February 12, 2016

Policy research report

How policy formation changes with evidence:

Trends toward evidence-based policy in the UK, as verified by randomized controlled trials, and some implications for Japan

Naoyuki IEKO, Yohei KOBAYASHI: Senior Analysts, Economic & Social Policy Departments

Natsuko MATSUOKA: Senior Analyst, Environment & Energy Department

Shinji NISHIO: Senior Analyst, Public Management & Regional Policy Department

Executive Summary

- Evidence-based policy (EBP), which advocates decision-making based on measures scientifically determined to be effective in achieving policy objectives, has been introduced and is progressing rapidly throughout the developed countries of Europe and the USA, as well as spreading to some developing countries via international development assistance. However, aside from a certain few policy fields, practical examples of EBP in Japan are extremely rare.
- In recent years, randomized controlled trials (RCT) have been of particular interest in regard to evidence for understanding the effectiveness of measures. Applying this kind of empirical approach enables more accurate measurement of the effective implementation of policy, although in Japan this approach is not yet in widespread use (other than in the area of medical research).
- For this paper, we undertook a study of the United Kingdom, in which EBP and empirical methods such as RCT are widely applied in the fields of social policy, as well as summarizing the results of a literature survey.
- Policymakers in the UK place a huge emphasis on "what works," and the national government recommends formation of evidence by taking an empirical approach. In doing so, by applying the theory and understanding of behavioral economics and by taking up small-scale measures and those having major financial impact, this approach has been empirically demonstrated to improve the effectiveness of policy measures and to make government spending more efficient, further accelerating the tendency toward an evidence-focused approach.
- In the UK, a number of intermediary support organizations have been established in the past few years, which have enabled groups such as local governments and non-profit organizations to undertake evidence-based practice and to make proper use of evidence. Of all these organizations, the public-private joint-funded What Works Centre is responsible for all three stages of evidence (the *generation*, *transmission,* and *adoption* of evidence), and it engages not only with the government, but with a wide range of stakeholders involved in policy, seeking to devise ways to encourage mutual cooperation.
- With a researcher-centered EBP foundation created more than 20 years ago and an ecosystem

Inquiries: Business Innovation & Creation Center, PR unit Phone: +81-(0)3-6733-1001 info@murc.jp

(of organic ties between various entities) for social implementation of EBP that has strengthened rapidly over the past few years, the United Kingdom has many recommendations to offer Japan. Particularly worthy of attention are the UK's efforts to seamlessly introduce EBP while staying on top of various issues such as devising empirical trials, selecting appropriate themes, and ensuring independence by utilizing public-private joint organizations, etc.

- In Japan too, there are opportunities to address this global trend and to expand the circle of dialogue and cooperation among all stakeholders.

Table of Contents

Inquiries: Business Innovation & Creation Center, PR unit Phone: +81-(0)3-6733-1001 info@murc.jp

1. Introduction

(1) Background

Faced with problems such as declining population, financial constraints, and growing future uncertainty, Japan needs more than ever to acquire a rapid grasp of social issues, to identify effective measures, and to implement these smoothly. However, for many policy measures in the real world, although acquiring a grasp of the issues has progressed to some extent, in the majority of cases we are still unable to identify effective measures, groping our way blindly toward measures while referring to the policy efforts of other countries or regional administrations when formulating our own policy measures[1]. While case studies of other countries and regional administrations may be effective in the context of those particular countries or regions, the same effect will not necessarily be exhibited in Japan or our other regions (lack of external validity). It is also undeniable that in groping our way toward policy, it is entirely possible that wasteful policies and inefficient policies will be furthered. Additionally, the speed of Japan's population decline and the severity of its financial constraints are unprecedented worldwide, such that many of the challenges facing Japan could be described as challenges that the world is facing for the first time, which makes seeking reference to foreign cases alone a difficult proposition.

As a result, for social issues on which opinions differ (or are opposed) among stakeholders, the formation of consensus and implementation of appropriate measures is delayed, leaving scope for guesswork or ideology to creep in. Given this situation, it is becoming increasingly important that we move ahead with policy based firmly on evidence: evidence-based policy (EBP). With regard to what is desirable policy, we are ultimately forced to rely on individual values. For example, to what extent should we ensure fairness of income and opportunity, and how do we strike a balance between efficiency and fairness? To what extent should we implement redistributive policies? Is it desirable to invest in the next generation, or should we provide lifestyles that ensure peace of mind for the elderly? There are diverse ways of perceiving the social issues that Japan has come to face in recent years, and our approaches to these issues must ultimately be considered on the basis of value judgments made by individuals. However, in order to implement these as policy measures, it is necessary to formulate social consensus, and to understand the evidence to determine which measures are needed in order to achieve certain policy objectives and what sort of social impact these policies might have. Having adequate material to make scientific and objective judgments has become a very important issue.

