2.2.4 STI policy evaluation and feedback

ISOGAWA Daiya[1]

First Published August 28, 2018 Final Updated July 2, 2021

Abstract

This paper explains the evaluation required at each stage of the STI policy process and the policy evaluation and feedback process. As the role of evaluation and the evidence relied on differs at each stage, there are unique issues for each of the pre-evaluation, process evaluation, and post-evaluation stages. STI policies are prone to positive externalities and can have long-lasting social and economic impacts. Consequently, target setting itself is a point of contention for socioeconomic evaluation. In order to conduct socioeconomic evaluation, including quantitative evaluation, it is important that scientific data are open to a wide range of actors.

Keywords

Policy process, ex-ante evaluation, process evaluation, ex-post evaluation, socioeconomic evaluation

Main text

The STI policy process is completed by an evidence-based evaluation of the policy and the ensuing feedback. Focusing on the policy for science, technology, and innovation (STI) within the STI policy process, this paper explains the evaluation required at each stage of the policy process and the policy feedback process. Following an overview of the role of evaluation according to the stages of the process, this paper sets out socioeconomic evaluation as a method of evaluation.

1 Associate Professor, Osaka City University Faculty of Economics.

The importance of evaluating STI policies based on evidence[2] has been highlighted in both the Science and Technology Basic Plan and the Comprehensive Strategy for Science, Technology and Innovation, which is prepared annually based on the Basic Plan. In the Fifth Science and Technology Basic Plan, a section entitled "Pursuing effective STI policy and enhancing the 'control tower' function" presents the goal of "promoting the planning, evaluation, and reflection of policies based on objective evidence" by "introducing a system for systematic observation and analysis of possible future developments in the economy and society based on objective evidence, and promoting the systematic development of data and information, indicators, and tools for evaluating and analyzing the effects of policies." Moreover, in a section entitled "Action to broaden effective public-private R&D investment based on objective evidence," the Strategy for Science, Technology and Innovation (2017) notes that "it is necessary to establish a PDCA cycle for policies based on objective evidence that clarifies inputs and outputs, as well as outcomes, by having the government 'make visible' R&D investment and policy effects, providing decision-making materials for policy formation, realizing appropriate resource allocation and evaluation, and disseminating and sharing information externally."

1 Evaluation at each stage of the policy process

As summarized in the previous section (2.1.2), the STI policy process begins with the discovery and setting of an agenda.[3] Needs assessment can be cited as a method of evaluation at this stage (e.g., Akiyoshi Takao, Ito Shuichiro, and Kitayama Toshiya, 2015). This is performed using existing statistical data, new surveys or communication with various actors to determine whether a policy response is needed to address the issue in question. While identifying the agenda is primarily the responsibility of the government, experts also play a significant role in this regard as the linkages between the economy and society and science and technology become increasingly complex (Arimoto Tateo, Sato Yasushi, Matsuo Keiko, and Yoshikawa Hiroyuki, 2016). By establishing a system that involves various stakeholders—government, experts, and citizens—in the discovery and setting of issues, it is possible to broaden the set of choices in the early stages of the policy process.

Once the agenda is established, preliminary evaluation is required in the process of setting policy options and making policy decisions. This is a measure to compare and weigh multiple options to determine what policy objectives should be set for an issue and what policy measures should be used to achieve these objectives. In preliminary evaluation, it is necessary to predict and estimate results from limited information, and it is especially difficult to make predictions and estimates for situations that have yet to manifest. If there are limitations on data collection, an approach that predicts policy effects based on certain assumptions, such as counterfactual simulation analysis, can be considered. Counterfactual simulation analysis is a way of thinking in which the outputs and outcomes to be realized if a policy is implemented

2 The term "evidence" used in this paper conforms to that used in Section 2.2.1. In downstream policy processes, we focus on "scientific facts" from the available types of evidence; this is not the case in the upstream process.

3 Note, this paper and "policy process" in Section 2.2.1 are different in terms of approach. This paper is concerned with the process of policy formulation and execution, and builds upon the development in Section 2.1.2.

are simulated in advance, and the difference from reality (i.e., baseline) is taken to be the effect/cost. At this stage, a policy's output or outcome has yet to be realized, which means that the analysis must necessarily capture the counter-factual situation. For this reason, there are methods to construct a model (economic model) that captures how the actors affected by a policy change their decisions and behavior as a result (e.g., Reiss and Wolak, 2007).

