P15 / Policy Evaluation and Policy Process Studies

Topic : POLICY DESIGN, POLICY ANALYSIS, EXPERTISE AND EVALUATION

Chair : Giancarlo Vecchi (Politecnico di Milano)

Second Chair : Hiroko Kudo (Chuo University)

Third Chair : Bruno Dente (Politecnico di Milano)

GENERAL OBJECTIVES, RESEARCH QUESTIONS AND SCIENTIFIC RELEVANCE

Following the debate developed in Singapore 2017, the objective of the panel is to develop stronger relations between evaluation research and public policy analysis (Geva-May and Pal; Weiss; Vedung; Peters). The aim is to help evaluation researches in taking advantage from the studies of the policy process to improve their use by decision-makers.

In the last decades, evaluation scholars devoted strong emphasis and efforts in the debate about methodologies; the discussions concerning qualitative and quantitative approaches appear now passed, with efforts to deal with the complexity of policies using mixed methods.

In any case, the policy evaluation literature – both oriented to policy making and to knowledge development - seems to lack, often, systematic connections and exchanges with the researches of the different policy analysis fields; for example, the studies devoted to the analysis of policy learning, policy change, policy design and instruments, and policy implementation.

With the aim to develop a fruitful integration between these two sides, the panel wants to deeper themes that can be eventually discussed in different sessions. In particular, the objective is to analyze

a) the connection between evaluation researches and policy learning, analyzing the relation among experts and decision-makers;

b) the progress of the debate on evidence-based policy making;

c) the use of big data in evaluation researches.

CALL FOR PAPERS

The debate on ‘Evidence-based policy’ underline the issue concerning the role of social science in supporting the policy-making processes e the design of public programmes.

The reflection of many scholars is often translated in methodological terms, and the conclusion is oriented to the adoption of approaches that improve the reliability of the causal inferences that sustain the attribution of an outcome to a programme; and the preference of the RCT designs is quite often the result. Indeed we need to sustain a wider pluralism and expansion in the use of social science techniques in evaluation. For example, it is important obviously to collect evidences on ‘what works’ about public programs, but at the same time policy makers and citizens are interested in the transferability issue, i.e. the effectiveness of a program if implemented in another site; in that case we need to add to the ‘what works?’ questions regarding ‘what works, for whom, in what context’, the mechanism at the base of the observed changes, the reasons for the success of the winners and the failure of the losers, the implementation gaps, etc.

The objective of the panel is to improve the impact of program evaluation on the decision makers, avoiding the risk of policy-based evidences, developing a better conceptualization of collective problems and the feasible paths of solution.

We invite paper-givers to propose both theoretical and empirical contributions regarding three themes:

- the integration between program evaluation and policy analysis studies, concerning in particular evaluation useful to improve the policy design and policy implementation phases;

- the use of methodologies to deal with both the causal and the transfer problems, and the capacity to improve the plurality of approaches and research techniques;

- the role of big data in evaluation research.

P15 / Policy Evaluation and Policy Process Studies

Chair : Giancarlo Vecchi (Politecnico di Milano)

Second Chair : Hiroko Kudo (Chuo University)

Third Chair : Bruno Dente (Politecnico di Milano)

Session 1Policy Evaluation and evidence-based policy making

Friday, June 28th 08:00 to 10:00 (MB 2.270)

Discussants

Bruno Dente (Politecnico di Milano)

Giancarlo Vecchi (Politecnico di Milano)

Is Bigger Municipality More Vibrant Municipality? - An Ex-Post Evaluation from Finnish Municipal Mergers

Niko Vartiainen (University of Eastern Finland)

In recent years in Finland, the government has operated a policy, which has striven to produce public services in bigger units. For example, the so-called PARAS reform, which took place between the years 2007 and 2013, decreased the number of municipalities significantly. The objectives of the PARAS reform were to obtain “a vibrant municipal structure”, to moderate the cost-increase of municipalities, and to improve the productivity of the output of public services. The government’s main instrument to achieve these goals was to enhance municipalities’ incentives towards municipal mergers with, for example, municipal merger grants. However, one substantial problem in the government’s policy has been that the decision-making, which has concerned the municipal mergers and bigger units, has not been based on evidence-based policymaking, at least as sufficiently as it should have been.

