Management
Journal of Management
2013, № 1(22)
ISSN 1648-7974

SOME ISSUES IN INNOVATION POLICY EVALUATION

Artūrs Prauliņš
Latvian State Institute of Agrarian Economics

Annotation
Nowadays, innovations facilitate both economic development and growth in national welfare. In most countries politicians are determined to improve the effectiveness of national innovation policies and to help entrepreneurs to be innovative and competitive, both in the national and the international market. The accountability culture, budget limitations and dedication to spending money as efficiently as possible are among reasons for developing an evaluation of innovation policy. Moreover, evaluations can assist decision-makers with identifying sources of competitive advantage and areas where public expenditure can reduce market fluctuations. The article explores the diversity of methods and approaches to the evaluation of innovation policy and traces their gradual transformation and historical development. Particular attention is paid to the hierarchical structure and temporal perspectives of evaluation as well as to the comparative analysis of different types of evaluation. Various sets of optional evaluation indicators are also examined. It has been concluded that a demand-driven innovation policy and diffusion-oriented measures lead to the use of more advanced and experimental evaluation methods going beyond previous practice. The design of an integral evaluation system is a challenging task due to the increasing complexity of evaluation objects and the impact of numerous factors outside of the evaluation model. Because of the politicians' particular interest in simple economic indicators and the accepted measures of the short-term impact of innovation policy, a conflict with evaluation experts arises. A systemic analysis and a benchmark approach could be used to mitigate this conflict. Recently, an adaptive learning approach has been considered to be the state-of-the-art method for evaluating innovation policy. In addition to exact measures, it could provide information for the recurrent implementation of successful innovation policy instruments and the cancellation of ineffective ones.
KEYWORDS: evaluation, indicators, innovation policy, methods, research and development.

Introduction

Nowadays, innovations are considered as one of the most important prerequisites of continual economic development and the growth in national welfare. It is not surprising that, in most countries, politicians are determined to improve the effectiveness of national innovation policies and to help entrepreneurs become more innovative and competitive, both in the national and the international market. This mission could be accomplished if the performance of present innovation policies is explicitly assessed and continuous improvements are implemented. The accountability culture, constrained budgets and necessity to spend money as efficiently as possible have been mentioned as obvious reasons for developing evaluation of innovation policy, which is an important tool of justifying or re-directing funding (Edler et. al. 2012). Moreover, evaluations can assist decision-makers with determining sources of competitive advantage and areas where public expenditure can reduce market fluctuations (Loikkanen et. al. 2006).

It is necessary to stress that some ambiguity still surrounds the definition of innovation. A number of dictionaries are extremely concise and define innovations as the development of something new (Statt 2003, p. 79; Statt 2004, p. 73). The opinion is expressed (Macmillan Dictionary 1986, p. 204) that innovations are often used as an alternative to 'inventions' and „cover both technological advances in production processes as well as the introduction of different attributes and attribute combinations in marketable products". However, an alternative point of view is that there is a fundamental distinction between these economic categories because working under laboratory conditions does not mean the same as working under commercial conditions (Freeman 1987). In other dictionaries which define innovation the time span is emphasised: „use of a new product, service, or method in business practice immediately subsequent to its discovery" (Dictionary of Business 2000, p. 338). The practical application of innovation has also been stressed: "innovation is the practical refinement and development of an original invention into a usable technique (process innovation) or product (product innovation)" (Pass et. al. 2005, p. 258). Some definitions are more focused on the commercial aspect of innovations. For example, "[..] the technical, industrial and commercial steps which lead to the marketing of new manufactured products and to the commercial use of new technical processes and equipment (Gilpin 1973, p. 109) or "putting new products and services on to the market or new means for producing them" (Bannock et. al. 2003, p. 191). In addition, the importance of changes caused by innovation has been highlighted ("significant step forward in technology, management, thinking etc.; something more than a mere extension or improvement of present methods") and their connection with entrepreneurial ability has been accentuated (Moffat 1983, p. 160).

The strategy and rationale are considered as the initial stage of any future-oriented innovation policy. Firstly, the reasons for public intervention in innovation process and financial support should be found. Secondly, the best way of implementation and enforcement of innovation policy, which would result in the highest possible output, should be specified. Finally, the evaluation of innovation policy should be carried out and proposals for improvements for more effective performance should be presented. Consequently, according to T. Loikkanen *et. al.* (2006), this process could be divided into the following stages:

the definition of rational of innovation policy,

the formulation of policy strategy and elaboration of a consequent action plan,

the implementation of policy using various instruments and measures,

the assessment of socio-economic impacts of policy, and the subsequent improvement.

Evaluation could be characterised as a social process due to the synergy of individuals, organisational methods, practices and routines (Papaconstantinou et. al. 1997). Both the singularity and complexity of evaluation was stressed by J. Edler *et. al.* (2012) who wrote that "Evaluations are unique, idiosyncratic exercises. Each evaluation has its specific political, stakeholder, and policy context, with specific requirements defined by those commissioning it and depending on its role in the policy cycle. The goals of evaluations differ and with this the dimensions that are covered and the methods that are used". Other researchers (Gans et. al., 2006; Molas-Gallart et. al. 2005; Smith 2006) concur with this view and note that the evaluation of innovation policy could be challenging due to various goals of the policy, its intricacy, implicit and complicated relations between policy input and output as well as its long-term impact. Therefore, R. Linnas (2008) is of the opinion that before the assessment and evaluation of any policy, including innovation policy, it is vitally important to analyse various factors such as the available and required inputs, the preconditions, obstacles and environment, which are followed by the implementation of the policy in question.

The **object of the research** is approaches to evaluation of innovation policy.

The **aim of the research** is to explore the diversity of approaches to the evaluation of innovation policy and to trace their gradual transformation and historical development.

The **tasks of the research** are:

1. to explore the historical transformation of methods of innovation policy evaluation;

2. to investigate the hierarchical structure, types and modern approaches to innovation policy evaluation;

3. to analyse the optional indicators for assessment of results of innovation policy implementation.

The novelty of the article is a critical comparative analysis of various approaches to innovation policy evaluation as well as assessment indicators, inter alia, for the use in emerging and transition economies. To accomplish the tasks mentioned above, the following research methods were used: literature review, comparative and logical analysis and synthesis.

