CORE VI – POLICY EVALUATION: SYLLABUS

2015-2016 HILARY TERM

Conveners: Professor Winnie Yip & Dr. Mara Airoldi

Course Objectives

Despite calls for public policy to be informed by rigorous research and evaluation evidence, it is often based on anecdote, analogies and media stories. Moreover, policymakers are not always able to appreciate the nature and limitations of different types of evaluation evidence. The aim of this core module is to provide students with an understanding of the potential contribution that research evidence can make to public policy and to become critical consumers of research evidence in public policy. The course is organized to answer the following questions:

- Does a policy achieve its intended outcome?
- Why or why not?

Learning outcomes

Following attendance of lectures, participation in seminars, completion of the key readings and assignments, students should be able to:

1. Conceptualize the design of evaluation studies

2. Assess the robustness of evidence that is used in public policy

3. Critically assess the advantages and disadvantages of alternative policy evaluation methods

Key readings

The main text for the course is: Paul J. Gertler, Sebastian Martinez, Patrick Premand, Laura B. Rawlings, Christel M. J. Vermeersch (2011) Impact Evaluation in Practice, the World Bank.
https://openknowledge.worldbank.org/handle/10986/2550

This will be supplemented by readings that will be assigned for each week.

Optional readings:

HM Treasury (2011) The Magenta Book: Guidance for Evaluation, The Stationery Office.

HM Treasury (2011) The Green Book: Appraisal and Evaluation in Central Government, The Stationery Office.

Faculty

There are six members of the core teaching team:

Mara Airoldi (seminar leader, co-convener), mara.airoldi@bsg.ox.ac.uk

Malu Gatto (seminar leader), maria.gatto@sant.ox.ac.uk

Aoife O'Higgins (seminar leader), aoife.ohiggins@education.ox.ac.uk

Ingo Outes-Leon (seminar leader), ingo.outes-leon@bsg.ox.ac.uk

Sara Usher (seminar leader), sara.usher@sant.ox.ac.uk

Winnie Yip (lecturer, convener), winnie.yip@bsg.ox.ac.uk

In addition, Drs. Edoardo Masset, Ezequiel Gonzalez Ocantos and Sara Shaw from the International Initiative of Impact Evaluation (3ie), Department of Politics and International Relations and Department of Primary Care will be guest lecturers for weeks 4, 7 and 8, respectively.

Module Programme

The course is eight week long. Each week consists of a lecture (Mon: 13.30-15.00) and a seminar (Thurs: 13.30-15.00 & 15.30-17.00, except for week 8 which will be held on Tues at the same time). The lecture covers core concepts of the week's topic and the seminar typically follows with a specific policy application, but some weeks are organized differently. There are two seminar sessions: A, and Q. The material covered in Q is more technical in nature, although there are important overlaps with A. For session A, weeks 1 and 2 will also introduce you to basic statistics that are essential building blocks for understanding evaluation methods.

Summative Assessment

The course will be assessed by a summative assessment. The assignment will be handed out during week 8 and it will be due at 12 noon on Tuesday 29 March. Hard copies of the assignment must be submitted to Examination Schools, with an electronic copy of the submission uploaded to WebLearn by the deadline.

Formative Assessment

In addition, there will be two WebLearn-based tests, to be submitted by the end of weeks 3 and 5-midnight of Sunday and a formative assessment, to be submitted during week 7. Solutions will be discussed during the seminars. These three assignments will not contribute to the overall mark of the course.

Week 1: Introduction to policy evaluation

This lecture introduces how evaluation evidence can affect policy decision and core concepts for designing evaluation studies, including result chains and counterfactuals.

Learning outcomes

By the end of this week, students should be able to:

- Appreciate why evaluation is needed and how evidence help with policy decision
- Understand the application of result chains in evaluation

Lecturer: Professor Winnie Yip

Required readings:

Gertler et al. Chpts 1, 2 and 3

The Impact of PROGRESA on Health in Mexico, Policy brief, Abdul Latif Jameel Poverty Action Lab

http://www.povertyactionlab.org/evaluation/impact-progresa-health-mexico

Additional readings

HM Treasury (2011) Magenta book

HM Treasury (2011) Green book

Week 2: Causal Inference, Counterfactuals and Impact Evaluation

This session introduces several counterfactuals commonly used in evaluation studies and their implications on drawing causal inferences, including pre-post comparisons, cross sectional treatment-control comparisons and randomization.