This presentation concerns my ongoing PhD thesis research. The objective of the thesis is to answer the following research question: *How should the government regulate the formation of the Finnish municipal structure in the future?* An ex-post impact assessment of the PARAS reform is conducted to answer that question. In other words, the impacts of Finnish municipal mergers between the years 2007 and 2013 are being studied from the perspectives of cost, productivity and vitality effects. The main purpose of the study is to generate scientific evidence whether the legislators “bigger-unit -policy” has achieved the objectives it was aiming to achieve. This information can be exploited in the future when contemplating subsequent municipal reforms.

The biggest problem when studying the impacts of municipal mergers is separating the impact of the merger from all other factors that have an impact on the variable that the researcher is interested in. The problem can be solved with a Difference-in-Difference layout, where the development of the treatment group (merger municipalities) is compared to a certain control group. Because the selection of the treatment and control groups cannot be conducted by random sampling, the selection of the control groups must be done otherwise. A very popular method is to use propensity score matching, which finds for every merger municipality a counterpart municipality, which is as similar as possible compared to the merger municipality, but has not been involved in a merger.

In addition to studying the impacts of municipal mergers, the thesis includes an article, which seeks to find “an optimal municipal size”. In other words, the purpose is to analyze, whether the costs of municipalities decrease and/or the efficiency of production of public services increase, when population size of the municipality increases. “The finding” of the optimal municipal size is an integral part of the thesis, because the results give additional information to the legislator, whether or not to form bigger municipal units. This part of the study uses panel data from the years between 2015 and 2017. The data is analyzed with regression analysis, which controls other factors that have an impact on the development of costs and efficiency, so the actual impact of the population level can be detected.

Variations of evaluation functions in different administrative settings

Tomasz Kupiec (University of Warsaw)

Dominika Wojtowicz (Kozminski University)

Introduction

One of the development of recent years in the field of evaluation use is the application of the system thinking (Leeuw, Furubo 2008), focusing not on single study, but streams of studies flowing through evaluation systems (Rist & Stame 2006), searching for the explanation of evaluation use types and functions in the organizational context, and design of the system (e.g. Hanberger 2011).

As a part of this approach, and drawing from earlier literature on organizational institutionalism Hojlund (2014)developed original framework explaining evaluation use types observable in particular organization as a resultant of two variables:

- external pressure on evaluation adoption - cultural constraints, uncertainty or normative expectation in environment) (DiMaggio, Powell 1983),
- internal propensity to evaluate - political or action organization (Boswell 2008).

Although interesting and promising that framework left some questions unanswered. First – not all types of use and evaluation functions are covered by the framework. It does not deal with the process or strategic (advocacy) use. It also does not predict when evaluation is oriented on accountability, learning or supporting the policy design or implementation process. Second and the more intriguing one is that in some contexts we can observe organizations facing the same level of external pressure, demonstrating similar internal propensity and yet orienting evaluation in a different manner, which potentially leads to different evaluation use.

Our paper deals with such case and its goal is to propose alternative / supplementary explanation for evaluation practice orientation depending on administrative / policy settings in which organization operates. We present the example of European Union cohesion policy implementation system as it is organized in 8 so called ‘new’ member states. Although dealing with the same external pressure (same rules imposed by EU and no domestic evaluation culture) and internal propensity (similar programs implementing same policy) they developed significantly different evaluation systems.

Research questions

· what are the differences in evaluation systems of studies countries

· what determines the shape of evaluation system

· how differences In evaluation systems influence the actual evaluation functions fulfilled

Methods

Comparative study of 8 European Union member states was comleted. In each country we conducted online survey with heads of all evaluation units in the field of cohesion policy (N=80). Additionally we conducted in-depth interviews with representatives of leading evaluation units, local evaluation experts or members of evaluation societies (4 IDI per country) and desk-research study.

