

CRITERIA AND METHODOLOGIES FOR ASSESSING EFFICIENCY OF ENVIRONMENTAL GOVERNMENT PROGRAMS IN THE RUSSIAN FEDERATION

Andrey Margolin¹

Russian Presidential Academy of National Economy and Public Administration

Abstract

Existing approaches to performance evaluation for environmental government programs require improvement. In the Russian context, the obstacles to objective evaluation include: target indicators for state programs are not set according to SMART (Specific, Measurable, Achievable, Relevant, Time-bound) criteria; the importance of budget efficiency indicators for investment decision-making is underestimated; and, some approaches to ex post evaluation of government programs are oversimplified. Specific recommendations are given that would allow improvement of the methodology for ex ante appraisal and ex-post evaluation of environmental programs. A flowchart is developed to guide decision-making on whether to terminate or continue the program on the basis of its overall evaluation rating, which is calculated using a modified Program Assessment Rating Tool (PART), and the degree of conformity between actual and planned volume of financing. The flowchart represents a formalized procedure for the adjustment of the program implementation period and schedules for the achievement of target values for individual indicators; review of target indicator values; funding amounts and schedules; and change of management. A case study of two Russian environmental programs, *Pure Water* and *Water Industry Development*, is used to test the approaches recommended by the author.

¹ Andrey Margolin is Vice-Rector at the Russian Presidential Academy of National Economy and Public Administration, Moscow, Russia.

The craving for immediate enjoyment will always be stronger than the voice of conscience. Even the fear that their own children will have nothing to breathe does not stop people.

—Bernard Werber, French writer

Introduction

Despite the broad consensus among socially responsible politicians, researchers and public figures on the urgent necessity of remediating accumulated environmental problems, everyday managerial decisions made by most public officials and business people are more often focused on short-term gain at the expense of possible long-term ecological implications.

The logic behind the well-known saying “after us, the deluge” is based on the common misconception that the assimilatory capacity of the environment is nearly limitless. It would not be an exaggeration to say that this contributes to the increased urgency of the issue in terms of the growing threat of a global environmental crisis. For example, over half of the planet’s population only has access to low-quality drinking water; moreover, according to the World Health Organization², the number of deaths caused by air pollution is nearly 7 million people per year.

One of the most effective ways to overcome these challenges is the development and implementation of government programs (GPs). Such programs require substantial investment resources and have no immediate effect; besides, in many cases their impact cannot be quantified in the usual monetary terms. To a large extent, this is why, despite the existence of extensive research in this area, as well as several approved evaluation methodologies and their practical applications demonstrated in various countries of the world, environmental problems are still far from being solved and the task of selecting objective criteria and methods for evaluation of environmental government programs retains its relevance.

What is the right way to choose indicators that would best reflect the goals of environmental programs? Which criteria should we focus on when substantiating the efficiency targets for a government program, especially if we take into account the importance of its non-financial outcomes? How can we objectively evaluate the degree to which a program’s intermediate results are achieved? How do the results of such evaluation affect the decision of whether program implementation should be continued or suspended? This article aims to search for answers to these questions, which are anything but simple. The recommendations provided in the article are illustrated by a case study of two government programs implemented in the Russian Federation.

Methodology

State programs are one of principal tools in the implementation of socio-economic policy in various fields, and primarily in the so-called public sector. It is nearly impossible to deal with such issues as the development of education, healthcare, culture, sport, prevention of environmental emergencies, and improving the condition of the natural environment without research-backed planning and efficient implementation of GPs. Since budget funds are always limited, the value-for-money requirements are rather strict and are usually reflected in relevant methodologies for

² “7 million premature deaths annually linked to air pollution.” World Health Organization, Media Centre, 25 March 2014. Accessed on November 25, 2016. <http://www.who.int/mediacentre/news/releases/2014/air-pollution/en/>

the evaluation of GP performance developed with due regard to best international practices (Robinson 2013; Kuzmin, O'Sullivan and Kosheleva 2009; Shepherd 2012; Afanasiev and Shash 2013).

A similar approach is employed in the Russian Federation, where there are currently several methodologies approved by laws and regulations on the federal (*Metodicheskie ukazaniya* 2013) and regional levels³. These and other documents reflect the consensus of government representatives who bear personal responsibility for GP quality and experts who are professionally involved in GP development and feasibility evaluation.

An analysis of efficiency⁴ evaluation methodologies for government programs requires consideration of the following three aspects or stages of evaluation:

- 1) selection of target indicators (TIs), or the quantitative and qualitative characteristics reflecting a program's objectives. TIs serve as the baseline for ex ante appraisal and ex post evaluation of a program;
- 2) ex ante appraisal, which is carried out at the stage of program development in order to make the decision about the program's viability; and,
- 3) ex post evaluation, which is performed in order to ascertain whether the implementation progress corresponds to initial plans, identify the reasons for possible deviations and underused resources, and ultimately to decide whether program funding should be ceased or continued.

A short review of existing approaches to these tasks is presented below.