(2) The current state of EBP in Japan

In Japan, EBP is still in its infancy, with very little effort being made outside the area of medical research. A scientific approach to innovation policy has started to be applied at the Ministry of Education, Culture, Sports, Science and Technology (MEXT), and the EBP concept has been incorporated by the Ministry of Health, Labour and Welfare to improve cancer screening rates, but aside from these efforts the accumulation of empirical research is not underway.

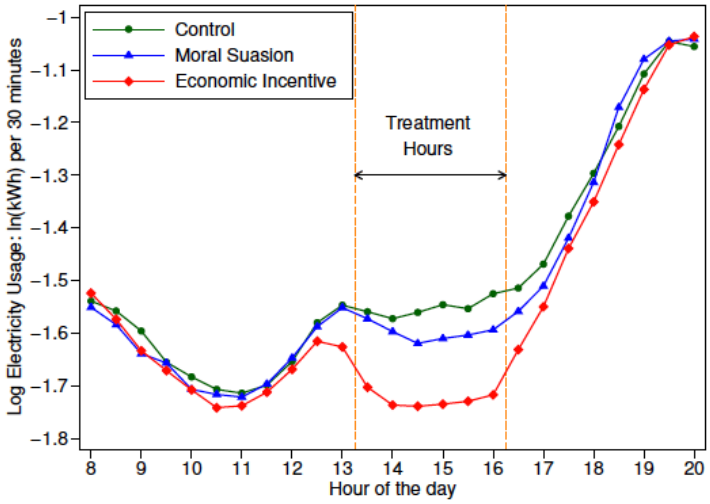
One of the few exceptions is the empirical research undertaken by Ito *et al.* (2015) in the energy sector.

---

[1] Referring to the Ministry of Internal Affairs and Communications (2012), "policy" is used here in the sense of "broadly defined administrative activities aimed at realizing basic principles in response to specific administrative challenges," while "measures" is used in the sense of “concrete measures and countermeasures used to achieve policy.”

Using randomized controlled trials (RCT), Ito *et al.* were able to verify the different effects on power consumption of economic incentives and a non-economic approach. In their study, about 700 Japanese households were randomly assigned to three groups—(1) a group motivated by economic incentive; (2) a group motivated by a non-monetary (moral) approach; and (3) a group that did not do anything (control group) — and the impact of this was verified. Figure 1. shows the results of this analysis. The red line labeled *Economic Incentive* shows Group (1), while the blue line labeled *Moral Suasion* shows Group (2) and the green line labeled *Control* shows Group (3). The analysis demonstrates that both moral suasion and economic incentive decrease power consumption, although the latter exerts a greater effect. In addition, the study also confirmed that the effect of policies motivated by moral suasion diminishes as they are repeated many times over, while the effect of policies motivated by economic incentive is sustained.

Figure 1. Power consumption of each group during policy implementation



(Source: Ito *et al.,* 2015)

Several reasons can be given for the lack of progress of EBP and RCT in Japan.

The first reason is the weakness of Japan's policy evaluation system and its low requirement for evidence. In policy formation, if not confronted with the pressing need to carry out effective measures, evidence is not particularly required, and as a result, there is less need for effective measurement using RCT[2]. As background to this, it is worth pointing out that PDCA is not well established in policy formation.

The second reason is the level of awareness of RCT. The RCT technique has been widely recognized in areas such as the medical field, but has had little recognition in government and policy formation. In addition, where simple before/after analyses and case studies of other regions are simply carried over, in many cases the results are misinterpreted and there is not adequate understanding of the need for verification or randomization.

[2] Professor Takanori Ida of Kyoto University, who conducted a RCT on power demand in cooperation with the Ministry of Economy, Trade and Industry, has pointed out that behind the Ministry of Economy, Trade and Industry's eagerness to measure the effect of policy lay an enthusiasm to pursue the budget of the smart community through division of business (Ida & Sawada, 2015). On the flip side, it can also be said that rigorous measurement of policy effects has not been required previously.

The third reason is the psychological resistance to experimentation. When measuring policy effects using RCT, it is necessary to divide subjects into treatment groups (benefiting from the policy measures) and control groups (not benefiting from the policy measures). However, since the government often requires fairness of policy, if only a section of the community benefits from the effects of the measures, some critical backlash is likely. In addition, there is often a strong feeling of resistance to the word "experimental" itself.

