

OVERARCHING VISION

Research-based policy advice to the G20

Prof. Rajshri Jayaraman, PhD (ESMT Berlin)

Prof. Jörg Rocholl, PhD (ESMT Berlin)

May 21, 2017

Abstract

A central aim of the T20 is to provide “research-based policy advice to the G20”. This policy brief provides a basic understanding of how research-based evidence can be integrated into policy design. We address three main questions:

1. What is evidence-based policy design and why is it important?
2. How does it work in principle?
3. How does it work in practice?

Challenge

Government policies in the G20 and elsewhere are typically aimed at improving their citizens’ well-being. For every problem that citizens face, there is a myriad of possible causes and potential solutions. Key challenges in this context are:

- How can evidence be brought to bear on policy design?
- How can policy makers assess whether or not policies have been successful?

Using evidence as the basis for policy design seems self-evident. After all, if policy decisions do not come from evidence, where do they come from? Too often, the answer is a variety of dubious sources including eminence, vehemence, or eloquence—an experienced elderly statesman, the loudest voice, or the most articulate person in a room, respectively. While decisions may be well-intentioned and even end up being effective, beliefs and opinions are precarious foundations for policy decisions whose stated motive is to improve citizens’ welfare.

How does evidence-based policy design differ? First – and this is perhaps the most obvious point – it is based on evidence. It is not based on who the architect of the policy is or what their beliefs or ideologies are. This evidence may take a variety of different forms – empirical, theoretical, practical, or a combination thereof – but it is grounded in reality, is subject to rigorous assessment, and is transparent, refutable, and verifiable. Second, evidence is brought to bear on every stage of the policy process: from problem identification, to design and implementation, and to evaluation. Third, it is a collaborative effort, between a wide range of actors including politicians, ministries, bureaucrats, implementing organizations (NGOs or private sector companies) and researchers.

Proposal

5 important steps for evidence-based policy design

In principle, evidence-based policy design involves five important steps.

1. Identify urgent policy problems. This is an obvious step, although the prioritization of problems that require policy redress is clearly open to debate even within government agencies.
2. Determine the potential source, or sources, of the problem.
3. Design feasible policy interventions which have a good chance of solving the problem, ideally by addressing one or more of its underlying causes. Steps two and three involve a combination of practical experience and intuition as well as domain expertise and “theory”.
4. Implement, monitor, and evaluate the intervention.
5. Modify and recalibrate the solutions, based on learnings from the monitoring and evaluation.

Problem definitions will obviously be context-specific, driven by a number of contending factors: economic, social, political, environmental, etc. The first step of the process outlined above is very much policy-maker or user- rather than researcher-driven. It is from the second step on, however, that research can really inform policy decisions. In the second and third steps, this transpires through domain expertise. In particular, the researcher can complement the knowledge of the practitioner in two ways. First, by confirming or refuting suppositions and intuitions with descriptive data analytics. Second, by translating practitioner knowledge into a conceptual structure, either formally or informally. This is useful because it corroborates and formalizes what is presupposed or assumed about the problem and its source. Moreover, the data and structure clarify how the proposed policy (at least in researcher’s model or practitioner’s world view) aims to remedy the problem. While data and structure are no panacea, in its absence policy prescriptions are more likely to be under-scrutinized, non-transparent, and (understandably) subject to personal bias.

In the G20, most policy decisions are subject to active debate and rigorous challenge, at least over the course of the political process, even in the absence of this active collaboration between researchers and policy makers. In that sense, evidence-based policy design complements, but does not radically depart from the status quo in terms of steps 1-3 outlined above. The proposition does, however, take a radical turn when it comes to the fourth and, by extension, the fifth steps. This follows from the fact that evidence-based policy design involves integrating monitoring and evaluation into the policy process. So, even if problem identification and policy design aimed at addressing root causes may be guided by experience, tradition, or political exigency, from the fourth step on, rigorous data analysis starts to play an important role. Monitoring and evaluation in this context may involve tracking inputs, but it focusses on results and outcomes of the policy intervention rather than the more common practice of tracking inputs and immediate outcomes.