The choice of target indicators

Target indicators (TIs) of any GP are qualitative and quantitative characteristics reflecting the degree of achievement of its main objectives. The key principles of TI selection and their brief descriptions are presented in table 1 (compiled by the author using the data from *Metodicheskie ukazaniya* 2013). TI values are normally calculated in one of the following ways: (1) using data obtained from governmental statistical surveys; (2) using methodologies adopted by international organizations; (3) using methodologies approved by regulation of the Russian government or a specific agency in charge of the program; and (4) using methodologies included in the program itself. The possibility of complete achievement of target values of these indicators within the established time limit is seen as a necessary requirement in ex ante appraisal of program efficiency. The existing framework guidelines give developers a lot of leeway when selecting target indicators for a specific GP. Such freedom seems excessive, especially since it often extends to the choice of efficiency criteria. In most cases, the developers focus on the TI construction methodologies included within the program itself. This approach may lead to the violation of the principles of

³ For example: Government of Saint Petersburg. "Poryadok prinyatiya reshenii o razrabotke gosudarstvennykh programm Sankt-Peterburga" [Procedure for decision making on the development of government programs in Saint Petersburg] December 25, 2013, last amendment September 28, 2016. <http://docs.cntd.ru/document/822402754>

⁴ The term *efficiency* is used in this paper in its interpretation as the ratio between outputs produced and amount of inputs used. It should be noted that, unlike private sector investment projects, government programs often have outputs and objectives that cannot be represented in monetary terms.

objectivity, unambiguity and credibility of TI selection, while carrying the additional risk of questionable quality of the results of the assessment of the GP's ex ante and ex post efficiency.

Table 1 Key principles of target indicator selection

Principle	Description
Social significance	All target indicators must reflect end users' degree of satisfaction with the public services as well as the quality and scope which the program is meant to improve
Objectivity and unambiguity	Each target indicator must unambiguously reflect progress toward achievement of the program's objectives; one TI cannot be improved at the cost of other TIs
Credibility	The accuracy of the source data used to calculate TIs must be verifiable
Cost effectiveness	Existing data collection procedures must be used to minimize the costs of TI calculation
Comparability	It must be possible to accumulate results of TI calculation and ensure their comparability throughout the whole implementation period

Source: Author compiled data from *Metodicheskie ukazaniya* 2013.

Ex ante appraisal of government programs

Ex ante appraisal for Russian federal and regional government programs is carried out at the GP development stage using the following criteria: (1) economic efficiency, which reflects the contribution of the GP to the economic development of the Russian Federation and its expected impacts on various spheres of the economy; and (2) social efficiency, which reflects the expected impact of the GP on social development, and is usually impossible to measure in terms of cost indicators.

Since neither federal nor regional laws contain any guidelines on the threshold values of economic and social efficiency criteria, program developers lack the necessary tools to answer the essential question: 'Is the contribution of this GP into the economic and/or social development of the Russian Federation valuable enough to recognize it as efficient?' This is why decisions on whether to launch or reject a certain program are often made regardless of its possible economic and/or social impact. Essentially, a positive result of ex ante appraisal currently depends to a much higher degree on the correlation between the prospects of TI achievement and the financial possibilities of the program's initiators (or, for a federal-level program, the capacity of the federal budget).

As is the case with TI selection, the methodological approaches to ex ante GP appraisal require further elaboration. Some suggestions are provided below in the *Results and discussion* section.

Ex post evaluation of government programs

The guidelines for government program evaluation (*Metodicheskie ukazaniya* 2013) state that the quintessence of ex post evaluation is the need to assess the extent to which the goals and objectives of separate program activities and the government program as a whole are achieved, as well as the ratio of actual to planned budget expenditure. The methodology presented in table 2 is currently used for the evaluation of most GPs funded from the federal budget.

Table 2 Standard criteria and methods for government program evaluation

Area of evaluation	Indicators	Efficiency criteria
1. The degree of achievement of program goals and objectives, defined as the simple average between actual and planned TI values	$DA^{total} = \frac{1}{n} \sum_{k=1}^n I_k^{total}$ <p>where:</p> <p>DA^{total} is the indicator of the degree of achievement of GP goals and objectives;</p> <p>n is the number of target indicators;</p> <p>I_k^{total} is the ratio of actual to planned value of the kth target indicator</p>	$DA^{total} \geq 1 \quad (1)$
2. Ratio of actual to planned budget expenditure	$RI = \frac{F^a}{F^p}$ <p>where RI³ is the indicator of the ratio of actual (F^a) to planned (F^p) volume of GP funding</p>	$RI \leq 1 \quad (2)$

An advantage of this system of indicators is its simplicity. In fact, no other tools are required to evaluate a program except for direct comparisons between actual and planned values of target indicators. However, is such a simplified approach sufficient for the solution of such a complicated task as government program evaluation?

For instance, in order for the indicator of the degree of achievement of GP goals and objectives to be greater than 1 (formula 1 in table 2), it is not at all necessary for all actual indicator values to meet or exceed target values. This means that if this methodology is applied, a program can be considered effective even if one or more of its indicators do not reach target values. The fact of

non-achievement of individual indicators of a program is not in itself a reason to label it as inefficient; however, it signifies the necessity of a thorough analysis of the underlying reasons. But if a program is already accepted as efficient, the question of further analysis will never arise. It is also worth noting that this approach to the determination of DA^{total} is based on the assumption of equal importance of all indicators. But in reality, this is often not the case.

One of the more promising directions for the development of GP evaluation methodologies is PART (Program Assessment Rating Tool), which was created in the USA and widely used in many countries, including the Russian Federation (Gilmour 2007; Margolin 2013). It is interesting to note that at the start of the PART initiative about 50% of all federal programs in the USA were assigned an “inefficient” status. After a year, this number went down to 30%, indicating an improvement in overall quality and outcome orientation of federal programs.

The PART approach is based on a generalized multi-criteria evaluation of expert opinions presented as answers to a series of questions grouped into four topical areas:

- 1) Program Purpose and Design – determines the clarity of program purpose and importance of the program, efficiency of the proposed problem solving mechanisms, and resource allocation;
- 2) Strategic Planning – determines the presence of long-term and interim program objectives and performance indicators;
- 3) Program Management – assesses the program’s management quality, including financial oversight and coordination of program activities; and,
- 4) Program Results – assesses program performance on achievement of intermediary outputs and long-term outcomes.