Historical development of approaches to evaluation

During the years after World War II, the scientific community aspired to not being involved in any social processes and preferred developing "pure" science or doing "basic research" unaffected by any social objectives and not being responsible for the possible use of scientific findings. The misuse of scientific achievements for military purposes in World War I and World War II could be mentioned as one of the main reasons for such segregation. A more balanced point of view is that science embraces two main modes of knowledge production (Arnold 2004). Unlike Mode 1, which is oriented towards disciplinary research within "basic science", Mode 2 contains both the applied science in university and research institutions and the development of knowledge in society (Table 1). As Mode 2 is closely connected to society, in addition to traditional quality control by peer review, it is subjected to social and economic evaluation. The modes also differ widely about the raising of funding. In general, Mode 1 is financed by funds allocated to university research using scientific council traditional procedures. Their evaluation methods have both merits and flaws. Transparent quality control can be mentioned as a good point whereas intellectual and institutional conservatism, reluctance to make changes and focus entirely on the scientific essence of the project or programme are examples of weaknesses. Sources of finance for Mode 2 are more varied because partly the resources are obtained form potential users of research results.

Table 1. Modes of knowledge production in science

Mode 1	Mode 2
Problems set and solved in the context of (academic) concerns of the research community	Problems set and solved in the context of application
Disciplinary	Transdisciplinary
Homogenous	Heterogeneous
Hierarchical, tending to preserve existing forms of organisation	Heterarchical, involving more transient forms of organisation
Internal quality control	More socially accountable quality control

Source: Arnold (2004, p. 8)

Up until the 1970s, the state agreed with the relative autonomy of science and authorised the scientific council to evaluate the quality of innovation and research projects and programmes and allocated the funding. However, this concept was undermined by the use of results of scientific research in the Vietnam War and the emergence of the strategic research paradigm. The paradigm stressed that basic research should result in a broad base of knowledge to find a solution to present and future problems. As result, research councils gradually modified the principles of finance appropriation and devoted much more attention to responsive-mode funding (Arnold 2004).

Unlike education and social policies, the evaluation of innovation policy was started rather late. The first evaluations were made in the United States at the end at the 1960s (Roessner 2002), but the first evaluation took place in Europe approximately ten years later. There were a lack of systematic approaches to the evaluation procedure and few general guidelines. Nevertheless, ex-ante assessment using peer review was quite popular and widespread prior to the development of other evaluation methods (Papaconstantinou et. al. 1997). In the following decades the situation changed dramatically because of the modern public management systems paying more attention to tight control of outputs as opposed to monitoring processes and inputs. This required the elaboration of appropriate evaluation indicators for performance and output measurement. In the 1980s, the further development of demand-pull theories of innovation and awareness of the links between the scientific community as a generator of new ideas and society, as a consumer, further intensified the search for modern evaluation methods and measures (Arnold 2004; Dudzevičiūtė 2011).

The main feature of innovation policy during these decades was the emergence of large-scale collaborative Research & Development programmes (hereinafter R&D) at national and international level. During the 1990s, the further development of evaluation theory and practice occurred with special focus on assessment of the EU Framework Programmes, EURECA Initiative and US Advanced Technology Programmes. Although the general conclusions about the effects of the abovementioned programmes were positive, due to the non-linear model of innovation progress which was described and analysed by N. Rosenberg and S. Kline in the early 1980s (Smith 2006). The evaluations were unable to provide information about return-on-investment. Politicians considered this indicator as extremely important. Firstly, it helped ascertain the positive return and expediency of further investments in innovative programmes. Secondly, the use of cost-benefit criteria provided policy-makers with information about relative profitability of different supported R&D areas. Instead, the evaluations demonstrated that, in addition to competitiveness and market effects (for example, growth in sales or reduction in production costs), the socio-economic effects of the innovation programmes can be evaluated using such individual and organisational learning effects as partnerships, networking, contribution to skills and research (Georghiou 1997; Molas-Gallart et. al. 2005; Papaconstantinou et. al. 1997; Powell et. al. 2006).

Since the 1980s the proliferation of different innovation system theories has shown that the innovation process is a collaborative interaction between users, producers, companies and different inputs to innovation. This development required the elaboration of appropriate evaluation techniques, which are quite dissimilar to methods used for the evaluation of other areas. The development was complicated by peculiarities of the assessed objects and dissimilarities between techniques which could be used for the evaluation of separate projects and social networks (Christensen 2012; Powell et. al. 2006; Bannock 2003).

Hierarchical structure and temporal perspective of evaluation

Traditionally, methods of conventional evaluation concentrated mostly on company level effects (Lenihan 2011). However, at present this approach is out of date as it ignores vast societal impacts and does not use appropriate performance measures of innovation programmes. Unfortunately, studies thus far have not given a definitive answer about what these impacts are (Christensen 2012). It has been discovered that the balance between traditional evaluation aspects depends heavily on the level where the assessment is carried out. At high levels of innovation policy system, the appropriateness is always analysed, whereas at lower system levels, the problems of implementation are resolved. Consequently, great stress is placed on impact and effectiveness. Evaluations at high levels are very likely to lead to remarkable changes in policymaking such as changes in national research policy resulting from the evaluation of the activities of the national research council (Arnold 2004).

The main conclusion to be made is that indicators for a description, analysis and evaluation of complex innovative systems are beyond the scope of indicators previously used in assessments. However, the evaluation still has to meet the requirements of appropriateness, impact and effectiveness. At the same time it should focus more on the impact at the aggregated level, level of programme portfolio and overall assessment of innovation system health (Arnold 2004; Georghiou 1997). This position is supported by K. Flanagan *et. al.* (2010), who are of the opinion that a new evaluation object should be policy mix and not individual initiatives.

According to E. Arnold (2004), nowadays the evaluation of research and innovation policy should be carried out on three levels. Firstly, a bottom-up element of evaluation concerned with the traditional evaluation of individual programmes. Secondly, a top-down element of evaluation dealing with the analysis of the thorough health of innovation system and doing more detailed analysis of research and innovation. For the suggested checklist of the optional criteria for the overall assessment of the health of innovation system and its performance, see Table 2. Finally, the evaluation of sub-systems or bottleneck analysis is carried out at meso-level in order to examine the systemic role of institutions, clusters and other elements. In literature these systems are also known as the system models, also known as the fourth generation of innovation models (Marinova et. al. 2003). For J. L. Christensen (2012), these levels are micro-level, meso-level and macro-level. The recommendation to use this multilevel approach is also made in other studies (Public Financial 2006).