The fourth reason is “infallibility of the government (bureaucracy).” Traditionally in Japan, custom dictates that the government is never wrong—and if it does make a mistake, it will never be admitted. This custom is known as "infallibility of the government (bureaucracy).” Due to the implicit assumption that if the government is using the budget, this must be a secure decision, there is considerable resistance to empirical techniques such as RCT where the presence or absence of effectiveness is not clearly known in advance. In particular, if costs (such as time and expense) are incurred, accountability to taxpayers is also incurred, making it difficult to take the first steps in a new direction.

Reasons such as these provide the background for the delayed progress of EBP, and in particular RCT, in Japan[3].

[3] Steven Levitt and Stephen J. Dubner, co-authors of the *Freakonomics* series, cite three reasons why RCT has trouble making progress. These reasons are: 1. Custom/habit; 2. Lack of know-how; and 3. Lack of courage to say "I don’t know." (Levitt & Dubner, 2015).

2. The progress of EBP in the UK and implications for Japan

Although the introduction of EBP has lagged in Japan, there are widespread efforts worldwide to make use of the high-quality evidence obtained through RCT in policy formation. This paper examines the United Kingdom as one example of a nation that emphasizes evidence in policy formation as well as in practice, and goes on to discuss the implications for Japan.

(1) Overall picture of EBP stakeholders in the UK: Mapping

Figure 2. shows a map of the overall picture of EBP stakeholders in the UK. The vertical axis shows function. *Producers* are entities that primarily produce evidence, while *Intermediaries* (intermediate support organizations) are entities that organize and sort this evidence and *Consumers* are entities that make use of this evidence. The horizontal axis shows types of organization, broadly classified into *Government*, *Public-private joint organizations*, and *Private sector*. The black-bordered fields indicate institutions that were visited as part of the field study for this paper.

As for the government, the UK Office for National Statistics (ONS) is responsible for the collection and publication of statistics across the United Kingdom, and can be described as an institution that mainly produces evidence to help understand current circumstances. The NAO (National Audit Office) is an independent parliamentary organization whose role is to examine the appropriateness of government expenditure. Aside from these, researchers and economists in the government are also responsible for collecting and organizing evidence. Solace (The Society of Local Authority Chief Executives and Senior Managers) is a network of local government officials that organizes and advocates good practice of EBP in local governments.

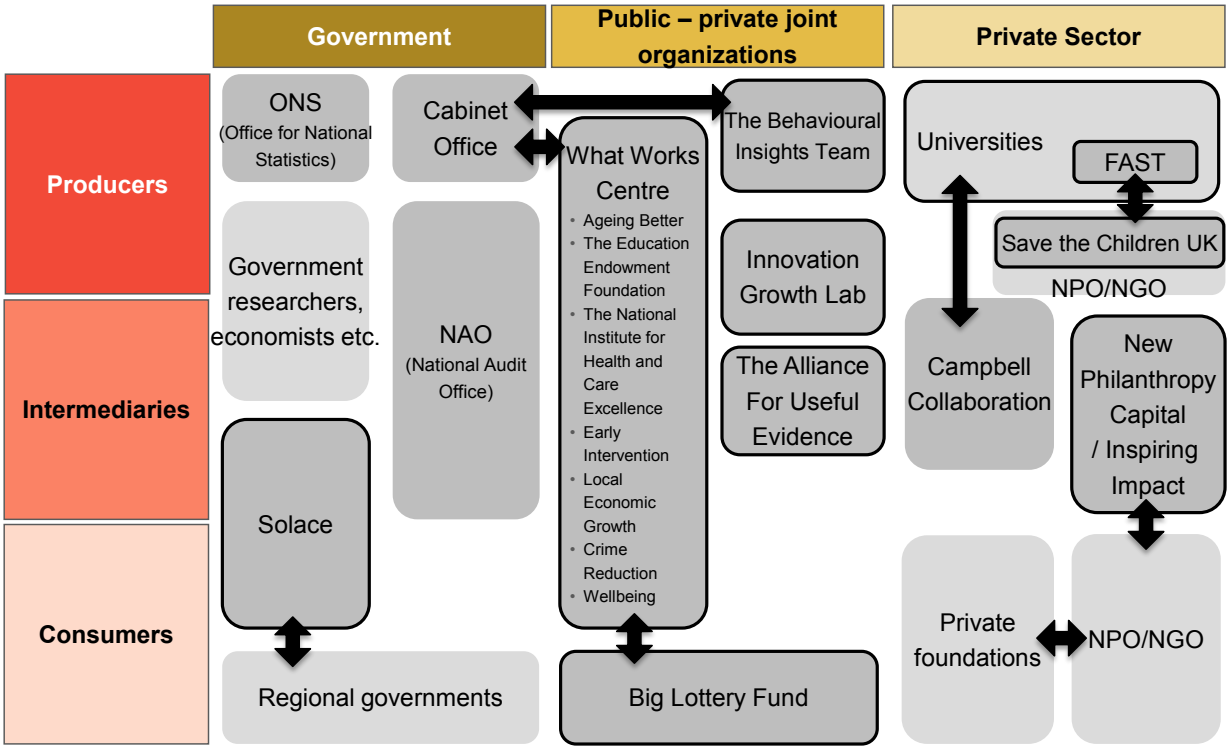
Among public-private joint organizations, the What Works Centre (WWC) also plays an important role as a producer, intermediary, and consumer of evidence. The Cabinet Office has certified the WWC as an organization that undertakes generation, transmission, and adoption of evidence. At present, the WWC is established in seven areas: Ageing Better, the Education Endowment Foundation, the National Institute for Health and Care Excellence, Early Intervention (for children and youth), Local Economic Growth, Crime Reduction, and Wellbeing. Unlike more general research institutions, the WWC's role is to organize evidence on cost-efficiency and effectiveness in order to contribute to decision-making by policy-makers and practitioners. The WWC's management format differs for each field, but basically in addition to funding from both public and private foundations (such as ESRC and Big Lottery Fund), it is also financed by funding from the respective ministries[4]. The Behavioural Insights Team (BIT, aka the Nudge Unit) is an organization created to promote effective and efficient policy-making based on the principles of behavioral economics, and it also serves as an advisory body to the WWC. The BIT was established as a Cabinet Office organization in 2010, and is currently operated as a public–private funded organization that keeps close relationships with various ministries and agencies, including the Cabinet Office. The Innovation Growth Lab (IGL) is an organization that supports RCT studies on innovation policy and industrial policy, and also consolidates and reviews this RCT evidence. The Alliance for Useful Evidence (AUE), which operates as a hub for evidence in the UK, is an organization established by funding from Big Lottery Fund, ESRC, and Nesta. Institutions such as BIT, IGL, and AUE