Individual ratings are calculated for each of these four areas and then combined into an overall rating used to assess the efficiency of a government program.

On the whole, although program performance evaluation methodologies have been tried and tested in many countries of the world, in practice, they often do not protect against misuse and inefficient use of budgetary and extra-budgetary funds, followed by program implementation results that can only be described as dismal. In such spheres as natural resource management and environmental protection, design flaws, construction, and maintenance mistakes can lead to major environmental disasters. Some examples worth mentioning are the industrial accident at the Sayano-Shushenskaya hydroelectric plant (Yenisei river, Russia, August 2009), the explosion on the Deepwater Horizon oil rig in the Gulf of Mexico (Louisiana, USA, April 2010), the Fukushima Daiichi nuclear disaster (northeast Japan, March 2011), and there are many others. Nevertheless, the PART approach still seems to be the most viable, which is why its modified version is used in this paper to improve the current Russian methodology of ex post GP evaluation.

Results and discussion

In order to formulate specific guidelines for the improvement of GP performance evaluation methodologies, let us look at the approved target indicators for two large-scale environmental

government programs: *Pure Water*⁵ and *Water Industry Development*⁶, presented in table 3. An analysis of the data in table 3 indicates the following:

(1) *By default, the target indicators are considered to be equally significant, which is by no means the case in reality.*

For example, the relative weight of such target indicators of the *Pure Water* program (column 1 in table 3) as “percentage of borrowed funds in the total volume of capital investment into water supply, disposal and wastewater treatment facilities” or “percentage of water supplied by utility providers operating on the basis of concession agreements” is significantly inferior to almost all indicators that are directly related to the performance and safety of water supply, disposal, and treatment facilities. On the contrary, such TIs as availability of centralized water supply and disposal services should receive higher weighting coefficients in comparison to the others.

Moreover, the assessment of program performance based solely on the ration between actual and planned indicator values (see formula 1 in table 2) can be insufficiently objective since it barely takes into account the actual progress in objective completion during program implementation. For example, if the value of the TI “Percentage of wastewater purified to meet approved standards in the total volume of wastewater passed through sewage treatment facilities” for the *Pure Water* program is 52%, a perfunctory consideration seems to indicate that the actual to planned TI value ratio is $52 / 53 = 0.981$. However, if we take into account the fact that the TI value was 46% at the start of program implementation (see table 3), this ratio becomes lower: $(52 - 46) / (53 - 46) = 0.857$.

In essence, the suggested approach to dealing with unequal values of a program’s TIs is as follows (see also the corresponding analytical formulas⁷):

- To use a weighted average rather than simple average to calculate the ratio of actual to planned TI values. This will solve the problem of unequal significance of different indicators and their corresponding impacts on the results of performance evaluation;
- To use the incremental values of TIs rather than their absolute values;
- To take into account the expected dynamics of TI values when calculating the ratios of actual to planned values. For example, in the *Pure Water* program, the value of the TI “Availability of centralized water supply services to the population” is supposed to

⁵ Government of the Russian Federation. “Federal'naya tselevaya programma “Chistaya voda” na 2011 - 2017 gody” [Federal target program “Clear Water” for 2011-2017]. December 22, 2010. <http://docs.cntd.ru/document/902256587>

⁶ Government of the Russian Federation. “Federal'naya tselevaya programma “Razvitie vodokhozyaistvennogo kompleksa Rossiiskoi Federatsii v 2012 - 2020 godakh” [Federal target program “Water industry development in the Russian Federation” for 2012-2020]. April 19, 2012. <http://docs.cntd.ru/document/902343713>

⁷ $DA^{total} = \sum_{n=1}^N \alpha_n \times DA_n$;

$$DA_n = \begin{cases} \frac{(TI_n^a - TI_n^s)}{(TI_n^p - TI_n^s)}, & \text{if the TI value is supposed to increase} \\ \frac{(TI_n^s - TI_n^a)}{(TI_n^s - TI_n^p)}, & \text{if the TI value is supposed to decrease} \end{cases} \quad (3)$$

where α_n , TI_n^s , TI_n^a , TI_n^p are respectively, the weighting factor for the n^{th} TI, the value of the TI at the start of GP implementation, its actual and planned value in the GP implementation year under consideration.

increase, whereas the value of the TI Percentage of polluted wastewater in the total volume of wastewater released into surface water bodies in the *Water Industry Development* program is supposed to decrease.

(2) *It is important to take into account the macroeconomic conditions of government program implementation.*

One objective of the *Pure Water* program was to increase the percentage of capital investment into water supply, wastewater disposal, and purification systems out of the total volume of water industry revenue from 10% in 2011 to 31% in 2017. However, this would lead to an imbalance of supply and demand, since it would require increasing the service tariffs to a level that customers would not be able to afford. The situation with the *Water Industry Development* program is similar. For example, the significance of the TI “Number of projects for the construction (reconstruction) of sewage disposal and water recirculation facilities implemented via the mechanism of interest rate subsidies” appears to be minimal, particularly in the context of a gradual reduction of inflation, investment risks and, therefore, of market interest rates⁸. It is also worth noting that in current Russian conditions, which are characterized by increased risks to investment activity, subsidizing interest rates is far from being the only mechanism of state support. Another such mechanism includes, among others, the provision of state guarantees on bonds issued for the implementation of priority projects or the conclusion of special investment contracts that include tax benefits received by investors for the duration of the project.

⁸ In the Fisher equation, the nominal interest rate is represented as the sum of inflation rate and the real interest rate, which, in turn, has a certain dependency on investment risks. Therefore, even in the hypothetical case in which the inflation rate is reduced to zero, the nominal interest rate can remain quite high if the investor considers the project to have a high enough level of risk to include a risk premium in calculations.