Example: On-going experiment on universal basic income in Finland

An example of evidence-based policy design in action can be found in Finland's on-going experiment on universal basic income.¹ The problem the Finnish government set out to address is unemployment, which (at the time of writing) has hovered around 8 percent for the last 3 years. While economists generally agree that benefits received by the unemployed have a bearing on the unemployment rate, the direction of this effect is theoretically ambiguous.² Moreover, payment of extant unemployment benefits is typically conditional on recipients actively seeking jobs. Putting these elements together, the Finnish government decided to institute a universal basic income policy.³ Under this policy, unemployed Finns were given a transfer of €460 every month, whether or not they looked for work. Implementation began at the beginning of 2017, with monitoring and evaluation built in. The intervention was introduced as a two-year trial, with 2,000 randomly selected unemployed Finnish citizens who received the universal basic income. Monitoring program participants and evaluating unemployment (among other things) of program recipients compared to non-recipients will allow the Finnish government (in 2019) to assess the impact of the program on unemployment, on the basis of which it can determine whether to modify, expand, or scrap it.

Putting monitoring and evaluation at the heart of evidence-based policy design is important for at least four reasons. First, effective implementation is key to policy success and this is arguably impossible without monitoring. Second, for any given problem, there are numerous potential solutions some of which are going to be more effective than others. Determining what the best feasible solution is requires evaluation of outcomes and results. Third, all policy interventions have costs and benefits. The optimal allocation of resources rests on measuring these and using this as a basis upon which to arrive at policy decisions. Finally, governments in the G20 are accountable to their citizens, and monitoring and evaluation provide credible evidence that citizens' tax money is well-spent and voters' trust is well-placed. As we explain below, what we mean by monitoring and evaluation goes well beyond how it is conventionally understood in policy circles.

Impact evaluation

Monitoring of inputs is a common feature of what is generically referred to as "policy evaluation". Impact evaluation goes beyond this; it involves assessing the causal impact of a policy on an outcome of interest – economic, social, political, environmental, etc., depending on the policy context. A detailed description of the empirical methods used in establishing causal inference, namely that policy A caused outcome B, is beyond the scope of this policy brief. The methods themselves, however, fall under two broad categories: experimental and quasi-experimental. The core idea behind both these methods is that in order to answer the question, "What is the effect of policy A on outcome B?", one has to be able to answer the counterfactual question, "What would have happened to outcome B in the absence of policy A?"

Experimental methods actively construct this counterfactual by running a "randomized controlled trial", in much the same manner that drug efficacy is tested in medicine. They are prospective evaluations in that experiments are developed contemporaneously with program design and built into

¹ The results of this experiment will only be available in 2019. Its description is based on public sources and supposition, rather than statements of motives. No interviews with the architects of this experiment have been conducted.

² It depends on the relative magnitudes of the income and substitution effects.

³ Two additional concerns are the prospect of a "jobless" future due to automation, and the fact that unemployment benefits are too complicated.

program implementation. Some subjects get randomly assigned to a “treatment group”, which is exposed to the policy intervention. Others are randomly assigned to the “control group”, which is not subject to the policy. The control group forms the counterfactual. Impact evaluation in this context involves comparing outcomes of subjects in the treatment group with that of the control group. The nice feature of experiments is that since treatment is randomly assigned, any difference in outcomes between the treatment and control groups can be attributed to the treatment; after all, the only systematic difference between the two groups is the policy intervention itself.

Policy experiments are common in developing countries, in part because budgets and capacity are more scarce, so being able to provide credible evidence of efficacy is important to access funding and support in this context. They are much less common in the G20, but there are notable exceptions. In the U.S., there have been a number of (local-)government led social experiments ranging from education (e.g. famous “Tennessee Star” experiment) to public housing (e.g. the “Moving to Opportunity” experiment in the Boston area). The previously mentioned Finnish experiment with a universal basic income is a contemporary European example.

In contrast to their experimental counterparts, quasi-experimental methods are typically retrospective. They involve using observational data from the past – such as administrative records, surveys, censuses or tracking data – to assess whether or not a policy was successful. Unlike experimental methods, randomization is not used to make causal inference. The main idea in this class of methods is therefore to make a credible case for a counterfactual using natural variation in the data.