[4] Cabinet Office "What Works Network" https://www.gov.uk/guidance/what-works-network

are in some aspects similar to government agencies, but they also maintain their independence and undertake their own work. In the UK, these sorts of organizations are known as “arm's length” or “skunkworks” organizations.

Among the private sector, universities are positioned as important players in EBP, as they provide academic knowledge and also produce evidence. They play a significant role in providing state-of-the-art theory as well as disseminating techniques to review empirical methods and evidence (including RCT), and they also cultivate human resources with practical skills for each stakeholder. Examples include York University’s York Trials Unit, which undertakes RCT in a broad range of policy areas, the University of London Institute for Education’s Evidence for Policy and Practice Information and Co-ordinating Centre, which teaches a method known as “systematic reviews,” and Oxford University’s MSc in Evidence-Based Social Intervention and Policy Evaluation, which offers a sequence of processes from social research through to practice. The Campbell Collaboration is an international effort to undertake reviews and summaries of evidence in the social sciences. New Philanthropy Capital (NPC) is a think-tank that connects the social sector with funders, while Inspiring Impact is a project being implemented by NPC in cooperation with seven other similar organizations, which creates guidelines and provides tools for the social sector to measure the social impact of its own activities. While NPOs/NGOs are organizations that make good use of evidence in their own activities, at the same time they also play a role in creating evidence. For example, Save the Children UK operates an early intervention program using RCT called FAST (Families and Schools Together) in collaboration with university researchers.

Figure 2. Mapping EBP stakeholders in the UK



(Source: Created by Mitsubishi UFJ Research & Consulting on the basis of a field study and the Alliance for Useful Evidence (2015))

(Note: Black-bordered institutions were visited as part of the field study. Arrows indicate institutions with particularly close relationships. The pale colors here indicate generic designations for entities/organizations, rather than specific groups.)

(2) Driving forces for EBP in the UK

In the United Kingdom, EBP is well underway in a broad range of fields, including medical care, education, poverty, industrial policy, regional revitalization, and more. The following factors can be singled out as driving forces for EBP.

The first factor is the thickness of the EBP “ecosystem” (the organic ties between various entities). Although EBP has made good progress in the UK, this has not necessarily been brought about by any single entity or event. When interviewed in the field study, many stakeholders agreed in describing the importance of this ecosystem. In the medical field, efforts to promote EBP (including RCT) have been underway for some time, and as a result of the international efforts in collecting and evaluating evidence in the medical field made by the Cochrane Collaboration in 1993, along with recognition of the importance of evidence, an international community was formed. In response to this process of evidence as utilized in the medical field, the Campbell Collaboration commenced in 2000, resulting in more widespread recognition of the importance of evidence as utilized in the fields of social science and policy. In addition, the UK government has also been aiming for outcome-oriented policy formation. These steady efforts in various fields have formed an EBP ecosystem that can be said to have culminated in the establishment of groups reviewed in this paper, such as the WWC, BIT, IGL, and AUE. If counted from the launch of the Cochrane Collaboration, this is the result of long-term efforts over about 20 years.