Table 3 Target indicators for the *Pure Water* and *Water Industry Development* government programs

<i>Pure Water</i>			<i>Water Industry Development in the Russian Federation</i>		
	Values at the beginning and end of the implementation period			Values at the beginning and end of the implementation period	
	2011	2017		2012	2020
Percentage of water samples taken from the water supply system that do not meet the sanitary requirements of hygienic standards, %	16.4	14.4	Increase in the population provided with improved water resources, million people	0.3	6.3
Percentage of water samples taken from the water supply system that do not meet the microbiological requirements of hygienic standards, %	5	4.4	Percentage of polluted wastewater in the total volume of wastewater released into surface water bodies, %	88.6	45.2
Percentage of outdoor water lines that require replacement, %	43	28	Percentage of population covered by measures aimed at increasing protection from negative water-related impacts in the total population of areas affected by detrimental water-related impacts, %	68.3	85
Percentage of outdoor sewage systems that require replacement, %	36	27	Percentage of water facilities with unsatisfactory and dangerous safety levels restored to safe operating conditions, %	17.6	92.3

<i>Pure Water</i>			<i>Water Industry Development in the Russian Federation</i>		
	Values at the beginning and end of the implementation period			Values at the beginning and end of the implementation period	
	2011	2017		2012	2020
Percentage of wastewater purified to meet approved standards in the total volume of wastewater passed through sewage treatment facilities, %	46	53	Ratio of renovated and new hydrological stations and laboratories out of the total requirement, %	7	83.8
Volume of wastewater passed through sewage treatment facilities in the total volume of wastewater, %	93	100	Number of newly created water reservoirs and hydrosystems on existing multipurpose reservoirs and water supply channels renovated to increase water yield, units	4	73
Percentage of centralized water supply services available to the population, %	77	85	Reconstruction and environmental rehabilitation of water objects, km	-	4010
Percentage of centralized water disposal services available to the population, %	73	84	Scope of new and renovated engineering protection and coast protection systems, km	31.5	1 763.9
Percentage of capital investment into water supply, disposal and wastewater treatment out of the total revenue of water supply, disposal and wastewater treatment facilities, %	10	31	Number of water facilities with unsatisfactory and dangerous safety levels restored to safe operating conditions, units	165	2 265

<i>Pure Water</i>			<i>Water Industry Development in the Russian Federation</i>		
	Values at the beginning and end of the implementation period			Values at the beginning and end of the implementation period	
	2011	2017		2012	2020
Percentage of borrowed funds out of the total volume of capital investment into water supply, disposal and wastewater treatment facilities, %	9	30	Number of renovated or reopened hydrological stations and laboratories within the state monitoring network, units	90	3 347
Percentage of water supplied by utility providers operating on the basis of concession agreements, %	2	35			
Percentage of water supplied by utility providers at tariffs set for the long-term regulation period, %	5	70			

Sources: Government of the Russian Federation. "Federal'naya tselevaya programma "Chistaya voda" na 2011 - 2017 gody" [Federal target program "Clear Water" for 2011-2017]; "Federal'naya tselevaya programma "Razvitie vodokhozyaistvennogo kompleksa Rossiiskoi Federatsii v 2012 - 2020 godakh" [Federal target program "Water industry development in the Russian Federation" for 2012-2020].

From the point of view of GP implementation, the actual mechanism of state support is not as important as the achievement of target indicators. In this respect, the TI mentioned above could steer the program managers to use only the mechanism of interest rate subsidies, limiting their capacity to attract investment resources; this is the reason why this TI is excluded from further consideration.

A necessary condition for the shortlist of target indicators is the exclusion of any indicators that lack macroeconomic prerequisites. The shortlist of TIs formed according to the above guidelines is the foundation for ex ante appraisal of GP performance. Besides obligatory achievement of TIs, it seems advisable to supplement the methodology by including the calculation of budget performance indicators, and above all, net present budget value (NPBV). Essentially, this is the ratio between tax and non-tax budget revenues and budget payments (such as co-financing of investment into environmental programs, subsidizing interest rates on loans, energy payments, etc.) It is calculated similarly to net present value (NPV), which is used to substantiate commercial efficiency of investment projects. However, while NPV is used to select the best investment option, the scope of practical application of the NPBV indicator is quite different: we believe it should be used not as a rigid criterion for acceptance or rejection of the program, but rather as an indicator that can increase the objectivity and validity of administrative decisions in the future.

Let us examine an algorithm of actions aimed at the coordination of long-term interests of all GP participants depending on NPBV values (discussed in more detail in Margolin 2012):

1. *NPBV > 0, which means that net budget revenue exceeds budget expenditure.* This is quite a rare case, since the outcomes of a GP quite often cannot be reduced to quantitative indicators. Nevertheless, this is possible if we take into account the multiplicative effects in related industries and cross-industry clusters that take the form of indirect tax payments from all program stakeholders⁹ into the consolidated budget. However, this situation is indicative of the fact that the possibilities of extra-budgetary funding attraction are far from being fully exploited, rather than of high program efficiency. Therefore, if the $NPBV > 0$ condition is met, it indicates the necessity to search for private sector investors and, possibly, to replace the initially planned budget funding by other instruments of state support (e.g. state guarantees, investment tax credits, interest rate subsidies on bank loans, etc.).
2. *NPBV < 0, that is, unlike the previous case, net budget revenue is lower than budget expenditure.* In this case, the task of attracting extra-budgetary financing seems to be extremely difficult. However, individual activities within the program can still remain socially and economically efficient and may be implemented in partnership with private sector companies with some preliminary planning.