Table 2. Optional criteria for the overall assessment of the health of innovation system

INSTITUTIONAL BLOCKS	
Business or "production" sector	Adequacy of resources and absorptive capacity Adequacy and performance of individual supply chains, clusters and sectors Levels of innovativeness, R&D performance and innovation activities Economic competitiveness
Education, research and intermediary structures (knowledge infrastructure)	Capacity and quality in research and education Participation in higher education and research training Presence of critical mass in key areas Speed with which capabilities can be (re)aligned with changing social and technical needs Strategic/managerial performance, especially in change management Effectiveness of interfacing with other parts of the innovation system
Political system	Effectiveness of policy intelligence and analysis functions Research and innovation policies; policy mix Effectiveness of institutional structures and the division of labour in implementing R&D and innovation policies Adequacy of governance
Infrastructure	Adequacy of provision Effectiveness and efficiency of operation
Framework conditions	Consistency of regulatory frameworks and their implementation with innovation and change Friendliness of tax and fiscal regimes to research and innovation Economic, social and cultural rewards and disincentives to innovation and entrepreneurship
Demand	Receptivity of buyers to innovation Harshness of the selection environment provided by domestic markets Extent to which markets are "leading edge" and able to stimulate supplier learning
CONNECTIVITY	
General connectivity	Extent of co-operation and networking across institutional boundaries, e.g. business-business networking
Knowledge connectivity	Extent of networking in knowledge production and use, and to which knowledge flows freely among parts of the innovation system
KNOWLEDGE AND CAPACITY	
Economic and technical capabilities	Adequacy of economic, technical and change capabilities across all institutions Adequacy of levels of education and training
Public understanding of science	Adequacy of general levels of understanding of research and technology to making good decisions about these, especially at a political level, but also more widely

Source: Arnold (2004, p. 15)

Sometimes it is problematic to differentiate between the abovementioned levels because national networks of research institutions can be alternatively classified as bottom or middle level (Powell et. al. 2006). Evaluations of systems and sub-systems on higher level can not usually be carried out in detail because of the intricacy and scale of the assessment objects. The possible solution is the application of such analytical tools as peer review of portfolio or benchmarking leading to aggregated judgements. At the same time this evaluation becomes less dogmatic and scrupulous because of its incompleteness. Thus, the elaborated evaluation system for assessment of innovative programmes at high levels should provide the opportunity to base the evaluation on incomplete sets of testimonies including elements of systemic and micro evaluation. A growth in the importance of socio-economic methods including scientometrics, surveys, case studies, best practice and cluster analysis is consistently reported to take place at higher system levels (Arnold 2004).

The temporal focus of evaluation has changed dramatically too. Whereby a traditional policy evaluation was previously focused on the immediate, measurable and direct effect within a certain period of time, the results are now expected to be much broader. Although "behavioural additionality", explained by T. J. Buisseret *et. al.* (1995), is a relatively young concept, it is a classical example of broader impact in innovation policy. The research can be started before the funding is received, finished afterwards and can contain many more activities compared to those stipulated by a contract. It has been emphasised, (Loikkanen et. al. 2006), that the segregation of impacts of innovation policy according to their temporal aspect (short, mid and long-term) is of vital importance for more precise analysis (Table 3).

Approaches and types of evaluation

Traditionally there were two distinctive basic approaches to the evaluation of innovation and research and development. Scientific peer review is the oldest and has various forms including *ex ante* and seldom *ex post*. The other is the analysis of publications including impact analysis and citation count (Papaconstantinou et. al. 1997). The assessment was usually carried out at the level of individual project, research or published paper. However, evaluation of programmes using peer review and other advanced assessment methods deals mostly with the allocation of resources and improvement in the quality of research, and not with the prioritisation of different research areas. As long as the peer review process is developed and supervised by the scientific community, it has a tendency to be impacted by current theories ("normal science" described by T. S. Kuhn (2012) as opposite to "revolutionary science"). At meso-level and macro-level this dominance could put obstacles to the application of non-scientific measures (for example, resource availability, cultural importance, socio-economic relevance).

Table 3. Outputs and impacts of innovation policy according to their temporal aspects

Short-term	Mid-term	Long-term
# Award-participant characteristics # R&D partnering # Acceleration of R&D # Development of innovative technology (patents, publications, competitiveness, prototype products and processes)	# Commercial activities (new products and processes, licensing) # Attraction of capital # Strategic alliances # Company growth	# National economic benefits ~ Return on investment ~ Diffusion (increase in GDP and collected taxes, societal impact)
Completion of project		
	Post-project period	

Source: Loikkanen (2006, p. 10)

Thus, due to disproof of the linear model of innovation development and a growth in popularity of "new public management" theory during the last 30 years, the innovation funding was subjected to both *ex ante* and *ex post* assessment using various measures in addition to quality criteria (Arnold 2004).

At present three different approaches to the evaluation of innovation policy are recognised. According to S. Kuhlmann, the use of a *formative approach* to the evaluation of innovation policy was caused by the development of a systematic approach to the policy elaboration. It is presumed that separate policy initiatives should be interrelated and become an integral part of a portfolio aimed at the performance of policy at various levels. A classical example of this integrity is a connection between knowledge creation, innovation and different actors involved in this process. Within formative context, the evaluator is a facilitator rather than an external expert. Formative approach does not focus extravagantly on the assessment of policy impact. Instead, it tries to involve policy-makers in the evaluation process. At present, the formative approach has many supporters who are of the opinion that prior ex-post and "hands-off" assessment carried by external experts does not suit any more the needs of modern budgetary control and society. Consequently, evaluation has to become an organic part of policy implementation and be carried out constantly in order to ameliorate the imminent imperfection of policy process (Molas-Gallart et. al. 2005; Papaconstantinou et. al. 1997).