The second factor is the impact of austerity measures. In response to the 2008 collapse of Lehman Brothers, a massive financial crisis and debt crisis occurred in Europe. Caught in the midst of this was David Cameron, appointed Prime Minister in 2010. Along with declaring his allegiance to the “big society” philosophy, Cameron sought to resolve the enlarged budget deficit and debt. In order to reduce expenditure, he needed to choose which areas and which policy measures would be subject to budget cuts, and how much funding to leave. More than even before, the efficiency and effectiveness of individual policies were subjected to strict oversight, with strong emphasis placed on utilization of evidence. In particular, following the BIT’s demonstration that financial effects far exceeding the additional costs can be achieved, the financial crisis coupled with the formation of the EBP ecosystem propelled EBP forward at a speed never before achieved.

The third factor worth pointing out is the environmental change of the social sector. Previously, little emphasis had been placed on the effects and social impact of the activities of the social sector. However, in recent years perspectives that consider the effects expected to result from individual activities, as well as their levels of social impact, have come to be emphasized in fundraising also. In our field interview with Save the Children UK, it was stated that demonstrating strict evidence is important in order to advance fundraising, with RCT in particular pointed out as a powerful tool for this.

It can be argued that this sort of multi-faceted environmental change lies behind the progress made by EBP in the United Kingdom.

(3) Advantages of EBP

The first advantage of EBP is that it can reveal what kind of policy is functional (*what works*) and what kind of policy is not functional (*what does not work*). As mentioned in 1. Introduction, although it is not always easy to accurately grasp the effects of policy, understanding the effects of policy contributes to more efficient use of tax funds. The fact that tools and processes such as RCT can be used to help understand the effects of policy offers a major advantage.

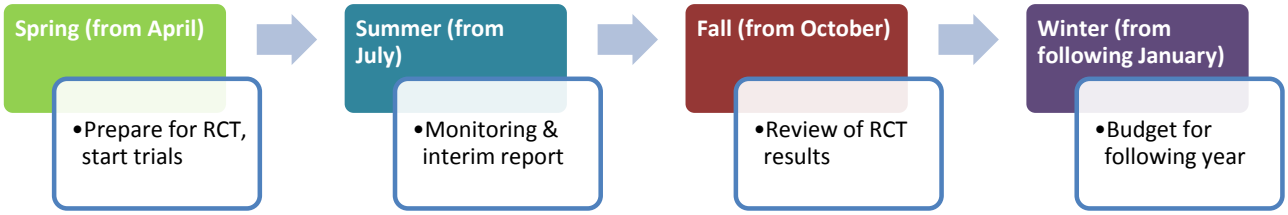
The second advantage of EBP is that it nurtures a culture of challenge in policy formation and administration, etc. Not only in Japan but in the UK also, accountability is required for all measures taken in politics and in governance, and attempting new things is not easy. However, with changing economic and social structures, new problems tend to arise one after another, and the correct answer is known in advance for only a very few policy issues. The role played by EBP in changing the culture of politics and government was pointed out in our field interviews, as well as the importance of constantly reviewing policies in both politics and administration through a process of trial and error ("experimental government").

The third advantage of EBP is that the losses incurred by doing nothing (maintaining the status quo) become more visible. As described above, undertaking new initiatives exposes policy-makers to the risk of failure and the risk of criticism. But this does not mean that doing nothing (maintaining the status quo) is justified. For example, in the case of the RCT on poverty measures for children, it is highly likely that the RCT method will be criticized, or that the effect of the measures cannot be detected. However, if neglected there is also a risk of leaving children in a state of poverty, or of implementing on a large scale policies of unknown effectiveness that place too much emphasis on fairness. Using evidence, EBP makes it possible to clarify the costs incurred by doing nothing and maintaining the status quo.

(4) Some points when promoting EBP

The first point when advocating EBP is to start with whatever can be achieved, even if very small in scale. If using RCT, evidence based on RCT is of the highest quality, as this enables elimination of various biases surrounding cause-and-effect relationship analysis. However, undertaking RCT requires a lot of time, effort, and money, making it difficult to undertake RCT for all policy measures. The social sector and local government sector in particular lack the capacity to undertake RCT, and it is not uncommon for evidence to be required in a short period of time. However, even evidence not produced through RCT is often useful practically. In order to proceed with EBP, rather than being strictly bound by the need for RCT, it is very important to make good use of evidence to any extent possible. Applying this to Japan's situation, since the national government and local authorities plan and implement their budgets in yearly units, conducting a small RCT over a short period of time to be reflected in the next fiscal year's budget is a good approach. It is also important to develop methods over a short-term cycle—for example, the Spring Test Autumn Review (STAR) format (planning and rapidly implementing a study in spring for verification in early autumn). Building up experience over short cycles can be expected to lead to more widespread practice of EBP.