In this context, the dynamic development of public private partnership (PPP) as an instrument of GP implementation in Moscow and the Moscow region is of particular interest. The pressing issue of waste management within the Environmental Protection program in the Moscow region requires the closure of 29 landfills and the construction of waste treatment plants with an annual capacity of 4-5 million tons of garbage using the PPP model. A particular feature of such projects is their considerable multiplicative effect: besides the direct effect of termination of solid waste disposal in landfills, there are indirect effects, including the possibility of

⁹ These include: providers of materials and equipment for the construction of facilities specified in the program; suppliers of raw materials and components for the manufacturing of goods or provision of services during program implementation; and companies that are consumers of these products or services.

producing and selling electricity and construction materials that are a byproduct of waste processing. An optimum balance between social, budgetary and commercial efficiency is obvious in this case.

PPP models have substantial potential for use in government programs, even when the prospects of attracting private sector companies seem questionable at a first glance. If a partnership is achieved, it may be necessary to review the initial result of NPBV calculation since its value can become positive due to the consideration of multiplicative effects.

3. *NPBV remains negative even after a thorough analysis of opportunities for private investment attraction.* In this case, it is necessary to estimate the significance of long-term impacts of the government program. If it is high enough (for example, the implementation of the program may prevent an environmental disaster in the future), NPBV is recalculated using a lowered discount rate (for example, equal to the refinancing rate divided by two¹⁰). It should be noted that this approach to the appraisal of socially significant projects is standard international practice and permits reduction of the effect of depreciation of cash receipts and payments over time.
4. *NPBV remains negative even if the discount rate is reduced.* In this scenario, it is necessary to conduct a thorough analysis of positive social returns of the program that are not reducible to quantitative values. If the expert consensus on the significance of social effects of the program is favorable, the program may be considered feasible even with a negative NPBV. If there are several options for program implementation that allow for the achievement of target indicator values within the stipulated time period, the preferable option is the one where NPBV is the smallest absolute value.

In accordance with the above arguments, we suggest that a key amendment to the current methodology of ex ante appraisal of government programs should be the addition of the NPBV indicator in order to assess opportunities for extra-budgetary funding and select the best option for program implementation.

If the ex ante efficiency of the government program is determined to be high enough to start funding, the public contracting authority faces the task of objectively evaluating intermediary outputs and determining the decision-making procedure with regard to whether to continue or to terminate the program. There is a need for insightful responses to the following questions:

- Is the program still relevant?
- Is it necessary to review the planned values of target indicators?
- Is it necessary to review the duration of the program and the timeline for the achievement of its target indicator values?
- Is it necessary to review the timeline and funding for program activities?
- Is it advisable to replace the managers responsible for program implementation?

In order to answer these questions for environmental programs, we suggest using a modified PART methodology, which, as mentioned above, is based on a generalized multi-criteria

¹⁰ The discount rate for GPs with a high social significance should be set as either equal to the Central Bank refinancing rate, or the rate of return on state bonds with a long maturity (20-30 years). In this approach, two components of the risk premium are excluded from the calculation, one of which is associated with the risks of the company initiating the project and the other one with the risks of the project itself.

evaluation of expert opinions presented as answers to a series of questions grouped into four topical areas (see also Office of Management and Budget 2008).

An individual rating is provided for each of these areas. For the first three sections, “Program Purpose and Design,” “Strategic Planning,” and “Program Management,” this rating is calculated based on experts’ answers to a number of evaluation questions recommended by the creators of PART, with some amendments made by the author to reflect the peculiar aspects of development and implementation of environmental government programs in the Russian Federation.

Tables 4.1, 4.2, and 4.3 below contain individual ratings calculated by the author for the Water Industry Development program. Of course, the list of questions provided here is not set in stone and can be reviewed. The Yes/No format can be expanded to reflect partial achievement of goals. Since environmental programs are difficult to evaluate in precise numerical terms, the approach used in these tables is to choose one of four answer options: “yes,” “to a large extent,” “to a small extent,” and “no.” On the one hand, this expansion does not change the general logic of the methodology while increasing the flexibility of individual ratings, but on the other hand, it could possibly lower the experts’ responsibility for their decisions.

Individual ratings (R^p) in tables 4.1-4.3 are determined by the following formula:

$$R^p = \sum_{n=1}^N \alpha_n \times B_n \quad (4)$$

where α_n , B_n are, respectively, the weight coefficient and the numerical score for the n^{th} question (a detailed calculation is demonstrated in table 4.1).

As for the Program Results section of PART, it contains the following standard questions for the experts to answer:

- Has the program demonstrated adequate progress in achieving its long-term performance goals?
- Does the program (including program partners) achieve its annual performance goals?
- Does the program demonstrate improved financial performance indicators each year?
- Does the performance of this program compare favorably to other programs, including private projects, with similar goals?
- Do independent evaluations of sufficient scope and quality indicate that the program is effective and achieving results?

As logical, non-contradictory, and relevant as these questions are, the answers overlook the actual target indicator values which reflect program performance more objectively than the opinions of even the most qualified experts. Let us emphasize that it is not advisable to give up expert appraisal altogether; however, the focus should shift from simple yes/no answers to a set list of questions that determine weight coefficients for target indicators, allowing one to assess both the degree of progress and the relative significance for each indicator.

With respect to the program under consideration, the final rating determined by formula 4 is presented in table 5.

Table 4.1 Program Assessment Rating Tool: Section 1. Program Purpose and Design

Questions reflecting the content of the evaluation procedures	Individual rating calculation			
	Weight coefficient	Answer	Numerical score	Share in total rating
		Yes	1	
		Large extent	2/3	
		Small extent	1/3	
		No	0	
		Yes	1	
		Large extent	2/3	
		Small extent	1/3	
		No	0	
		Yes	1	
		Large extent	2/3	
		Small extent	1/3	
		No	0	
		Yes	1	
		Large extent	2/3	
		Small extent	1/3	
		No	0	
		Yes	1	
		Large extent	2/3	
		Small extent	1/3	
		No	0	
Total rating for Section 1				0.634

¹¹ SMART stands for Specific, Measurable, Achievable, Result-oriented, Time-bound.