Systemic analysis was proposed by E. Arnold (2004) and embraces a multilevel system of assessment of innovation policy including acknowledged evaluation of systems and programmes, the systemic analysis of innovation system viability and additional analysis of bottlenecks. Special attention is devoted to institutional environment, collaboration and networks within institutional boundaries from the aspect of knowledge, economy and technology. It is stressed that systemic analysis is oriented towards the evaluation of policy portfolios and the connection between evaluation and the complicated procedure of policy development (Molas-Gallart et. al. 2005).

The *benchmark approach* suggests sets of easily communicative measurements whose elaboration is based on the systemic perception of the innovation and technological novelties. It is believed that the benchmark approach can evaluate whether the innovation policy has anticipated positive influence on competitiveness and employment. In addition, this approach enables several dimensions to be evaluated simultaneously, for example, such dimensions as social and human capital, research capacity, technological and innovative performance and absorptive capacity. Another example is an analysis of the interaction between organisations (companies, research institutes etc.), incentive structures and institutional arrangements (Molas-Gallart et. al. 2005).

Despite obvious merits, the benchmark approach could be difficult to apply due to some drawbacks. Firstly, the considerable number of system components complicates the identification of correlation between R & D as inputs and competitiveness and employment as outputs. Secondly, the precise assessment of innovation policy impact on overall performance is almost impossible to carry out because of the unpredictable reaction of the private sector to public policies. Finally, due to the uniqueness of the environment where the innovation policy is implemented, the same political instruments can result in different outcomes. Although the calculation of correlation between government programmes and changes in measures of innovative performance can provide a general insight into programme application effectiveness, it can not answer a question about the effectiveness of specific policy mix elements. Thus, the main conclusion is that the use of the benchmark evaluation method does not result in an unequivocal assessment of the impact of innovation policy or mix of its elements. Consequently, the benchmarking is more suitable for defining policy. However, in practice it is applied to make a comparison of national innovative performance if just a few of the most common indicators are considered (Molas-Gallart et. al. 2005; Papaconstantinou et. al. 1997).

Theoreticians (Edler et. al. 2012) recognise several types of evaluation, namely supporting, verdict and holistic. The supporting evaluation is mostly formative, planned beforehand and is a part of the measure cycle. Its main focus is on the consistency, coherence and efficiency of the programme. The "verdict" evaluation is diametrically opposed to a supporting one as it is summative, generally ex post and not integrated in policy cycle. Verdict evaluations are usually made for the purpose of justification and accountability. Instead of taking into account the context of the programme, they concentrate on economic impact and input additionality. The third type integrates formative and summative features. Although it is planned, it deals with a much vaster range of topics compared with the traditional formative approach. In addition, it measures goal attainment, effectiveness and other facets of impact as logical parts of the entirety. Common methods used for carrying out holistic evaluations are case studies, network analysis and group comparisons.

According to analysis of statistics on evaluation of innovative policy instruments in EU25 countries from 2002 to 2007 (Edler et. al. 2012), most evaluations comprise both formative (learning and improvement) and summative (judgemental) components. Just a third of the evaluations is entirely formative and a fifth is completely summative. Recommendations made in evaluation reports seldom result in such considerable changes as termination or redesign of programmes because usually it is the consequence of general political decisions. However, high quality positive evaluations could influence the decision to expand or lengthen the programme. Summative evaluations are found to be more intensively discussed by policy makers and able to cause terminations, redesign or expansion of the programme, whereas formative evaluations cause less remarkable changes. Presumably, results that are quantitative, easy to communicate and easy to perceive have a greater impact on the future of evaluated programmes. In general, the more intensively the evaluation reports are discussed by the government and stakeholders, the higher the probability that they will make a difference.

Sets of optional indicators

J. Gans *et. al.* (2006) have emphasised that measuring innovative performance is crucially important for the implementation of effective innovation policy and economic development. At the same time the opinion is expressed (Smith 2006) that some aspects of innovation a priori are impossible to quantify and measure. However, general characteristics of innovation and key dimensions could be subjected to measurement. It is believed that the significance of measuring the innovation process results from the interrelation between innovation, genuine enhancements of competitiveness, economic progress and the welfare of society. Thus, innovation indicators are a basic instrument for decision-making both in the private and the governmental sector. In the private sector, they can be used for formulating competitive strategy. In the public sector, they are crucially important for the elaboration and assessment of innovative policies (Lugones 2009). In spite of the popularity of the evaluation of innovation policy, Loikkanen *et. al.* (2006) have noted that no generally accepted or harmonised methodology and definitions have been introduced so far.

The main problem with the development of innovation indicators is the necessity to devise the object being measured, the meaning of the measurement concept and overall usefulness of different types of measurement (Smith 2006). Indicators for the evaluation of innovation policy should meet three general criteria (Lugones 2009), belonging, reliability and comparability. Belonging means the ability of indicators to act as benchmarks when the users (policy-makers and companies) analyse, formulate and design current and potential strategies of innovative development. Reliability characterises the quality of indicators and the tools used in gathering and further processing the appropriate information. Comparability is a prerequisite of comparative analysis of the same object (for instance, sector of national economy or company) from the perspective of time. In addition to the peculiarities of the assessed process, the availability of resources (time, finance and human) for the collection of data used in the evaluation also determines the selection of indicators. One of the studies (Lugones 2009) suggests that all the indicators could be grouped into separate modules, namely innovation strategies, outcomes of innovation activities and appropriateness as well as obstacles to innovation, sources of funding and use of public instruments. The extensive list of variables in the Table 4 provides a better insight into the nature of proposed evaluation methodology. There are also alternative classifications. For instance, the one suggested by K. Smith (2006) includes, firstly, R&D data, secondly, data on patent applications, grants and citations and thirdly, bibliometric data. Moreover, as additional classes of indicators are mentioned - technometric indicators, which analyze the technical performance characteristics of products; synthetic indicators for scoreboard purposes and databases on specific topics created and maintained by individuals and groups for research purposes.