Contribution to the panel

Our paper contributes to the theme “the integration between program evaluation and policy analysis studies” by providing evidence on how different policy / administrative settings shape evaluation systems and functions.

We propose a two stage framework where administrative settings shape evaluation system, and later the system determines evaluation practice orientation. Therefore it is a policy / administrative context that in an indirect way decides what evaluation functions are fulfilled - accountability, learning, feeding into policy design or implementation stage.

We deal not with single organizations but large evaluation systems consisting of number of knowledge

producers and potential user organizations.

Reforming Statistics, Innovation Policy and Evaluation -Evidence-based Policymaking in Japan and the Crisis of Evidence-

Kazuya Sugitani (Iwate Prefectural University)

As in many developed countries, the Japanese government has attempted to promote evidence-based policymaking. One aspect of this effort is reforming government statistics.

Led by the Cabinet Office (Naikaku-fu), Japan's government has been supporting the reform of government statistics since 2016. However, the Cabinet Office has not formulated specific policies, as policymaking takes place at the ministry level. Accordingly, to promote evidence-based policymaking, the Cabinet Office has dispatched a "Director General for EBPM (Evidence-based Policy Making)" to the ministries. These directors are selected from among the ministry's bureaucrats. The overall aim is to improve statistics regarding specific policies and to encourage evidence-based policymaking.

Can such actions succeed? To answer this question, cases dealing mainly with scientific technology policies and scientific innovation policies are examined. For Japan's government, effectively utilizing science and technology is a highly important national goal in the age of economic decline; scientific innovation and economic growth represent the hope of politicians, bureaucrats and the Japanese people.

Science and technology policies and scientific innovation policies exemplify the need for policies that are evidence-based. In the government's "5th Science and Technology Basic Plan," the importance of innovation and evidence-based policymaking is made clear. Indeed, the plan established specific indicators for monitoring policy outcomes. This plan took effect in 2016 and is strongly related to reforming statistics. According to the plan, using indicators and sophisticated statistics will ensure that effective policies leading to fruitful innovation will be formulated and followed.

To achieve these goals, policy evaluation is necessary. However, Japan's evidence-based policymaking faces significant resistance from bureaucrats. Because in many cases they have already implemented policy evaluation, government bureaucrats have tended to view evidence-based policymaking as a matter of double-handling. In this presentation, the nature of evidence-based policymaking is discussed and important differences with existing policy evaluation are described.

Ominously, the resistance of ministry bureaucrats and initiatives taken by politicians have given rise to serious problems regarding evidence and policy. It has been recently revealed that many of Japan's ministries have falsified their data and that some bureaucrats have been involved in political corruption.

Why have these problems occurred? Is evidence-based policymaking simply naive thinking? To answer these questions, this paper examines the various policy processes and instruments of the Japanese ministries and administrations. The present situation in Japan regarding evidence-based policymaking as it relates to reforming statistics is described and the relationship of evidence-based policymaking to policy evaluation, with an emphasis on scientific innovation policy, is presented. Finally, the crisis of evidence and how to overcome it is addressed through an analysis of policy processes and institutions.

How Data and Evidences were ignored by the Policy-makers, while they reminded the Importance of Evidence-Based Policy-Making nationwide: Analysis of the 2020 Tokyo Olympic Game related policy making of the Tokyo Governor

Hiroko Kudo (Chuo University)

When the controversial parliamentary member Yuriko Koike candidate herself to the Tokyo Governor election in 2018, some predicted a type of leadership similar to that of the former Osaka Governor and City Mayor Toru Hashimoto, although they never shared the political parties nor political believes. Both conservatives, but never belonged to traditional parties and serial founder of new movements and new groups, they both share another aspect: a former McKinsey partner turned management professor Shinichi Ueyama and his team of consultants and academics. However their styles to which extent adopting the recommendations of the consulting team as well as "data" or "evidences" were completely different.