Impact evaluation: The case of financial regulation

The Financial Stability Board (FSB) has recently developed a *Proposed Framework for Post-Implementation Evaluation of the Effects of the G20 Financial Regulatory Reforms*. Subsequently to and as a consequence of the financial crisis of the years 2008/9, policy-makers and regulators around the world have implemented numerous reforms to increase the stability and resilience of financial markets, thus to avoid or at least reduce in the future the significant costs that the financial crisis has created for taxpayers and societies so far. The FSB within the proposed framework seeks to better understand the question of whether the implemented reforms have contributed to reaching the goals set out in the regulatory process and likewise to better understand the unintended side consequences that these reforms may have triggered, i.e. to make costs and benefits of the reforms visible and transparent.

The FSB attempt has to be regarded in the context of the need for research-based policy evaluation and advice outlined in this paper and seeks to achieve a more evidence-based assessment of reforms in financial markets. The framework describes appropriate methods for the analysis of the impact of financial reforms and provides guidance on their interpretation.

The framework thus provides an important first step to create a context in which the costs and benefits of reforms in financial markets can be assessed. The equally important next step will be to identify which reforms shall be analyzed in which priority. The positive aspect from an evidence-based policy perspective is that for most reforms following the 2008-2009 financial crisis a sufficiently long period of data has become available by now, enabling regulators, policy-makers, and academics alike to empirically assess the outcome of these reforms, i.e. to provide monitoring and evaluation, as argued and described in this paper.

The framework was first made public in the official G20 communiqué following the Baden-Baden meeting of the the G20 finance ministers and central bank governors in March 2017. In the next step, it shall be approved by the heads of states and governments in the G20 meeting in Hamburg in July 2017. It opens the opportunity to stakeholders to voice their views, including the contribution by

academics.

What method – experimental or quasi-experimental—is used for impact evaluation will obviously depend on at least three feasibility criteria.

1. The intervention. For example, micro-level interventions such as education and healthcare may be amenable to experimentation because randomization is possible, whereas macro-level interventions such as financial regulation or trade policy are not.
2. Capacity constraints. Experimental methods are often expensive and time-consuming, not least because they involve active collaboration through the policy design, implementation and refinement phase, waiting through policy implementation and gathering high-quality primary data. The stakes clearly need to be high enough to warrant experimentation – as in the Finnish case. Non-experimental data may do better on both of these dimensions, but observational data are not always readily available and not necessarily of high quality.
3. Credibility. Can the method be used to credibly assess policy impact?

At the end of the day, the architecture of evidence-based policy design rests on 5 pillars:

1. Political Support: A conviction on the part of government and bureaucrats that solid evidence is a necessary foundation for good policy design; a commitment to take the evidence seriously; and a willingness to translate the evidence to policy.
 2. Institutional Capacity: Being willing and able to devote the necessary human, physical and monetary resources to this endeavor.
 3. Data access: Collecting representative and comparable high-quality data availability that can be accessed and evaluated by policy makers and researchers is a necessary condition for building a body of evidence.
 4. Collaboration between researchers and policy makers: The latter have policy questions and contextual knowledge and the former have the analytic tools needed to build evidence. Collaboration between these two parties is therefore key.
 5. Openness: Being open to the idea of experimentation and the possibility of failure, in the service of better policy design. Transparency as well as independence in monitoring and evaluation are also integral to the process.
-

References

6. European Commission (2017). Support Mechanisms for Evidence-Based Policy-Making in Education, Eurydice Report.
7. Financial Stability Board (2017). Proposed Framework for Post-Implementation Evaluation of the Effects of the G20 Financial Regulatory Reforms.
8. Finland trials universal basic income of €560 every month (3 January, 2017). WIRED magazine.
9. Gertler, Paul et al. (2011). Impact Evaluation in Practice, Washington D.C.: The World Bank.
10. Karlan, Dean and Jacob Appel (2016). Failing in the Field: What We Can Learn When Field Research Goes Wrong, Princeton: Princeton University Press.
11. Moffat, Peter (2016). Experiments: Econometrics for Experimental Economics, London: Palgrave