Table 4.2 Program Assessment Rating Tool: Section 2. Strategic Planning

Questions reflecting the content of the evaluation procedures	Individual rating calculation			
	Weight coefficient	Answer	Numerical score	Share in total rating
		Yes	1	
		<u>Large extent</u>	<u>2/3</u>	
		Small extent	1/3	
		No	0	
		<u>Yes</u>	<u>1</u>	
		Large extent	2/3	
		Small extent	1/3	
		No	0	
		Yes	1	
		<u>Large extent</u>	<u>2/3</u>	
		Small extent	1/3	
		No	0	
		Yes	1	
		Large extent	2/3	
		<u>Small extent</u>	<u>1/3</u>	
		No	0	
		Yes	1	
		Large extent	2/3	
		<u>Small extent</u>	<u>1/3</u>	
		No	0	
		Yes	1	
		<u>Large extent</u>	<u>2/3</u>	
		Small extent	1/3	
		No	0	
Total rating for Section 2				0.633

Table 4.3 Program Assessment Rating Tool: Section 3. Program Management

Questions reflecting the content of the evaluation procedures	Individual rating calculation			
	Weight coefficient	Answer	Numerical score	Share in total rating
		Yes	1	
		Large extent	<u>2/3</u>	
		Small extent	1/3	
		No	0	
		Yes	1	
		Large extent	<u>2/3</u>	
		Small extent	1/3	
		No	0	
		Yes	1	
		Large extent	2/3	
		Small extent	<u>1/3</u>	
		No	0	
		Yes	1	
		Large extent	2/3	
		Small extent	<u>1/3</u>	
		No	0	
		Yes	1	
		Large extent	2/3	
		Small extent	<u>1/3</u>	
		No	0	
		Yes	1	
		Large extent	2/3	
		Small extent	<u>1/3</u>	
		No	0	
Total rating for Section 3				0.466

Table 5 Calculating individual ratings: Section 4. Program results

Target indicators for the <i>Water Industry Development</i> program	Target indicator values			TI achievement ratio, as a decimal	α_n in formula 3 (expert appraisal)
	Initial (from Table 3)	Planned (for 2016)	Actual		
1. Increase in the population provided with improved water resources, million people	0.3	2.7	2.4	0.875	0.2
2. Percentage of polluted wastewater in the total volume of wastewater released into surface water bodies, %	88.6	73	75	0.872	0.2
3. Percentage of population covered by measures aimed at increasing protection from negative water-related impacts in the total population of areas affected by detrimental water-related impacts, %	68.3	75.9	74.8	0.855	0.2
4. Percentage of water facilities with unsatisfactory and dangerous safety levels restored to safe operating conditions, %	17.6	47.5	42.7	0.839	0.1
5. Ratio of renovated and new hydrological stations and laboratories out of the total requirement, %	7	34.8	29.4	0.806	0.05
6. Number of newly created water reservoirs, hydrosystems on existing multipurpose reservoirs, and water supply channels renovated to increase water yield, units	4	25	21	0.81	0.05
7. Scope of reconstruction and environmental rehabilitation of water objects, km	-	1 310	1 100	0.84	0.05
8. Scope of new and renovated engineering protection and coast protection systems, km	31.5	763.9	640	0.831	0.05
9. Number of water facilities with unsatisfactory and dangerous safety levels restored to safe operating conditions, units	165	1 005	850	0.815	0.05
10. Number of renovated or reopened hydrological stations and laboratories within the state monitoring network, units	90	1 265	1 070	0.834	0.05
Results of DA ^{total} calculation (formula 3)	DA^{total} = 0.2×(0.875+0.872+0.855)+0.1×0.839+0.05×(0.806+0.81+0.84+0.831+0.815+0.834) = 0.851				

Since Russian budget expenditures were cut by at least 10% in the wake of the 2014-2015 economic downfall, the calculations are based on the assumption that program objectives will be approximately 80-90% completed by the end of 2016. Ultimately, neither estimated TI values nor expert appraisals of weighting coefficients are to be taken as completely accurate. The main purpose of the data in tables 4.1, 4.2, 4.3 and 5 is merely to illustrate the methodology proposed in this paper to evaluate GP performance.

The last stage of ex post GP evaluation is to determine the overall integrated rating (see formula 4 and table 6) and compare it to the scale in table 6 (Office of Management and Budget 2008). The result determines the decision of whether to continue or cease program implementation.

$$R_{\text{int}} = \sum_{i=1}^4 \mu_i \times R_i^p, \quad (5)$$

where: μ_i , is the weighting coefficient for each individual rating; R_i^p is the individual rating for each of the 4 areas (i=1,2,3,4).

The representative estimation of weighting coefficients is a substantive aspect of the methodology. The value ranges in table 7 are based on the original PART guidelines; however, this is not the only possible approach. In fact, the author's experience of GP evaluation case studies with various and representative groups of students shows that most Russian specialists tend to increase the weighting coefficients for the Program Purpose and Design and Program Management areas while lowering the coefficient for Program Results. The final choice is always determined by the specific features of the program under consideration and the expertise of the decision makers.