Learning outcomes

- Understand the importance of counterfactuals in drawing causal inference
- Understand the concepts of internal and external validity

Lecturer: Professor Winnie Yip

Required readings

Gertler et al. Chapter 4

Martin Ravillion (2001) 'The mystery of the vanishing benefits: an introduction to impact evaluation', *World Bank Economic Review*, vol.15, no.1, pp.115-140.

Additional readings

Banerjee, A. 2006. "Making Aid Work," Boston Review, July/August. URL: http://bostonreview.net/BR31.4/banerjee.php and Deaton's reply: http://www.bostonreview.net/deaton-evidence-based-aid.

Banerjee, et al. 2010. "The Miracle of Microfinance? Evidence from a Randomized Evaluation." Mimeo, MIT. URL: http://econ-www.mit.edu/files/5993

Gertler, Paul and Simone Peart Boyce. 2001. "An Experiment in Incentive-Based Welfare: The Impact of Mexico's PROGRESA on Health." University of CA, Berkeley mimeo.

http://www.povertyactionlab.org/publication/experiment-incentive-based-welfare-impact-progesa-health-mexico

King, Gary, Emmanuela Gakidou, Kosuke Imai, et al. "Public Policy for the Poor? A Randomised Assessment of the Mexican Universal Health Insurance Programme." *The Lancet* 373 (2009): 1447-1454.

Olken, Benjamin. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." Journal of Political Economy. 115(2). pp. 200–49.

http://www.povertyactionlab.org/publication/monitoring-corruption-evidence-field-experiment-indonesia

Video

You might be interested to watch the following two videos about RCTs:

Esther Duflo's TED talk:

http:/www.ted.com/talks/esther_duflo_social_experiments_to_fight_poverty.html

Angus Deaton's lecture:

http://media.bloomsburymediacloud.org/media/angus-deaton-epidemiology-randomised-trials-and-th

Week 3: Randomized Design

This week introduces several randomization design: randomized assignment with complete compliance, randomized assignment with incomplete compliance, randomized promotion.

Learning objectives: by the end of the lecture, students should be able to undertand:

1. The differences between Intent to Treat (ITT) and Treatment on the Treated (TOT) estimates and their interpretations
2. The use of randomized promotion when eligibility is universal and random assignment of treatment is infeasible
3. The underlying assumptions of each method

Required readings

Gertler et al. Chapter 4

Martin Ravillion (2001) 'The mystery of the vanishing benefits: an introduction to impact evaluation', *World Bank Economic Review*, vol.15, no.1, pp.115-140.

Additional readings

See week 2's readings

Week 4: Using randomized control trial in policy evaluation

This lecture discusses the main advantages, limitations and applicability of randomised controlled trials, including:

- Randomised trials in international development
- Randomisation in practice: outcome indicators, sampling and household surveys
- Practical problems and limitations of RCTs

Learning outcomes:

- Apply the statistical theory of random assignment to impact evaluation
- Explain how randomisation is implemented in practice

Lecturer: Dr Edoardo Masset

Required readings:

Gertler et al. Chapter 4

Abhijit Banerjee, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. A multifaceted program causes lasting progress for the very poor: Evidence from six countries. Science, May 2015 DOI: 10.1126/science.1260799

Additional readings

Duflo, E. Glennerster, R. And Kremer, M. (2007) Using Randomization in Development Economics Research: A Toolkit, Discussion Paper Series No. 6059, Centre for Economic Policy Research. Freely available online: http://is.gd/Duflo2007

For a critical perspective on randomisation and development see:

Deaton, A. (2010) Instruments, Randomisation and Learning about Development, Journal of Economic Literature, 48, 424-455

Week 5: Quasi-experimental Approached: DinD, PSM, RDD

This session introduces three quasi-experimental impact evaluation methods, difference-in-difference, propensity score matching and regression discontinuity design. These are commonly used approaches in policy evaluations when randomization is not feasible. We

Learning objectives

By the end of this week student should be able to understand:

- The underlying assumptions of each method
- The pros and cons of each method in generating robust evaluation evidence
- The contexts in which each method is most applicable

Lecturer: Professor Winnie Yip

Required readings

Gertler et al. Chapter 5-7

Additional readings

Wagstaff A, Lindelow M, Jun G, Ling X, Juncheng Q. "Extending health insurance to the rural population: an impact evaluation of China's new cooperative medical scheme." J Health Econ. 2009 Jan;28(1):1-19

Duflo, Esther. 2003. "Grandmothers and Granddaughters: Old Age Pension and Intra- household Allocation in South Africa," World Bank Economic Review, 17(1): 1–25.