Figure 3. Overview of the Spring Test Autumn Review (STAR) format



(Source: Created by Mitsubishi UFJ Research & Consulting)

The second point when advocating EBP is cooperation with stakeholders. Although EBP is understood as a scientific approach to policy, if EBP is promoted only among the central government and researchers, it can tend to focus on policy themes with low need or urgency in the real world, mistakenly affirming the effects of policies that are already evidently ineffective and leading to the concern that EBP might unnecessarily further increase the burden on real-world practitioners. On the other hand, without access to the state-of-the-art knowledge, resources, and advanced theory that the central government and researchers possess, there is always the concern that strictly “bottom-up” efforts by practitioners will not lead to the formation of effective evidence. In order to promote EBP effectively, cooperation and dialogue with all stakeholders (including local governments and practitioners) is essential. This cooperation and dialogue can reveal what kind of problems are being confronted in the real world, the quality of the evidence currently known as well as the quality of the evidence sought, and the challenges for implementing EBP. Transplanting this to the Japanese context, we believe that the prefectures can play a greater role as wide-area administrations.

The third point when advocating EBP is the importance of highly independent administrative bodies (so-called “arm's length” or “skunkworks” organizations). Although promoting EBP and RCT has major advantages, given the aforementioned accountability requirement and the risk of failure, it is not easy for government to be directly involved from the beginning; thus, if public-private joint organizations that are highly independent from the government (such as WWC and BIT in the UK) are established, these will be capable of acting more flexibly than the government, as well ensuring neutrality of evaluation. Unfortunately, Japan has very few examples of collaborative organizations created by joint public–private sector investment. However, in the future, we can expect to see some consideration of this issue for specific social issues.

(5) Some points when promoting RCT

Although it is desirable to make use of a variety of evidence when promoting EBP, RCT are without doubt one very powerful tool for identifying cause-and-effect relationships in policy. On the other hand, RCT require a great deal of time, effort, and labor, and they expose practitioners to possible criticism in regard to fairness. Our field interviews also pointed out various ideas for promoting RCT.

Firstly, it is important to launch with small, cheap, non-political projects with huge payback. For example, the initial project conducted by BIT was a RCT on payment reminder letters for tax delinquency. BIT sent various types of letters to tax delinquents to verify what kind of message makes the most effective reminder. Since payment reminder notices for tax delinquents were already sent in the past, this RCT did not generate any additional cost by being “experimental.” In addition, since the RCT targeted people who were already behind on paying their tax, the fairness issue was only minor and there was no political confrontation in regard to prevention of tax delinquency. In addition, if the tax rate could be improved (even by just a few percent) by simply changing how reminder messages are sent, the impact on finances could be very significant. We believe that when promoting RCT in Japan, addressing these themes on a priority basis is important.

The second point is devising methods for randomization. With means of randomization, those assigned to the control group will often reject the policy as unfair, even if they wish to benefit from it; but by devising particular methods of randomization, it is also quite possible to mitigate this rejection. For example, if the control group subjects and the treatment group subjects can be swapped/reversed over different periods, the fairness of the policy is ensured. Also, instead of implementing policy all at once, if

the treatment group is selected randomly in the initial stages and then this is extended in a gradual rollout, it is possible to ensure fairness to a certain extent. If budget constraints make it difficult to extend the policy to all persons who wish to benefit from it, there are also methods to assign beneficiaries at random among the applicants (using wait lists), within the limits of the budget. Devising these sorts of randomization methods, combined with communicating effectively to participants, makes it possible to alleviate to a considerable extent the rejection of policy due to randomization.

The third point is ensuring the neutrality of research design and evaluation. When undertaking RCT, it is not uncommon to end up verifying the effects of the programs already being provided by service providers. But in such cases, if the service provider itself designs the study and evaluates the program, concerns arise that neutrality is no longer guaranteed. In fact, with the FAST program and the efforts of IGL, the service provider and evaluator roles are clearly separated. In order to maintain neutrality when measuring the effect of a program using RCT, the service provider and the evaluating body should be separate entities.