Indeed, Governor Koike announced and appointed a team of policy consultants above mentioned, in order to carry out the strategies, plans, and projects as well as to review the on-going projects started by the previous administration, which ended with early election caused by financial scandal of the previous Governor. In reality, however, the revision was slow and strongly influenced by political situation, not by data and evidences provided by the consulting team.

The paper is a direct observation of the author, who joined the above mentioned team for a year and a half,

from the summer 2016 to March 2018. The paper analyses a case of revision of the on-going projects, to which the new Governor promised revision during her campaign: the choices of several locations as well as the facility constructions for the Tokyo Olympic and Paralympic Games in 2020. The consultant team was asked to investigate the projects. They did the research and announced the facts with possible alternative solutions. The communication style of both the consultant team and the governor was initially embraced by media and the public, thus making the alternative proposals of the governor popular among citizen.

However, technical procedures as well as negotiations with stakeholders were stuck, and many members of the provincial Council opposed the Governor. The locations and facilities for Olympic Games remained as they were planned, despite the unreasonable choices and those evidences, while the media started to turn their back to the Governor and the approval rate of the Governor started to decline sharply.

The attempt by the team was to bring data and evidences in policy-making. The intention seemed to have been shared with the media and the citizen, but not with the traditional policy makers. Once the political negotiation proved difficult, even the Governor herself ignored the facts and pushed the previously approved projects, although they were clearly irrational choices.

The paper tries to analyse the story, focusing on the major stakeholders, especially the consultant team, the Governor, and the Council, and the relationship among them. The paper tried to find why evidence-based proposals failed to convince policy makers in Tokyo, using administrative techno-populism (Kudo and Sancino, 2016) and Blame Game framework proposed by Hood (2011).

P15 / Policy Evaluation and Policy Process Studies

Chair : Giancarlo Vecchi (Politecnico di Milano)

Second Chair : Hiroko Kudo (Chuo University)

Third Chair : Bruno Dente (Politecnico di Milano)

Session 2Evaluation: policy learning and policy design

Friday, June 28th 10:30 to 12:30 (MB 2.270)

Discussants

Bruno Dente (Politecnico di Milano)

Hiroko Kudo (Chuo University)

Seeking causal explanations in policy evaluation: applying process tracing to the case of the Barbados sugar-sweetened beverage tax

Miriam Alvarado (University of Cambridge)

Tarra Penney (York University)

Jean Adams (University of Cambridge)

Paper focus:

In this study, we tested two competing causal theories that may explain the results of an impact evaluation. Finding an association does not prove a causal relationship, nor does a statistical association necessarily provide policymakers with the evidence needed to inform future policy development. Instead, developing a stronger understanding of the causal explanations behind an association may prove more useful in guiding policy design.

Research questions:

Given evidence of a statistical association between a policy and an outcome of interest, 1) can we identify mechanisms through which the policy operated to increase our confidence in causality, 2) if so, what were these mechanism(s), and 3) what are the implications for policy design?

Case study:

We used the Barbados sugar-sweetened beverage tax (introduced in 2015) as a case study. Sugar-sweetened beverage (SSB) taxes have been proposed to address rising rates of obesity and type 2 diabetes. According to dominant theory, price change is the sole mechanism through which SSB taxes dampen consumer demand. However, SSB taxes may also have a signaling effect, stimulating awareness around the risks of SSBs and dampening demand independently of price. Some SSB taxes are introduced with little fanfare, while others have been associated with a lot of formal/informal signaling. Clarifying these mechanisms would have implications for the design of future taxes.

We selected the Barbados case based on evidence of an effect (a big-data based finding), and because we had access to a wide range of relevant data sources (60 semi-structured interviews with stakeholders and the public, archived newspaper articles, direct observation, and 200+ weeks of detailed grocery store retail data).