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," American Economic Review, 91(4), pp. 795–813.

Jayachandran, Seema. 2009. "Air Quality and Early-Life Mortality: Evidence from Indonesia's Wildfires." Journal of Human Resources, 44(4), pp. 916–54.

Bertrand, et al. 2004. "How Much Should we Trust Difference-in-Differences Estimates?" Quarterly Journal of Economics, 119, pp. 249–75.

Week 6: Monitoring and process evaluation

This lecture considers how monitoring evidence can help you assess whether your policy is implemented correctly and on the right track and introduces the concept of process evaluation.

Learning objectives

By the end of this week student should be able to understand:

- The purpose and audience of monitoring
- The purpose and audience of process evaluation
- The need for a clear result chain to monitor and process evaluate
- The role of monitoring and process evaluation in assessing *why* a policy did or did not work

Lecturer: Dr. Mara Airoldi

Required readings

HM Treasury (2011) Magenta book. Chapter 7 and Chapter 8

Additional readings

Adato, Coady and Ruel (2000). An operations evaluation of PROGRESA from the perspective of beneficiaries, Promotoras, school directors, and health staff. Final report.

Rossi, Lipsey et al (2004) Chapter 6. Assessing and monitoring program process, in Evaluation: a systematic approach.

Week 7: The logic of process tracing

Because of the limited number of cases in qualitative, small-N designs, the strongest empirical foundation for making valid causal inferences comes from the analysis of causal mechanisms within cases. What are causal mechanisms and how do we trace them empirically? The lecture will offer an introduction to the design of different types of process tracing tests, and provide students with the tools to evaluate the leverage of individual causal

process observations gathered to document any given step of the causal mechanism under scrutiny. Students will learn to judge the strength of process tracing tests using the Bayesian framework developed by Bennett.

Lecturer: Dr. Ezequiel Gonzalez Ocantos

Required readings:

Goertz, G. and J. Mahoney. 2012. A Tale of Two Cultures: Contrasting Qualitative and Quantitative Research. Princeton, New Jersey: Princeton University Press. Chapters 7 & 8

Gerring, J. 2010. "Causal Mechanism: Yes, but…", in Comparative Political Studies, 43: 1499-1526.

Bennett. A. 2010. "Process Tracing: A Bayesian Perspective." In Box-Steffenmeier, J., H. Brady and D. Collier eds. Oxford Handbook of Political Methodology. Oxford: Oxford University Press.

Waldner, D. 2014. "What makes process tracing good? Causal mechanisms, causal inference, and the completeness standard in comparative politics." In Bennett, A. and J. Checkel eds. Process Tracing: From Metaphor to Analytic Tool. New York: Cambridge University Press.

Beach, D. & B. Pedersen. 2012. Process Tracing Methods: Foundations and Guidelines. New York: Cambridge University Press. Chapter 6

Optional readings and Examples:

George, A. and A. Bennett. 2004. Case Studies and Theory Development in the Social Sciences. Cambridge, MA: MIT Press. Chapter 10

Bennett, A. 2010. "Process Tracing and Causal Inference." In Collier, D. and H. Brady eds. Rethinking social inquiry: diverse tools, shared standards, 2nd edition. New York: Rowman & Littlefield

Collier, D. 2011. "Understanding Process Tracing," in PS: Political Science & Politics, 44(4): 823-830

Tannenwald, N. 2005. "Stigmatizing the Bomb: Origins of the Nuclear Taboo," in International Security, 29(4):5-49

Gonzalez-Ocantos,E. 2014. "Persuade Them or Oust Them: Crafting Judicial Change and Transitional Justice in Argentina," in Comparative Politics 46(4): 279-299

Slater, D. & J. Wong. 2013. "The Strength to Concede: Ruling Parties And Democratization in Developmental Asia," in Perspectives on Politics, 11(3): 717-733

Week 8: Understanding complex interventions – what role do qualitative approaches play?