The fourth point is the construction of hypotheses for effective measures. Since the implementation of RCT generates additional costs (whether minor or major), and since RCT are carried out in the field of actual policy, "evaluation for evaluation’s sake" must be avoided. Therefore, measures that do not require proof (those already proven effective, or proven ineffective) need not be re-verified by RCT. However, it is not easy to come up with hypotheses for policy measures that "probably are effective, but have not been verified." Another important point in the practice of RCT is obtaining the involvement of experts in policy areas and researchers who strive on a daily basis to add the latest facts to the wealth of knowledge accumulated over many years.

The fifth point is the inclusion of local stakeholders and practitioners. The effectiveness of programs is dependent on specific regional circumstances, and even if the program is confirmed effective in a certain region, we cannot know if it will actually exert a similar effect in other regions. In addition, the designers of RCT, even though they may be experts in designing trials, are not necessarily experts in any given region or policy area. In order to properly design trials and programs in such circumstances, input from stakeholders and practitioners in the region is essential. If these stakeholders participate in the discussion from the design stage, this will help mitigate any rejection of the RCT. Rather than undertaking RCT in haste, it is important to invite stakeholder participation, taking into account the needs of the specific site, and to make steady progress.

3. Recommendations for Japan

Recommendation 1: Promoting introduction of EBP for more sophisticated policy formation

In Japan, the importance of policy evaluation has been argued for some time, and since the Government Policy Evaluations Act came into force in 2002, this area has been the subject of countless overhauls and efforts[5]. However, with regard to accumulated evidence about the cause-and-effect relationships of policy or outcome-oriented evaluations, we cannot really say that the PDCA process is functioning as effectively as intended. The UK has made significant progress on EBP, and one factor that sparked this off was fiscal austerity. In Japan too, coupled with an aging and declining population, financial constraints are becoming more severe every year. Given such circumstances, the "scrap-and-build" of policy measures (repealing of measures with relatively low effect and establishing of new measures expected to be effective)—in other words, greater selection and concentration of policy—is now essential. Japan's population constraints and financial constraints are an opportunity to promote the introduction of EBP, and as a result of the increasing outcome-orientation of policy, if we place a greater emphasis on "what works" in policy formation, we will also naturally achieve formation of more sophisticated policy through selection and concentration of policy.

Recommendation 2: Start from small-scale projects—establish a "Japanese BIT"

In order to make EBP (especially RCT) more widespread, greater awareness of its effects is essential. By taking advantage of the principles of behavioral economics, the BIT in the UK embarked on a small, cheap, non-political RCT project with huge payback, the effectiveness of which was highly visible. Subsequently, it undertook a strategy to broaden its scope to fields in which EBP is more difficult to implement and verify effectiveness, such as employment support. Even if awareness of the usefulness of EBP does spread in Japan, it will be important to start from small and easy-to-implement projects with verification over a short time-cycle, such as the Spring Test Autumn Review (STAR) format. However, since effective accumulation of experience and knowhow is key here, we believe that following the successful example of the UK—for example, creating a "Japanese BIT"—would also be effective.

Recommendation 3: Create guidelines or manuals aimed at local governments and the social sector, and offer support for "in the field" work

In Japan, awareness of the importance and need for EBP and RCT are still undeniably low. Because implementing new RCT and understanding the details of those RCT already implemented require a certain degree of statistical skill, this presents significant hurdles for various local governments and social sector organizations. In order to proceed with the introduction of EBP and RCT for local governments and the social sector, one good method is to create manuals of easy-to-use and accessible guidelines for workers and practitioners in the field. In fact, organizations such as the WWC, BIT, NPC, and AUE all summarize their RCT and systematic review methodologies etc. in an easy-to-understand form. Such guidelines or manuals would contain not only the technical aspects of EBP and RCT and

[5] Hatake (2005) etc.

various ways to devise randomization etc., but would also need to incorporate practical knowledge such as methods of engagement and communication for work in the field. In addition, since simply creating guidelines or manuals alone would run the risk of these not being used in practice, it is also necessary to establish a framework for intermediate support in the field, as well as efforts to raise public awareness through seminars, workshops, etc.

Recommendation 4: Accumulate, collect, organize, and advise on evidence—establish a "Japanese WWC"

In order to make best use of evidence in policy, it is essential to accumulate, collect, and organize this evidence. In the UK, the WWC (1) supports primary research through RCT (accumulation); (2) collects and organizes evidence through systematic reviews; and (3) advises desirable directions in the form of guidance—in other words, it plays the role of an "evidence center." In order to make best use of evidence in policy, it will be necessary to establish a body such as the WWC in Japan. In doing so, in order to ensure its independence and neutrality, this body should be not an internal government agency but should instead take the form of a public-private joint organization that enjoys a certain degree of independence. The creation of a Japanese version of the WWC can be expected also to contribute greatly to the formation of an EBP ecosystem in Japan.

