference (with discussion)”, *Journal of the American Statistical Association*, 1986 vol. 81, 945-970.

Imbens, Guido W., “Semiparametric Estimation of Average Treatment Effects under Exogeneity: A Review,” Technical Report, UC Berkeley and NBER 2003.

Kahneman, D. and Krueger, A. B. 'Developments in the Measurement of Subjective Well-Being', *Journal of Economic Perspectives*, 2006, 20(1): 3-24.

Kaufmann, Peter; Stagl, Sigrid; Zawalinska, Katarzyna; Michalek, Jerzy, "Measuring quality of life in rural Europe - A review of conceptual foundations ", *Eastern European Countryside*, 2007, 13, pp.1-21.

Layard, R. 2006 'Happiness and Public Policy: a Challenge to the Profession', *The Economic Journal* 116(510): 24-33.

Lechner, Michael, “Program Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labor Market Policies,” *The Review of Economics and Statistics*, May 2002, 84 (2), 205–220.

Liu, Ben-chieh. "Quality of the life indicators: A preliminary investigation." *Social Indicators Research*, 1974, 1, pp.187-208.

Michalos, Alex. C. "Migration and the quality of life: A Review Essay," In: Michalos, A. C. (eds.), *Essays on the quality of life.* , 2003, Kluwer Acad. Publ.: Dordrecht [u.a.],pp. 89-123.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A. and Giovannini, E. 2005 'Handbook on Constructing Composite Indicators: Methodology and User Guide' OECD Statistics Working Paper, Paris: OECD.

OECD 2002 'Aggregated Environmental Indices - Review of Aggregation Methodologies in Use', Paris: Organization for Economic Co-operation and Development.

Rosenbaum, Paul R. and Rubin, Donald B. "The central role of the propensity score in observational studies for causal effects." *Biometrika*, 1983, 70(1), pp. 41-55.

Sen, A. 1979 'The Welfare Basis of Real Income Comparisons: A Survey', *Journal of Economic Literature* 17: 1-45.

— 1985 'Well-Being, Agency and Freedom: The Dewey Lectures 1984', *Journal of Philosophy* 82: 200-1.

— 1993 'Capability and Well-being', in M. Nussbaum and A. Sen (eds) *The Quality of Life*, Oxford: Clarendon Press.

Schultz, T. Paul. "Lifetime Migration within Educational Strata in Venezuela: Estimates of a Logistic Model ", *Economic Development and Cultural Change*, 1982, 30(3), pp. 559-593.

Smith, Jeffrey, "A Critical Survey of Empirical Methods for Evaluating Active Labor Market Policies," *Schweiz. Zeitschrift für Volkswirtschaft und Statistik*, 2000, 136 (3), 1–22.

Smith, Jeffrey and Petra Todd, "Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?," University of Western Ontario, CIBC Human Capital and Productivity Project Working Papers 20035, University of Western Ontario, CIBC Human Capital and Productivity Project 2003..

Tiebout, Charles. "A Pure Theory of Local Expenditures," *Journal of Political Economy*, 1956, 64(5, Oct.), pp. 416-424.

Van Praag, B. M. S. and Frijters, P. 1999 'The Measurement of Welfare and Well-Being: The Leyden Approach', in D. Kahneman, E. Diener and N. Schwarz (eds) *Well-Being: The Foundations of Hedonic Psychology*, New York: Russell Sage.

Zhong Zhao, "Sensitivity of Propensity Score Methods to the Specifications," *IZA Discussion Papers* 1873, Institute for the Study of Labor (IZA), 2005 revised



THE SIXTH FRAMEWORK PROGRAMME

“Development and Application of Advanced Quantitative Methods to *Ex-ante* and *Ex-post* Evaluations of Rural Development Programmes in the EU”

– Mid-term Workshop –

Date: 17th January, 2008

Place: Hanse-Office

Avenue Palmerston 20, 1000 Brussels

Partners involved:

1. University of Kiel: *Institute of Agricultural Economics and Institute of Statistics and Econometrics*, Kiel, DE – Project Coordinator
2. University of Bonn: *Chair of Economic and Agricultural Policy of the Institute for Agricultural Policy, Market Research and Economic Sociology*, Bonn, DE
3. University of Sussex: SPRU: *Science and Technology Policy Research*, Brighton, East Sussex, UK
4. University of Utrecht: *Department of Sociology / ICS*, Utrecht, NL
5. University of Mannheim: *The Mannheim Centre for European Social Research*, Mannheim, DE
6. Institute of Agricultural and Food Economics, Warsaw, PL
7. Research Institute of Agricultural and Food Economics, Bratislava, SK
8. Institute of Rural and Agricultural Development, Polish Academy of Sciences (IRWiR PAN), Warsaw, PL
9. Europrojects-LBV, Kiel, DE



Programme for the Mid-term Workshop in Brussels, 17th January 2008

9.00-9.05	Arrival and welcome, Hanse Office
9.05-9.20	Opening remarks <ul style="list-style-type: none"> • Government of Schleswig-Holstein • Dr H-J. Lutzeyer, European Commission
9.20-9.50	Evaluation of EU rural development policies: Expectations from the viewpoint of policy makers <ul style="list-style-type: none"> • Mr Guido Castellano, European Commission
9.50-10.20	Introduction to the Project ADVANCED-EVAL <ul style="list-style-type: none"> • Professor Christian Henning (University of Kiel, DE)
10.20-10.30	Coffee break
10.30-11.00	Evaluation of SAPARD: Experiences in the Slovak Republic <ul style="list-style-type: none"> • Dr Gejza Blaas, RIAFE Evaluation of SAPARD: Experiences in Poland <ul style="list-style-type: none"> • Professor Klank, IRWiR
11.00-11.30	Application of advanced methods to <i>ex-post</i> evaluation of RD programmes <ul style="list-style-type: none"> • Dr Jerzy Michalek (University of Kiel, DE)
11.30-12.00	Advanced methods of <i>ex-ante</i> evaluation <ul style="list-style-type: none"> • Professor Christian Henning (University of Kiel, DE)
12.00-12.30	Discussion
12.30-13.30	Lunch break
13.30-14.20	Conception of Quality of Life in rural areas <ul style="list-style-type: none"> • Dr Peter Kaufman, University of Sussex Measuring of Quality of Life (QoL) and application of QoL Index to evaluation of RD and structural programmes <ul style="list-style-type: none"> • Dr Jerzy Michalek (University of Kiel, DE)
14.20-14.40	Discussion
14.40-14.50	Coffee break
14.50-15.50	Social networks and rural development <ul style="list-style-type: none"> • Professor Vincent Buskens (University of Utrecht, NL) <i>Social Capital and labour market performance in rural regions</i> • Dr Volker Saggau (University of Kiel, DE) Information networks, regional knowledge spillovers and regional economic growth: An agent-based approach • Professor Thomas Heckelei (University of Bonn, DE) <i>Business networks and regional transaction costs: An extended CGE-approach</i> • Professor Eric Linhart (University of Kiel, DE) <i>Elite Networks and local government performance in rural region</i>
15.50-16.15	Discussion
16.15-16.30	Coffee break
16.30-17.20	Towards advanced policy evaluation: discussion opening <ul style="list-style-type: none"> • Dr Peter Wehrheim, European Commission
17.20-17.30	General discussion
	Closing remarks