Table 4. Set of optional indicators for innovation policy evaluation

INNOVATION STRATEGIES
1.1. Innovation activities ***Analyzed variables****:* a) Absolute and relative number of innovative enterprises; b) Absolute and relative number of enterprises involved in R&D activities; c) Proportion of spending on innovation activities and R&D in turnover; d) Proportion of total spending on innovation activities in each sector. These variables which are calculated for different sectors and years are used to determine the size and structure of expenditure on innovation activities, for example, internal and external research and development; acquisition of machinery and equipment, technology, hardware and software; contracting consultancy firms and technical assistance; engineering and industrial design activities; staff training and marketing.
1.2. Cooperation ***Analyzed variables:*** a) Purpose (type of innovation – product, process, radical or incremental; objective - R&D, engineering and design, training etc.); b) Duration and degree of formality (contractual agreement, monetary compensation); c) Agent (clients, suppliers, universities, training centres, technological centres, consultants etc.); d) Geography (local or international); e) Obstacles.
1.3. Sources of information for innovation ***Analyzed variables:*** a) Internal sources (marketing, production, distribution); b) External sources (Internet, clients, research centres, journals, consultants, conferences and exhibitions, databases, competitors, suppliers etc.).

1.4. Determining factors in innovative efforts ***Analyzed variables:*** identification of total or partially unsatisfied demand in the market; use of an idea or scientific and technical novelties etc.
1.5. Objectives of innovative efforts ***Analyzed variables:*** a) Reductions in production costs (use of new materials, more effective use of labour, reorganisation of the production process, new marketing strategy or distribution channels); b) Product differentiation (modification of existing products, launch of new products etc.).
1.6. Technological and absorption capacities ***Analyzed variables:*** a) Total number of personnel, composition by education and qualification; b) Average seniority of personnel; c) Existence of R&D department or laboratory, engineering and industrial design department, information technology and systems department; d) Number and qualification of the employees assigned to abovementioned departments and in quality management, full and part-time employment; e) Employee training.
1.7. Quality management ***Analyzed variables:*** control points, follow-up charts, certified processes and products.
OUTCOMES OF INNOVATION ACTIVITIES AND APPROPRIATNESS
2.1. Introduced innovations ***Analyzed variables:*** product (new or significantly improved) and process (new or significantly improved) innovations, innovations in organization and marketing.
2.2. Scope of the innovations ***Analyzed variables:*** novelty for the company, national or international market.
2.3. Impact of the innovations ***Analyzed variables:*** a) Changes in sales and growth in export; b) Changes in productivity (sales-to-employment, sales-to-consumption of energy etc.); c) Proportion of sales from new or improved products in total sales; d) Changes in average wage in company; e) Evolution of average seniority of workers.
2.4. Appropriateness ***Analyzed variables:*** a) Methods of formal protection (brands, patents, industrial design, copyright); b) Methods of strategic protection (control distribution networks, secret, design complexity etc.); c) Number of patents requested and obtained; d) Obstacles to patenting (costs, complexity, administrative difficulties). These variables help to identify a relation between the characteristics of introduced innovations (depth, scope etc.) and a competitive strategy. Therefore, they should be compared with indicators of innovation activities.
OBSTACLES, SOURCES OF FUNDING AND USE OF PUBLIC INSTRUMENTS
3.1. Endogenous and exogenous obstacles ***Analyzed variables:*** a) Endogenous obstacles (shortage of personnel with the necessary qualification in the company, problems in administrative or production organisation, excessively long period of return, uncertainty about the successful introduction of innovations or appropriate protection by patents etc.); b) Exogenous obstacles (shortage of personnel with the necessary qualifications in the labour market, problems with access to the exogenous knowledge, reduced market size and diseconomy of scale, barriers to entry into market, inadequate available infrastructure, inefficient protection of intellectual property, difficulties of access or excessive cost of funding, shortcomings in public policies).
3.2. Sources of funding for innovation ***Analyzed variables:*** a) Inner resources (contributions by shareholders, reinvestment of profit); b) Outer resources.
3.3. Knowledge and use of public instruments of innovation promotion ***Analyzed variables:*** a) Knowledge (absolute and relative number of companies which have knowledge of innovation support programmes); b) Use (absolute and relative number of companies which use funds provided by support programmes); c) Obstacles to obtaining finance (rejection of projects, high interest rates, excessive demand for guarantees, bureaucracy, difficulty in preparing innovation projects etc.).

Source: adopted from Lugones (2009, p. 21-28)

R. Linnas (2008) maintains that a special integral system of innovation policy evaluation comprising different sets of indicators should be elaborated for each of three stages of the process, namely starting stage, running stage and mature stage. For the evaluation of the policy during starting stage, the indicators have to reflect the different types of input, for example, the financing or funding of innovation, foreign capital investments in R & D. These indicators are important for *ex ante* assessment of the available financial resources as well. In later phases, the indicators characterise the dynamics and progress or regress of innovation policy in compliance with programmatic documents. In addition, they help to predict the effectiveness of the policy both in economic and societal terms. The list of appropriate indicators comprises the total amount of funding allocated to the implementation of innovation policy, the amounts spent on innovation related processes and products, the number of individuals (students, academicians, researchers, civil servants, employers and employees) and research institutions involved in innovative activities, the number of patents, trademarks and inventions.

Indicators for the running stage are more various and could be divided into several subgroups. Firstly, these are indicators for the assessment of the overall situation in economics (innovation potential, role of the government, public and private sector etc.). Secondly, indicators to measure changes in environment (economic, political, legal and cultural) where the innovation policy is implemented. Finally, indicators for the evaluation of performance in terms of output could include such measures as the amount of total and direct investments in R & D, the number of patents, trade marks and inventions, improvement in the innovation model, system and processes, the transfer of knowledge, the innovation supporting programmes and projects, the number of innovations in an organisation, process, product or technology, the development and application of technology. Some of these indicators could also be used to measure input.

The set of measures for the mature stage is most important as they show the effectiveness of the process regarding mostly societal impact and less in terms of economy. The list could consist of absolute and relative numbers of volume of exported products, services and technology, level of unemployment, productivity in different sectors of the national economy, energy consumption, popularity of e-services, life expectancy, the human development index, pollution and the availability of fresh water etc.

Recently, in the EU the necessity to analyse the contribution of innovation to the sustainable development has been recognised and the problems of assessment of social and environmental dimensions of innovation policy has been explored (Gjoski 2011). The EU has set a target to adopt much more strategic approach to innovation, when all policy instruments, measure and funding are designed to contribute to it (Europe 2020 2010). It has been stated that progress towards the Innovation Union as one of the seven flagships announced in the Europe 2020 Strategy should be measured by such main indicators as the R&D investment target and a new Innovation indicator. A panel of experts has explored two options: a list of three immediately available indicators (based on patent applications, the contribution of medium-high-tech and high-tech products to the trade balance, and employment in knowledge-intensive activities) and a single indicator (reflecting the successful development and dynamism of innovative entrepreneurial activities). After the analysis of conclusions made by the panel the Commission decided to use the single indicator defined as fast growing, innovative firms.