Methodology:

We chose to use process tracing (PT) for several reasons. First, the method is well-suited to within-case studies with a focus on causal explanation. Second, PT can improve transparency and efficiency of the research process, by requiring an assessment of prior confidence in hypotheses and guiding data collection efforts. Third, while PT has primarily been used with qualitative data, it is flexible enough to accommodate quantitative data (including big data insights).

Following PT best-practice, we developed causal theory, operationalized the theory into a series of tests,

and identified the probative value of each test in advance. We considered two competing causal explanations for the observed change in SSB sales. We assessed prior confidence in both hypotheses and pre-specified the evidence we would expect to find if each hypothesized mechanism occurred, drawing on all available data (both quantitative and qualitative). We identified what kind of test each piece of evidence would provide (straw-in-the-wind, hoop, smoking gun or doubly decisive). We prioritized tests with the highest probative value and updated our confidence in the hypotheses using results of the empirical tests.

Panel Fit:

This study demonstrates that an innovative small-n evaluation method can produce useful policy-relevant evidence beyond traditional impact evaluations. Understanding which mechanism(s) are in operation can help policymakers understand if results from elsewhere are relevant and inform policy design. We illustrate how big data insights can be integrated into evaluation.

Process Evaluation and Policy Design Improvement. The case of a welfare programme in Italy

Giancarlo Vecchi (Politecnico di Milano)

Marco Di Giulio (Università degli Studi di Genova)

The paper will reflect on the connection between the recent studies on policy design with the policy evaluation, with the aim to foster the theoretical basis of process evaluation and its usefulness.

The domain of policy design has been characterized in the last decade by a relevant debate, proposing different perspectives based on the analysis of policy instruments, the mechanisms to improve the engagement of users and implementors, the learning processes analysis, the ideational wave to analyze the role of ideas in supporting policy change, etc. On the other side, evaluation research is still discussing the ways to develop the theory of change approach, with little references to the policy analysis debate.

The paper is based on the evaluation of multi-level programme realized in Italy by the National Agency for Pensions and Welfare in the recent years, to help some categories of non-sufficient subjects. The multi-level characteristic of a national scale intervention introduced a relevant complexity in the implementation phase.

The objective of the paper is to develop, starting from the analysis of this implementation, some lessons to progress the connections between evaluation analysis and policy re-design.

References

Béland D. and R.H. Cox. (Eds). 2011. Ideas and Politics in Social Science Research. Oxford: Oxford University Press.

Dorst K. 2016. “Design practice and design research: finally together?” Design research Society 50th Conference, Brighton 27-30 June.

Howlett M. 2019 (2d ed.). Designing Public Policies. Principles and Instruments. Abingdon and New York: Routledge.

Howlett M. and I. Mukherjee (eds.) 2018. Routledge Handbook of Policy Design. Abingdon and New York: Routledge.

Pawson R. 2013. The Science of Evaluation. London: Sage.

Peters, G. etc. 2018. Designing for Policy Effectiveness. Cambridge: Cambridge University Press.

Rose R. 2005. Learning from Comparative Public Policy: A Practical Guide. London: Routledge.

Vedung E. 2009. Public Policy and Program Evaluation. New Brunswick: Transaction Publishers.

Weiss C.H. 1998. “Have We Learned Anything New About the Use of Evaluation?” American Journal of Evaluation 19(1): 21-33.

Toward the use of integration methods to support health policy learning in an evaluation context: a case of the UK Soft Drinks Industry Levy evaluation

Tarra Penney (York University)

Miriam Alvarado (University of Cambridge)

Jean Adams (University of Cambridge)

Martin White (University of Cambridge)

Paper focus

This paper is a conceptual piece that uses the case of the UK Soft Drinks Industry Levy (SDIL) evaluation to critically examine how embracing a wider pluralism and complex political context during evaluation can support health policy learning, but also present a set of challenges.