Recommendation 5: Make RCT obligatory where possible for model projects

In Japan, although a number of model projects are already underway, these often come about due to local governments declaring their wish to implement such a project. In many cases, such local governments are highly motivated or have abundant resources in the first place, and so these are considered by other local governments to be special cases, which presents an obstacle when trying to horizontally deploy the results of such model projects to other regions. Furthermore, in almost all cases, randomization is not carried out and the cause-and-effect relationship of policy cannot be measured. As has been described in this paper, although some policies are not suitable for RCT, if implementation of RCT is made obligatory for those model projects for which national government assistance is available, this will bring progress toward the accumulation of high-quality evidence, which will bring significant benefits for the entire country. If we measure the effects of policy using rigorous methods such as RCT, it will be possible to predict the effects of deploying it to other regions more accurately than ever.

Following the process of making policy formation more sophisticated (Recommendation 1); undertaking small-scale projects with short-term cycles (Recommendation 2); creating guidelines for using evidence in the field (Recommendation 3); establishing a "Japanese WWC" as an evidence center to collect and organize the knowledge gained (Recommendation 4); and having projects for which evidence has been confirmed once again verified as national projects (Recommendation 5), we can expect that EBP will be applied across a wider range of policy areas.

4. In conclusion

This paper was written from a position of expecting full-scale utilization of policy based on evidence, particularly evidence revealing cause-and-effect relationships on the effects of public policy in Japan. Already in Japan, evidence is being used in a variety of policy areas to understand current circumstances, and in some policy areas (such as medical/healthcare) evidence is being used to understand cause-and-effect relationships from the policy-making stage. However, when compared to nations such as the UK and the USA, policy formation using evidence created in verification trials, as well as the set of processes leading to the effects of policy as intended by the policy-makers defining implementation on the practitioner level, is still in the untouched or experimental stage for many policy areas, and the EBP ecosystem is accordingly weak.

As mentioned in this paper, until the UK cross-agency organizations were fully established within government agencies with the objective of better utilization of evidence, the various stakeholders accumulated a lot of practical experience as well as a considerable period of trial and error learning. On top of such accumulation, due to a variety of overlapping factors (such as changing social issues, academic development, compromises between researchers and policy makers, conversion to fiscal austerity policies, and the strengthened foundation of the social sector), the past few years have seen a variety of new players, including public-private joint organizations, emerging one after another in the ecosystem. Social implementation of policy based on evidence is on the way to being fully realized. The austerity policies of recent years are only one trigger for the promotion of EBP; the ecosystem has been gradually nurtured over more than 20 years through the tireless efforts of the parties involved, with high quality evidence being gradually accumulated. However, this does not stop at simple cost-cutting under severe financial constraints—we should not overlook the fact that this will lead to the selection of effective policy.

The current trend worldwide is to discuss and determine policy on the basis of evidence, instead of determining policy only through ideology, with an emphasis on policy that is functional (what works). This approach is not only adopted in the UK and the USA, but is also rapidly being introduced in other developed countries too. Additionally, these efforts are also being deployed to developing countries via international development assistance, rendering this approach the global mainstream for policy formation. In Japan, in order to hasten the fully-fledged realization of policy based on evidence, by widening the circle of dialogue and cooperation to include not only policy-makers (including national and local governments), but also researchers, private companies, and related organizations (such as grant funding bodies and non-profit organizations), we can expect gradual development of this system across society as a whole.

5. Bibliography

1. Introduction

Centre for Public Scrutiny website: http://www.cfps.org.uk/# Accessed December 16.

European Venture Philanthropy Association (2013). *A Practical Guide to Measuring and Managing Impact.*

Koichiro Ito, Takanori Ida, & Makoto Tanaka (2015). *The Persistence of Moral Suasion and Economic Incentives.* RIETI Discussion Paper Series 15-E-014

Takanori Ida & Yasuyuki Sawada (2015). *Recommendations for Field Experiments*. Economic Seminar 2015 June–July issue (in Japanese)

Steven Levitt & Stephen J. Dubner (2015). *Think Like a Freak*. Diamond Publishing. (in Japanese)

2. The progress of EBP in the UK and implications for Japan

Cabinet Office (2015) Government – guidance What Works Network
https://www.gov.uk/guidance/what-works-network Accessed December 1.

3. Recommendations for Japan

Motoaki Hatake (2015) *The Current State of Policy Evaluation and Issues: 10 Years after the Revised Resolution, Based on its Verification.* In Legislation and Investigation (in Japanese)

Inquiries: Business Innovation & Creation Center, PR unit Phone: +81-(0)3-6733-1001 info@murc.jp