Table 6 Overall rating calculation

	Area of evaluation	Weighting coefficient	Individual ratings
1	Program Purpose and Design	0.2	0.634
2	Strategic Planning	0.1	0.633
3	Program Management	0.2	0.466
4	Program Results	0.5	0.851
	Overall rating, R_{int}	$0.2 \times 0.634 + 0.1 \times 0.633 + 0.2 \times 0.466 + 0.5 \times 0.851 = 0.709$	

Table 7 Qualitative assessment of government program performance based on its overall rating

Quantitative rating (R_{int})	Qualitative assessment
$R \geq 85\%$	Effective
$85\% > R \geq 70\%$	Moderately effective
$70\% > R \geq 50\%$	Adequate
$R < 50\%$	Ineffective

The integrated overall rating in the above example is 0.709, which places the program into the moderately effective range. However, this conclusion is not definitive and requires detailed elaboration. For instance, it is advisable to consider the ratio of actual to planned volume of program funding (F^a and F^p in table 2). If this ratio substantially exceeds the R_{int} rating ($F^a / F^p \gg$

R_{int}), this means that program outputs are disproportionate to program funding, and the agency in charge of the program may make a justified decision to replace the program managers. Conversely, if $F^a / F^p \ll R_{int}$, we can come to the conclusion that the managers were able to maintain a high level of performance despite funding cuts, and the adopted program management approaches deserve replication.

In the Russian context, lowering administrative expenses often turns out to be an effective way to improve program performance, since this helps avoid the rather too common situation where a large part of the program budget is used to pay administrative personnel rather than to fund program activities. Monitoring and control of GPs is currently the responsibility of the Ministry of Economic Development and the Ministry of Finance of the Russian Federation; a more scientifically grounded approach to the development of monitoring guidelines (such as *Metodicheskie ukazaniya* 2013) would allow for early prevention of possible problems and deviations in program implementation. Additionally, a system of citizen oversight would be beneficial, for example, in the form of an online platform that would accumulate, summarize, and systematize information received from the population, who are, after all, key stakeholders of government programs.

The flowchart in Figure 1 shows the overall process of decision-making by the agency responsible for program implementation based on intermediate program performance assessment at any given moment in time. While this particular flowchart is based on the assumption that the program qualifies as “moderately effective,” the overall logic of the approach will remain the same for any other performance range in table 7.

It should be noted that using the algorithm represented in the flowchart would allow the program managers to avoid any significant deviation of actual TI values and expenditures from their planned values due to annual adjustment of planned TI values, timelines for their achievement, and required levels of funding based on the evaluation of intermediate program results. Nevertheless, if by the end of program implementation the actual TI values turn out to be significantly lower than those initially planned, while the expenditures are significantly higher, there is every reason to conclude inefficient management of the program by the government agency in charge.

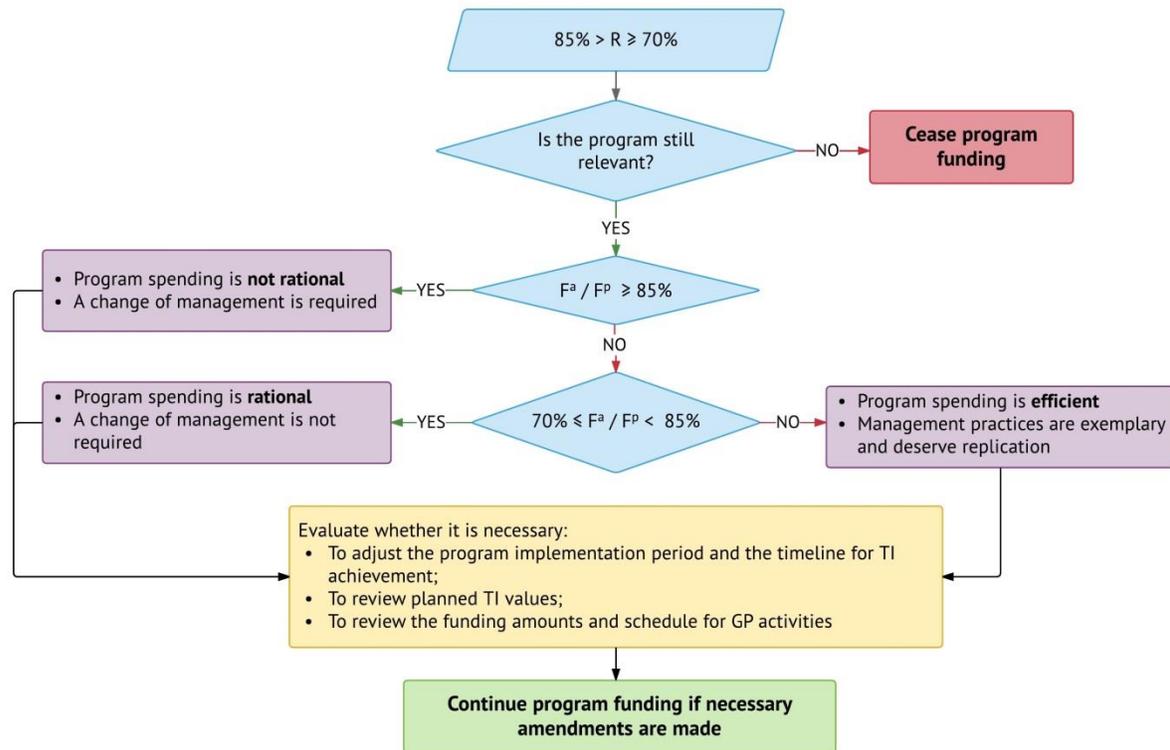


Figure 1 Ex post evaluation of a government program

Conclusions

1) Improving the quality and performance of environmental protection programs implemented by the government is one of the most important factors in preventing environmental disasters. Using theoretically grounded methodologies for program evaluation would allow us to debunk the myth that the assimilative capacity of the natural environment is practically endless, and to minimize the chances of managerial decisions being focused on short-term gains rather than informed by the assessment of possible long-term impacts.