Complex interventions are widely used in the health service, in public health practice, and in areas of social policy such as education, transport and housing. Conventionally defined as interventions with several interacting components, they present a number of special problems for policy evaluation. The first part of this lecture will take a close look at these problems, setting out the different approaches and research paradigms that have typically been used to understand and evaluate complex interventions, and paying particular attention to 'process evaluation' (i.e. evaluation of the actual development and implementation of a particular programme). It will raise questions about the role that qualitative research plays and, drawing on examples from health and social care, guide students to reflect on the potential value of different qualitative methods. The second part of the lecture will focus on specific examples of complex interventions such as national IT programmes, development of electronic patient records and introduction of rapid HIV testing to explore what qualitative methods have been employed in evaluating such programmes, how, why and to what ends. The lecture will conclude by considering the challenges and opportunities of using qualitative research when seeking to understand complex interventions.

Lecturer: Dr Sara Shaw, Senior Researcher in the Nuffield Department of Primary Care Health Sciences; and Visiting Senior Fellow at the Nuffield Trust.

Required Readings

Greenhalgh, T., Russell, J. Why Do Evaluations of eHealth Programs Fail? An Alternative Set of Guiding Principles. *PLoS Medicine* 2010, 7(11): e1000360. doi:10.1371/journal.pmed.1000360.

Moore, G.F., Audrey, S., Barker, M. et al. Process evaluation of complex interventions: Medical Research Council guidance. *British Medical Journal*, 2015; 350:h1258

Additional Readings

Blackwood, B., O'Halloran, P., Porter, S. On the problems of mixing RCTs with qualitative research: the case of the MRC framework for the evaluation of complex healthcare interventions; *Journal of Research in Nursing*, **Nov 2010,** vol. 15, no. 6: **511-521**

Datta, J. and Petticrew, M. Challenges to evaluating complex interventions: a content analysis of published papers. *BMC Public Health*, 2013, **13**:568. **DOI:** 10.1186/1471-2458-13-568

Greenhalgh, T. and Swinglehurst, D. Studying technology use as social practice: the untapped potential of ethnography; BMC Medicine 2011, 9:45 http://www.biomedcentral.com/1741-7015/9/45.

Kushner, S. Program evaluation and case study. In: Mills, Albert and Durepos, Gabrielle and Wiebe, Elden, (eds.) Encyclopaedia of Case Study Research. Case Study & Narrative Analysis. 2009, Sage, Thousand Oaks, California, pp. 740-743.

Martin, S. and Sanderson, I. Evaluating Public Policy Experiments: Measuring Outcomes, Monitoring Processes or Managing Pilots? *Evaluation,* 1999, Vol 5(3): 245–258.

May, C. A rational model for assessing and evaluating complex interventions in health care. *BMC Health Services Research* 2006, **6**:86 doi:10.1186/1472-6963-6-86.

Patton MQ. Qualitative evaluation and research methods. London: Sage, 1990.

Pawson, R., Greehalgh, T., Harvey, G., Walshe, K. Realist review – a new method of systematic review designed for complex policy interventions. *Journal of Health Services Research & Policy,* **2005** vol. 10 no. suppl 1 **21-34.**

Pope, C. and Mays, N. Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. British Medical Journal, 1995;311:42-5

RAMESES website at http://www.ramesesproject.org/, giving information on the growing interest in theory-driven, qualitative and mixed-method

approaches to systematic review as an alternative to (or to extend and supplement) conventional Cochrane-style reviews. These approaches offer the potential to expand the knowledge base in policy-relevant areas - for example by explaining the success, failure or mixed fortunes of complex interventions.

Spencer L, Ritchie J, Lewis J and Dillon L. Quality in Qualitative Evaluation: A framework for assessing research evidence. 2013, London, The Strategy Unit, Cabinet Office. Available at: http://dera.ioe.ac.uk/21069/2/a-quality-framework-tcm6-38740.pdf.