In the policy implementation process, process evaluation is required to assess whether a policy is being implemented properly, especially for STI policies that tend to be implemented over long periods of time. At this stage, it is necessary to review the relationship between the policy objectives and policy instruments set during the preliminary assessment, and to take into account changes in constraints and the occurrence of problems that could not be foreseen at the outset.

Depending on the results of the evaluation, there may be options to modify or change the policy, or even cancel it. However, in doing so, care must be taken to prevent falling into the "sunk cost fallacy" (e.g., Arkes and Blumer, 1985). Sunk costs are the costs of resources that have already been invested and cannot be diverted or recovered when a project is discontinued or rolled back.[4] Although a comparison of the "costs and benefits of continuing" and "costs and benefits of discontinuing" is of interest when making decisions about business continuity, both costs include sunk costs, which should be separated from any rational decision to continue or discontinue. However, in the real world, we often find ourselves in situations where we are unable to stop a project because we regret the resources that have already been invested in it, even if these are a sunk cost. This phenomenon is known as the sunk cost fallacy or Concorde fallacy.[5] It is important to be aware of sunk costs even when judging whether a policy should be continued or not—a point reflected in some policy evaluation guidelines and manuals.[6]

Ex-post evaluation is conducted a certain period of time after the policy is implemented. In this regard, it is necessary to evaluate the effects of policies by measuring changes in outputs and outcomes, and to examine the causes of any differences from preliminary evaluations. One of the roles of ex-post evaluation is to fulfill the obligation of accountability to various stakeholders, including the public. Accordingly, it is necessary to conduct evaluations using indicators that are easy to understand and handle. The results of the ex-post evaluation itself can also be used as new evidence. While process evaluation is positioned as a means of providing feedback to the policy in question in the process of policy implementation, ex-post evaluation serves as a means to provide feedback to another policy. In this regard, it should be recognized

4 For example, investments paid for R&D and human resource development are difficult to recover when withdrawing, and thus tend to be sunk costs. On the other hand, because purchased cars, land, and so on can be put to other uses (e.g., sold on the market), their sunk costs are small.

5 An example of this is the circumstances surrounding the development of the supersonic airline Concorde, from which the term originated (e.g., Nagase Katsuhiko, 2008). In the 1960s, the United Kingdom and France began joint development of the Concorde, but it later became clear that the development costs would far exceed initial estimates. At that time, the decision could have been made to cancel development and withdraw from the project, but the governments of both countries—having already invested huge sums of money—decided to make further investments. This project was considered a commercial failure, and all Concorde aircrafts were retired in 2003.

6 For example, the "Manual for Cost-Benefit Analysis of Water Supply Projects" (Ministry of Health, Labour and Welfare) clearly states that sunk costs should not be considered when calculating the cost of discontinuing a project.

that the evaluation of a policy process is not limited to the policy in question, and that the evaluation of a policy can lead to the setting of a new agenda.

2 Socioeconomic evaluation

In conducting STI policy evaluations, it is necessary to draw upon evidence to assess the various impacts of the policy on society and the economy. As Japan and other countries face limitations in their fiscal resources, socioeconomic evaluation is playing an increasingly important role in making the policy process more rational. Moreover, in terms of accountability to citizens, we can point to the importance of policy evaluations based on objective evidence.

In socioeconomic evaluation, it is necessary to identify the human and financial resources (i.e., inputs) needed to implement the policy and the outcomes of the policy, respectively. Policy outcomes can be classified in terms of outputs, outcomes, and impacts.

Shineha Ryuma (2017) regards outputs as the direct products of policy implementation (e.g., publication of papers), outcomes as the short- to medium-term results of outputs (e.g., creation of new drugs and patents), and impact as the broader effects that become visible over a longer period of time.[7] STI policies tend to have positive externalities (i.e., spillover effects) and can have long-term social and economic impacts. Accordingly, the qualities of STI policy evaluation can be placed within the framework of the Impact Assessment (TA), which evaluates the broader social and economic impacts of planned and implemented actions and policies. Impacts vary depending on the field, times, society, and so on. Therefore, it is essential to always be aware of what the target of policy evaluation should be.