The analysis of data from the INNO-Appraisal database which contains evaluations of innovation policy instruments in EU25 countries from 2002 to 2007 has revealed (Edler et. al. 2012) that the most popular innovation policy measure is that of direct financial support for innovation activities (52 %), which is followed by such measures as networks and clusters, collaboration and technology or knowledge transfer, innovation management support and dissemination, innovation culture as well as science and industry cooperation (respectively 29 %, 28 % and 27 %). Support for the uptake and diffusion of innovation and mobility of personnel are less widespread (just 19% and 10%). The creation of start-ups and spin-offs as well as development of intermediary bodies and agencies are much less popular (7 %), but indirect measures (tax) are least popular (just 3 %).

Separation of some aspects of evaluation

The increasing complexity of innovation policy programmes and projects led to the separation of certain aspects of the innovation process for assessment purposes. This is true of the evaluation of fiscal incentives for R & D (Corchuelo et. al., 2009; Haegeland et. al. 2007; Hall et. al. 2000). The measurement of the efficiency of fiscal incentives is considered to be an important aspect of modern macroeconomic policy. It is usually evaluated using such indicators as an increase in expenditures on innovation and research induced by favourable tax regime and the consequent introduction of innovative products, processes and technology. The number of enterprises which used fiscal incentive instruments could also be used (Entrepreneural 2003). The studies (Czarnitzki et. al. 2005; de Jong et. al 2007) have shown that fiscal incentives may have a positive impact on the sales or range of innovative products. The positive correlation between fiscal incentives and a growth in productivity was also evidenced (Lokshin et. al. 2007).

The next aspect separated for evaluation purposes deals with institutions involved in implementing technology and innovation policy. This assessment is usually carried out as a part of the evaluation of opportunity-enhancing innovation policies. This type of evaluation focuses on the role of public sector research institutions, their commercial orientation and linkages with innovating companies, which are another popular object of innovation policy measurement (Georghiou 1997). The process of knowledge and technology transmission between research institutions and economic sectors is complicated and consists of various elements including the creation of networks and links. Usually two assessment methods are used. A survey of scientists in order to get information about their contacts with practitioners is followed by interviews with some representatives of industry. Another method is based on scientometric indicators, particularly co-authorship, funding acknowledgements and data and equipment acknowledgements.

Diffusion-oriented policies facilitating the transfer of knowledge and technologies from research centres to industry and especially small and medium enterprises have been a popular evaluation object for quite a long time (Georghiou 1997). Alternative methods could be used to reach a target. For example, in Germany, which is one of the countries where the evaluation of diffusion-oriented policies is widespread, the existing transfer activities are determined and classified in the first stage and in-depth analysis of some selected activities is carried out in the second stage. In Denmark, the evaluation of the research centres takes place every three years and is based on the reports prepared by the centres and reviewed by independent experts.

Conclusions

The evaluation of innovation policy needs constant improvement and modification as the applied methods have to be updated in order to keep up with the rapid development of new, more complicated support programmes and instruments. The necessity to focus on demand-driven innovation policy and diffusion-oriented measures is expected to lead to the use of more diversified and experimental methods going beyond previous practice (Edler et. al. 2012).

Studies have recognised that the design of an integral system for the evaluation of innovation policy is a challenging task. This task is complicated by the fact that innovation policy regulates programmes, projects and processes at different stages of completion. In addition, the results are affected by previous R & D programmes. No single method of creating this system exists. However, one of the prerequisites of successful evaluation is a consideration of the peculiarities of a national innovation system and a match between all the components of the system and the interests of the country and national economy. Moreover, the components should comply with the general trends of evaluation of innovation policy accepted by the European Union and such international organisations as OECD, World Bank and others (Linnas 2008; Papaconstantinou et. al. 1997).

The previous experience of carrying out evaluations of innovation policy has shown that the best results could be obtained if the evaluation is initially elaborated as a part of the assessed programme or policy. Besides that, policy-makers should react to the results of evaluations and should ideally make them available for public discussion. Evaluations should be more focused on meeting the needs of the users because it determines the application of specific methods for yielding the most appropriate results (Papaconstantinou et. al. 1997). Following these recommendations in practice is not always possible. Many politicians are still more interested in simple economic indicators as well as traditional quantitative measures of short-term and direct impact of innovation policy because they assist them in justifying the expenditure on R & D (Papaconstantinou et. al. 1997). Therefore, they ignore the fact that high returns in innovation and research do not always correspond to the rapid growth in GDP and the level of welfare in society (Dimza 2003). The differences in innovation culture also play an important role (Brynteson 2010).

Biased orientation of policy-makers implicitly leads to conflict with professional experts who apparently give preference to more advanced theory-led types of evaluations, so in order to meet the demand for certain evaluation output, experts are compelled to base their analysis on an out-of-date concept of linear development of innovation. Obviously, this tension can not be easily alleviated (Arnold 2004). Hypothetically, a systemic analysis could be a panacea and produce required measures. The benchmark approach is another good example of method synthesising a theory-led systemic approach and measurement. Nevertheless, it fails to identify the direct relationship between a certain policy instrument and outcome (for instance, the impact of an innovation programme on employment or increase in productivity). This limitation is primarily partly due to the discrepancy between the political demand for measurement immediately after finishing innovation programme and actual long-term effects and, secondarily, the existence of numerous other impact factors which are not included in the systemic model because of its anticipated excessive complexity. Consequently, the majority of measurements of innovation policy thus far have been based on linear models (Molas-Gallart et. al. 2005). From the historical perspective, a linear model (or so called "technology push" innovation) represents just the second out of a total of six generations of innovation models developed so far (Marinova et. al. 2003). In theory, the more advanced models that emerged soon substituted the linear models. Despite its apparent flaws, in practice a linear model established a solid ground for the schemes of research financing and nowadays still remains relatively popular.