2) An examination of two Russian government programs, *Pure Water* and *Water Industry Development*, allows to identify specific drawbacks of the current approaches to ex ante appraisal and ex post evaluation of environmental programs, such as:

- non-compliance of the target indicator selection methodology to the SMART principle;
- underestimation of the significance of budget efficiency indicators for funding-related decision-making; and,
- application of insufficiently justified simplifications when evaluating the ex post performance of government programs.

3) The key directions for the improvement of ex ante appraisal of government programs are as follows:

3.1. To expand the existing approaches to TI selection by:

- shortlisting target indicators in a way that takes into account their relative weights and excludes overlaps both between indicators within a single program, and between indicators for different related programs;
- departing from the common practice of using target indicators that lack macroeconomic prerequisites.

3.2. To include a new mandatory requirement into the ex ante appraisal methodology to calculate net present budget value (NPBV) and use it as an indicator of possibilities for the attraction of extra-budgetary financing and one of the criteria for the selection of the optimum approach to program implementation.

4) The author proposes a modified PART approach to the evaluation of environmental programs, which has not yet gained widespread acceptance in Russia. The main differences between this modified approach and the original methodology are as follows:

4.1. To introduce additional answer options besides the traditional yes/no dichotomy for the first three PART sections: Program Purpose and Design, Strategic Planning, and Program Management. Due to the particular nature of environmental programs, rigid adherence to the “yes or no” dilemma merely increases the degree of imprecision, which is why it is advisable to use multiple, more flexible options for expert appraisal.

4.2. To replace the expert survey by an analysis of TI values when evaluating direct program outputs. The individual rating score of the Program Results section at any point in its implementation (generally at the end of the calendar year) is calculated with due regard to the relative significance of different indicators, which is determined by expert appraisal of weighting coefficients for each of them at the planning stage of the program with the possibility of subsequent adjustment.

4.3. To introduce a flowchart to guide decision-making on whether to terminate or continue the program on the basis of its overall evaluation rating and the degree of conformity between actual and planned volume of financing. This logical algorithm illustrates a formalized procedure for the adjustment of the program implementation period and schedules for the achievement of target values for individual indicators; review of target indicator values, funding amounts and schedules; and for change of management. In most cases, the application of the proposed flowchart would prevent a significant mismatch between actual and planned TI values and expenditure by the end of program implementation.

5) In general, the proposed recommendations for improving the ex ante and ex post evaluation of environmental government programs are aimed at enhancing the objectivity of decisions taken by authorized government agencies concerning the advisability of budget funding of these programs, as well as the continuation or termination of their implementation.

REFERENCES

- Afanasiev, Mstislav and Natalia Shash. 2013. "Instrumentarii otsenki effektivnosti byudzhetykh programm." [Tools for evaluating the efficiency of budget programs]. *Voprosy gosudarstvennogo i munitsipal'nogo upravleniya* 3.
- Gilmour, John B. 2007. "Implementing OMB's Program Assessment Rating Tool (PART): Meeting the Challenges of Integrating Budget and Performance," *OECD Journal on Budgeting* 7(1).
- Kuzmin, Alexey, Rita O'Sullivan and Natalia Kosheleva (eds.). 2009. *Otsenka programm: metodologiya i praktika* [Program evaluation: methodology and practice]. Moscow: Presto-RK.
- Margolin, Andrey. 2012. "Sovershenstvovanie metodov otsenki planiruemoi effektivnosti regional'nykh gosudarstvennykh program" [Improving the methodology of ex ante appraisal of regional government programs]. *Gosudarstvennaya Sluzhba*, 6(80).
- . 2013. "Kriterii effektivnosti pri realizatsii gosudarstvennykh program" [Efficiency criteria in the implementation of government programs]. *Gosudarstvennaya Sluzhba*, 2(82).
- Metodicheskie ukazaniya po razrabotke i realizatsii gosudarstvennykh programm Rossiiskoi Federatsii* [Guidelines for the development and implementation of government programs in the Russian Federation]. 2013. Moscow.
- Office of Management and Budget. 2008. *Guide to the Program Assessment Rating Tool (PART)*.
- Robinson, Marc. 2013. *Program Classification for Performance-Based Budgeting: How to Structure Budgets to Enable the Use of Evidence*. Washington: The World Bank.
- Shepherd, Robert P. 2012. "In search of a balanced Canadian evaluation function: getting to relevance." *The Canadian Journal of Program Evaluation* 26(2): 1-45.

Published by the Centre for European Studies at Carleton University, Ottawa, Canada
Available online at: journals.carleton.ca/rera

RERA is an electronic academic peer-reviewed journal. Topics relate to the European Union, its Member States, the former Soviet Union, and Central and Eastern Europe. The journal is a joint project supported by the Canada-Europe Transatlantic Dialogue—a cross-Canada research network supported by the Social Sciences and Humanities Research Council of Canada (SSHRC)—along with the Institute of European, Russian and Eurasian Studies (Carleton University) and its associated research unit, the Centre for European Studies.

RERA aims to provide an accessible forum for research, to promote high standards of research and scholarship, and to foster communication among young scholars.

Contact:

Carleton University
The Centre for European Studies
1103 Dunton Tower
1125 Colonel By Drive
Ottawa, ON K1S 5B6
Canada

Tel: +01 613 520-2600 ext. 3117; E-mail: rera-journal@carleton.ca

Creative Commons License



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/) (CC BY-NC-ND 4.0).

Articles appearing in this publication may be freely quoted and reproduced, provided the source is acknowledged. No use of this publication may be made for resale or other commercial purposes.

ISSN: 1718-4835
© 2017 The Author(s)