Socioeconomic evaluation often requires that quantitative evaluation be conducted. I will leave the introduction of the specific framework of the evaluation method to Chapter 4, but I would like to point out here that data, the basis for scientific facts, are essential for quantitative socioeconomic evaluation. In order for discussions to be open and based on scientific evidence, it is particularly important that data are open to actors such as academia and think tanks. More specifically, we need data on both the costs and the effects of policies. With regard to data concerning costs, in addition to direct cost expenditures, there may be situations where there is interest in social costs for policies with large spillover effects. This is also true in respect to the effects, especially in STI policies, where spillover effects are often hard to ignore. For example, when considering measures to promote R&D investment in the private sector, given the possibility of social and economic spillovers from the results of R&D investment (e.g., Griliches, 1992), cost-benefit data may be needed not only for the direct impact of the policy on the target firm, but for broader targets such as markets and industries.

Naturally, the data used at each stage of the policy process will be different. In the preliminary evaluation stage, projections have to be made based on pre-implementation data (e.g., R&D expenditures of the target company or industry). On the other hand, it is not always easy to identify the effects and costs of policies,

7 There are also concepts that include long-term consequences in outcomes (e.g., "Toward the Promotion of Social Impact Evaluation," the Cabinet Office), so care should be taken not to confuse this point.

although it is possible to use information on the costs and effects that have already emerged in the process evaluation and ex-post evaluation stages. Given the possibility of being affected by other policies and external factors (e.g., exchange rates), it is necessary to be creative in estimating the net effect and net cost of the policy in question. The counterfactual simulation analysis mentioned above is one promising approach in this regard. The idea is to predict the counterfactual situation supposing the policy is not enforced, and examine the difference in costs/effects with the actual situation in which the policy was implemented (i.e., the baseline).

Finally, I would like to reiterate that we need to keep in mind the timeframe of the policy evaluation. This is because with STI policies, where outcomes tend to take time to be realized, there may be situations where short-term costs and benefits diverge from long-term costs and benefits.

References

Arkes, H. R. and Blumer, C. (1985). The psychology of sunk cost. Organizational Behavior and Human Decision Processes, 35(1):124–140. http://www.communicationcache.com/uploads/1/0/8/8/10887248/the_psychology_of_sunk_cost.pdf

Griliches, Z. (1992). The search for R&D spillovers. Scandinavian Journal of Economics, 94:29–74. https://www.researchgate.net/publication/4747485_The_search_for_RD_spillovers_Scand_J_Econ_94Suppl29-47

Reiss, P. C. and Wolak, F. A. (2007). Structural econometric modeling: Rationales and examples from industrial organization. In Heckman, J. and Leamer, E., editors, Handbook of econometrics, volume 6, 4277–4415. Elsevier. https://pdfs.semanticscholar.org/f0c5/8744d6db36843a9c7c42906c41cd2c25206e.pdf

Akiyoshi Takao, Ito Shuichiro, Kitayama Toshiya (2015). Kōkyō seisaku-gaku no kiso (shinpan) [Foundations of public policy studies, New Edition]. Yuhikaku Books, Tokyo. http://www.yuhikaku.co.jp/books/detail/9784641184282 [In Japanese]

Nagase, Katsuhiko (2008). Ishi kettei no manejimento [Managing decision-making]. Toyo Keizai Inc. https://books.google.co.jp/books?hl=ja&lr=lang_ja|lang_en&id=URIFBAAAQBAJ&oi=fnd&pg=PA5&dq=意思決定のマネジメント&ots=6eGxGW73H9&sig=Y52SG90dmbat4xsgnf8wIE50I-g [In Japanese]

Shineha Ryuma. (2017). Assessment of “impacts”: Science and technology policy and research evaluation in transformation of science and technology policies in the post-Cold War era, 39–53. National Diet Library http://dl.ndl.go.jp/info:ndljp/pid/10314914?itemId=info%3Andljp%2Fpid%2F10314914&__lang=en [In Japanese]

Arimoto Takeo, Sato Yasushi and Matsuo Keiko (2016). Kagaku-teki jogen: 21 Seiki no kagaku gijutsu to seisaku keisei [Scientific advice: Science, technology, and policymaking in the 21st century]. University of Tokyo Press.

https://www.jstage.jst.go.jp/article/johokanri/59/10/59_716/ _article/-char/ja/ [In Japanese]