It is believed (Georghiou 1997) that an adaptive learning approach is the most suitable answer to the latest development trends in innovation policy. Evaluation should give feedback not only on exact measures but also provide information for the recurrent implementation of successful innovation policy instruments and the cancellation of ineffective ones. As long as evaluations are an integral part of the innovation process, well-established institutional framework is a guarantee of unbiased assessment.

Literature

Arnold, E. (2004). Evaluating research and innovation policy: a systems world needs systems evaluations. *Research Evaluation*, volume 13 (1), pp. 3-17.

Bannock, G., Baxter, R. E., Davis, E. (2003). *The Penguin Dictionary of Economics*. Penguin Group, London.

Brynteson, R. (2010). *The Manager's Pocket Guide to Innovation*. HRD Press.

Buisseret, T. J., Cameron, H. M., Georghiou, L. (1995). What difference does it make? Additionality in the public support of R&D in large firms. *International Journal of Technology Management*, volume 10 (4/5/6), pp.587-600.

Christensen, J. L. (2012). Innovation policy evaluation – challenges and roads ahead. Paper for DRUID Society Conference „*Innovation and Competitiveness. Dynamics of organisations, industries, systems and regions*". June 16-18, Copenhagen. [revised 2013 05 10], http://druid8.sit.aau.dk/acc_papers/82614umcol2atftycac71silebqy.pdf

Corchuelo, M. B., Martínez-Ros, E. (2009). The effects of fiscal incentives for R&D in Spain. *Universidad Carlos III de Madrid Working Paper*, *Business Economic Series*, No. 09-23. [revised 2013 05 10], http://e-archivo.uc3m.es/bitstream/10016/3870/1/wb092302.pdf

Czarnitzki, D., Hanel, P., Rosa, J. M. (2005). Evaluating the impact of R&D tax credits on innovation: A microeconometric study on Canadian firms. *Groupe de Recherche en Économie et Développement International Working Paper*, No. 05-01. [revised 2013 05 10], http://pages.usherbrooke.ca/gredi/wpapers/GREDI-0501.pdf

De Jong, J. P. J., Verhoeven, W. H. J. (2007). *WBSO Evaluation 2001-2005: Impacts, target group reach and implementation*. Maastricht Economic and Social Research and Training centre on Innovation and Technology, Ministry of Economic Affairs, the Hague.

Dictionary of Business Terms (2000). J. P. Friedman (ed.). Barron Educational Series, New York.

Dimza, V. (2003). *Inovācijas pasaulē, Eiropā, Latvijā*. LZA Ekonomikas institūts, Rīga.

Dudzevičiūtė, G., Tvaronavičienė, M. (2011). Measurement framework of innovation activity: theoretical approaches' analysis. *Journal of Security and Sustainability Issues*, volume 1 (1), pp. 63-75.

Edler, J., Berger, M., Dinges, M., Gök, A. (2012). The practice of evaluation in innovation policy in Europe. *Research Evaluation*, volume 21, pp. 167–182.

Entrepreneurial innovation in Europe. A review of 11 studies of innovation policy and practice in today's Europe. (2003). European Commission, Office for Official Publications of the European Communities, Luxembourg.

Europe 2020 Flagship Initiative Innovation Union. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. (2010). SEC 1161. European Commission, Brussels.[revised 2013 05 29], http://ec.europa.eu/research/innovation-union/pdf/innovation-union-communication_en.pdf

Flanagan, K., Uyarra, E., Laranja, M. (2010). Reconceptualising the 'policy mix' for innovation. Revised version of the paper "The 'policy mix' for innovation: rethinking innovation policy in a multi-level, mult-actor context. *Manchester Institute of Innovation Research Working Paper*, No. 52 / Manchester Business School Working Paper No. 599. [revised 2013 05 10], https://www.escholar.manchester.ac.uk/api/datastream?publicationPid=uk-ac-man-scw:119191&datastreamId=POST-PEER-REVIEW-NON-PUBLISHERS.PDF

Freeman, C. (1987). Innovation. In: J. Eatwell, M. Milgate, P. Newman (ed.) *The New Palgrave Dictionary of Economics*. volume 2, pp. 858-860. Palgrave Macmillan.

Gans, J., Hayes, R. (2006). Measuring innovative performance - Essential for effective innovation policy and economic growth. *The Melbourne Review: A Journal of Business and Public Policy*, volume 2 (1), pp. 70-77.

Georghiou, L. (1997). Issues in the evaluation of innovation of innovation and technology policy. *OECD Conference on policy evaluation in innovation and technology*. 26-27 June, Paris, pp. 19-33. [revised 2013 05 10], www.oecd.org/sti/inno/1822541.pdf

Gilpin, A. (1973). *Dictionary of Economic Terms*. Butterworths, London.

Gjoski, N. (2011). Innovation and sustainable development: Linkages and perspectives for policies in Europe". *European Sustainable Development Network Quarterly Report*, June 2011. [revised 2013 05 29], http://www.sd-network.eu/quarterly%20reports/report%20files/pdf/2011-June-Innovation_and_sustainable_development.pdf

Haegeland, T., Moen, J. (2007). The relationship between the Norwegian R&D tax credit scheme and other innovation policy instruments. *Statistics Norway Report*, No. 2007/45. [revised 2013 05 10], http://www.ssb.no/emner/10/02/rapp_200745/rapp_200745.pdf

Hall, B., van Reenen, J. (2000). How effective are fiscal incentives for R & D? A review of the evidence. *Research Policy*, volume 29 (4), pp. 449–469.

Kuhn, T. (2012). *The Structure of Scientific Revolutions*. The University of Chicago Press, Chicago.

Lenihan, H. (2011). Enterprise policy evaluation: Is there a 'new' way of doing it? *Evaluation and Program Planning*, volume 34 (4), pp. 323-332.