Research questions

- How can evaluation better support the development of causal explanations of if, how and why fiscal policy impacts health?
- What methods are available to support an integrative approach within a broader policy learning process?

Case details

In March 2016 the UK government announced the SDIL, designed to incentivise soft drinks manufacturers and importers to reformulate the sugar content below a set of thresholds. Unlike the majority of the taxes implemented globally to reduce the demand for soft drinks, industry was given a two-year window to reformulate eligible drinks to avoid paying the levy. However, other reactions such as passing on the cost or changing marketing strategies were possible. The introduction of the SDIL was covered extensively in the UK print and trade press, and the focus of a great deal of attention from public health advocacy groups, industry advocacy groups and the general public, possibly impacting public attitudes. Additionally, the levy was not implemented in isolation; the UK government has continued to develop a set of related policy actions to address childhood obesity, and is actively reviewing the SDIL structure and targets to determine if this policy instrument is 'working' and appropriate for other food and drinks products.

Due to the uniqueness of the levy design and the level of uncertainty surrounding industry reactions, the evaluation uses multiple data sources (e.g. population level surveys, interviews, policy documents and media) and methods to test a range of impacts on industry, the economy and the population. However by embracing a wider pluralism, the evaluation has had to move beyond 'mixed methods' toward seeking analytic approaches that can straddle a wide epistemological divide and bring together theory and various lines of evidence to ensure emergent results are not only robust, but suitable for ongoing policy learning.

Methodology

A brief assessment of the current challenges within the domain of population health and policy evaluation is provided, along with a critical reflection on notions of causal explanation, integration and the role of theory. This critique is used to derive a hierarchy of principles and research activities to guide the selection of integrative methods to support explanations of if, how and why fiscal policy impacts health. A selection of 'integrative methods' from multiple disciplines is then assessed for their alignment with these principles and suitability to support fiscal policy learning.

Panel fit

This paper touches on all elements of this panel by exploring the challenges embodied by an ongoing evaluation of a national level fiscal policy, and proposes a framework and set of methods that can support the use of empirical evidence (including big data) as part of the policy learning process.

Accessing Effects of Digital Government and Examining Methodologies of Evaluations

Tong-yi Huang (Department of Public Administration, National Chengchi University)

Yen-Chen Huang (Department of Public Administration, National Cheng-chi University.)

The emergence of digital-government has changed the communication channels between government and citizens. However, digital divides widen the gap between resource owners and the underprivileged groups, and they are undoubtedly an obstacle to the realization of equality. As a vibrant democracy and front runner in various world e-government rankings, Taiwan has put great efforts in bridging the gap between the information "haves" and "have nots." However, whether these policy efforts took effects have not been examined systematically. To fill such void, this study analyzes a survey conducted annually from 2002 to 2018, entitled "Digital Opportunity of Individual/Household" and conducted by National Development Council in Taiwan. By investigating the longitudinal data, the study makes an inventory of 17 years of the usage of government websites. It further analyzes the demographic characteristics of users to demonstrate the effect of digital government. In addition to assessing the effect of digital government, this study will also discuss the limitation of using longitudinal survey data to conduct evaluation research.

Results reveal that net citizens' participation through government website and e-government usage is

decreasing from 2002-2018. Digital divide has been exiting in terms of education and it has not get better over time. Regarding gender inequality, digital divide replicates the situation of inequality between males and females in the real world. Males still dominate the discourse on the Internet. The analysis also shows that even with rapid development of information technologies, digital divides still remain between rural and urban areas. It is worth noting that there are still digital gaps between generations. The Net Generation, who are constantly exposed to computer-based technologies, are seldom involved in e-government services. However, although it is not easy for the middle-aged and older generations to use government website, the gap between different generations is dwindling over time. The results of the present study will be useful for our government to develop a new policy, enhance access to technology among underprivileged groups and broaden the scope and complexity of building digital government services. The discussions on the strengths and weakness of using longitudinal survey data also will contribute to policy evaluation methodologies and evidence-based research.