Linnas, R. (2008). An assessment and evaluation of innovation policy: the case of Estonia. *Discussions on Estonian Economic Policy*, volume 16, pp. 98-118. [revised 2013 05 10], http://www.mattimar.ee/publikatsioonid/majanduspoliitika/2008/05.pdf

Loikkanen, T., Kutinlahti, P., Eerola, A. (2006). Towards an integrated framework of impact assessment and foresight studies in innovation policy analysis. Paper submitted to the 2nd International *Seminar on Future-oriented Technology Analysis*. September 28-29, Seville. [revised 2013 05 10], http://forera.jrc.ec.europa.eu/documents/papers/Foresight_ImpactAss_FINAL.pdf

Lokshin, B., Mohnen, P. (2007). Measuring the effectiveness of R&D tax credits in the Netherlands. *CIRANO Working Paper*, No. 29. [revised 2013 05 10], http://www.cirano.qc.ca/pdf/publication/2007s-29.pdf

Lugones, G. (2009). Training module for the recollection and analysis of innovation indicators. *Inter-American Development Bank Working Paper*, No. 8. [revised 2013 05 10], http://docs.politicascti.net/documents/Doc%2008%20-%20capacitacion%20lugones%20EN.pdf

Macmillan Dictionary of Modern Economics (1986). D. W. Pearce (ed.). The Macmillan Press, London.

Marinova, D., Phillimore, J. (2003). Models of innovation. In: L. V. Shavinina (ed.), *The International Handbook of Innovation*, pp. 44-53. Pergamon.

Moffat, D. W. (1983). *Economics Dictionary*. Elsevier, New York.

Molas-Gallart, J., Davies, A. (2005). Toward theory-led evaluation. The experience of European science, technology, and innovation policies. *American Journal of Evaluation*, volume XX (X), pp. 1-18. [revised 2013 05 10], http://dipartimenti.unicatt.it/dises_sem_060531.pdf

Papaconstantinou, G., Polt, W. (1997). Policy evaluation in innovation and technology: an overview. *OECD Conference on policy evaluation in innovation and technology*. 26-27 June, Paris, pp. 9-14. [revised 2013 05 10], www.oecd.org/sti/inno/1822393.pdf

Pass, C., Lowes, B., Davies, L. (2005). *Collins Dictionary of Economics*, Harper Collins Publishers, Glasgow.

Powell, W.W., Grodal, S. (2006). Networks of innovators. In: J. Fagerberg, D. C. Mowery, R. R. Nelson (ed.), *The Oxford Handbook on Innovation*, pp. 56-85. Oxford University Press.

Public financial support for commercial innovation. Europe and Central Asia knowledge economy study. Part I. (2006). *Europe and Central Asia Chief Economist's Regional Working Paper Series*, volume 1 (1). [revised 2013 05 10], http://siteresources.worldbank.org/INTECA/Resources/KE_Study_Final.pdf

Rip, A. (2003). Societal challenges for R&D evaluation. In: P. Shapira, S. Kuhlmann (ed.), *Learning From Science and Technology Evaluation. Experiences from the United States and Europe*, pp. 32-53. Edward Elgar, United Kingdom.

Roessner, D. (2002). Outcome measurement in the United States: State of the art. *Annual meeting of American Association for the Advancement of Science*, 17 April, Boston, Massachusetts. [revised 2013 05 10], http://www.prism.gatech.edu/~sc149/reseval/html/boston.pdf

Smith K. (2006). Measuring innovation. In: J. Fagerberg, D. C. Mowery, R. R. Nelson (ed.), *The Oxford Handbook on Innovation*, pp. 148-177, Oxford University Press.

Statt, D. A. (2003). *Concise Dictionary of Business Management*. Routledge, London.

Statt, D. A. (2004). *The Routledge Dictionary of Business Management*. Routledge, London.

SOME ISSUES IN INNOVATION POLICY EVALUATION

Summary

The paper explores the diversity of approaches to the evaluation of innovation policy and traces their gradual transformation and historical development. Particular attention is paid to the hierarchical structure and temporal perspectives of evaluation as well as to the comparative analysis of different types of evaluation. Various sets of optional evaluation indicators are also examined.

Nowadays, innovations facilitate both economic development and growth in national welfare. In most countries, politicians are determined to improve the effectiveness of national innovation policies and to assist entrepreneurs with innovation and competitiveness both in the national and international market. The accountability culture, budget limitations and dedication to spending money as efficiently as possible are among reasons for developing evaluation of innovation policy. Moreover, evaluations can assist decision-makers with identifying sources of competitive advantage and areas where public expenditure can reduce market fluctuations.

It is believed that the evaluation of research and innovation policy should be carried out on three levels: a bottom-up element of evaluation concerned with the traditional evaluation of individual programmes, a top-down element of evaluation dealing with the analysis of the thorough health of innovation system and the evaluation of sub-systems or bottleneck analysis to examine the systemic role of institutions, clusters and other elements. For higher level evaluations, analytical tools such as peer review of portfolio or benchmarking leading to aggregated judgements should be used. The application of socio-economic methods including scientometrics, surveys, case studies, best practice and cluster analysis is also gaining in popularity.

At present, formative approach, systemic analysis and benchmarking are recognised as dominating approaches to evaluation. Most evaluations comprise both formative and summative elements. Summative evaluations are found to be more intensively discussed by policy makers and are able to cause terminations, redesign or expansion of the programme whereas formative evaluations cause less remarkable changes. In addition to the peculiarities of the assessed process, the availability of resources (time, finance and human) for the collection of data used in the evaluation also determines the selection of indicators, which, in general, should meet such criteria as belonging, reliability and comparability.

It is concluded that a demand-driven innovation policy and diffusion-oriented measures lead to the use of more advanced and even experimental evaluation methods going beyond previous practice. The design of an integral evaluation system is a challenging task due to increasing complexity of evaluation objects and the impact of numerous factors outside the evaluation model. Because of the politicians' particular interest in simple economic indicators and traditional measures of the short-term impact of innovation policy, a conflict with evaluation experts arises. A systemic analysis and a benchmark approach could be used to avoid this conflict to a certain extent. Recently, an adaptive learning approach has been considered the state-of-the-art method for evaluating innovation policy. In addition to exact measures, it could provide information for the recurrent implementation of successful innovation policy instruments and the cancellation of ineffective ones.

KEYWORDS: evaluation, indicators, innovation policy, methods, research and development.

Artūrs Prauliņš, *Dr. oec.*, researcher at the Department of Agribusiness information and analysis, Latvian State Institute of Agrarian Economics. Scientific interests: financial and management accounting, analysis of farm performance, microeconomic analysis. Address: Struktoru Street 14, Riga, LV-1039, Latvia. Phone: (371) 67552786; fax: (371) 67541789; e-mail: arturs@lvaei.lv.