Comparative Policy Analysis
and
Practical Program Evaluation

Preliminary syllabus

Instructors
Douglas J. Besharov
Douglas M. Call

This course provides an overview of the process of policy analysis and of program evaluation in comparative perspective. It could serve as a student's sole (or stand-alone) course on policy analysis and program evaluation, or it could also supplement other courses on policy analysis and program evaluation that students will take or may have taken. The policy areas to be considered include employment, health, families and children, social policy, pensions, migration, education, and, perhaps, the environment. The course will address these topics for developing as well as developed countries.

As a comparative course, a major theme running through the course is the conduct of policy analysis in various settings because of differences in types of governance, a country or region's research infrastructure, politics, and culture. In addition, overarching the course is a concern for the moral dimensions of social programs and decision making in the face of substantive uncertainty and political pressures.

In the United States, there are two main types of policy analysis courses: (1) one is more akin to political science, and aims to explain why particular policies reach the policy agenda and how one or another is selected; and (2) the other is more program-based, and seeks to identify and apply the programmatic elements of a technical planning process—using the tools of program evaluation. This course adopts the latter, program-based approach, and focuses on four main topics: (1) planning programs, including specifying the problem, selecting a theory of change, and designing programs; (2) implementing programs; (3) assessing program impacts; and (4) monitoring the ongoing operations of programs. (For the purposes of this syllabus, "programs" include "policies.")

As such, compared to many other courses, it spends relatively less time on how to perform regression and other econometric analyses. Instead, it focuses on providing students with a broad understanding of the full range of impact evaluation methodologies (including qualitative, pre-/post, comparison group, randomized experiment, and natural experiment studies) and the practical skills needed to assess and apply them. In doing so, it seeks to balance the actualities of conducting studies with the need to maximize generalizability as well as causal validity.

Classroom procedures

All students are expected to come to class prepared, having read and digested the readings. (Supplemental readings are optional.) If some unavoidable circumstance prevents you from being prepared, please inform the professors. If you have not done so before class begins, we will assume that you are affirming that you are fully prepared to participate in the class discussion.

Assignments

On the first day, students (either collectively or individually) will select a policy question or policy topic for which they will write a policy analysis, using the format described in class. During the course, students will prepare two 4-6 page memos that will be used to develop their final policy memos. The topics of the two memos are:

> *(1) specifying and describing the problem, and identifying the causes of the problem,* including its nature and severity, size or extent, the population/s affected, the effects on the individuals involved and society as a whole; and their immediate/proximal and long-term/distal causes, and

> *(2) identifying the realistic policy options and recommending one,* by first describing all of the realistic options (including "doing nothing") in sufficient detail to judge their viability and utility, and then making a detailed recommendation (including a summary of reasons, with pros and cons) with an articulated program theory or logic model and the results of past evaluations.

The first memo will be due at **5:00 p.m.** on Sunday, July XX and the second memo will be due at **5:00 p.m.** on Sunday, July XX.

Students will take the comments and suggestions from the professors into account when preparing their final policy analyses. These analyses will be due four weeks after the final class session.

In addition, students will give 12-15 minute presentations on their combined memos 1 and 2. The use of PowerPoint is strongly encouraged. Students should be succinct, limiting what they say to what really matters, and what they display on the screen to information that can be

easily grasped and that supports their key points.

For the file name of memos, please use the following naming convention: CPA and PPE_Memo #_Your Last Name (example: CPA and PPE_Memo #2_Smith). Please make sure to send emails to both professors. (Use exactly the same naming convention for the subject line of your e-mail.)

Grading

Memo #1	10%
Memo #2	10%
Presentation	20%
Class participation	10%
Final examination	20%
Final policy analysis	30%

Required texts and assigned readings

Douglas J. Besharov, Program Evaluation in a Nutshell (College Park, MD: Center for International Policy Exchanges, January 2017).

Rachel Glennerster and Kudzai Takavarasha, *Running Randomized Evaluations: A Practical Guide* (Princeton, NJ: Princeton University Press, 2013), ISBN: 9780691159249.

CLASS MEETINGS AND TOPICS

Class #1. Monday

Introduction to the course and the OECD. The outline of the course, the deliverables for the course, the selection of memo topics, and the purpose and mission of the OECD.

Policy analysis. The ten steps of a policy analysis and where those steps fit in the framework of the course.

Reading

Douglas J. Besharov, *Modern Policy Analysis* (College Park, MD: University of Maryland School of Public Policy, Welfare Reform Academy, 2016) (6 pages). This reading depicts the steps in conducting a policy analysis and their causal linkages and where logic models, program planning/development, process evaluations, summative evaluations, and project redesign fit in modern policy analysis.

Logic models. The construction and use of logic models that incorporate inputs, activities, outputs, outcomes, and impacts.

Readings

Douglas J. Besharov, *Logic Models* (College Park, MD: University of Maryland School of Public Policy, Welfare Reform Academy, 2016) (10 pages). This reading discusses the purposes and the elements of logic models, emphasizing the relationship and causal linkages between the elements.

Douglas J. Besharov and Douglas M. Call, "Using Logic Models to Strengthen Performance Measurement," in *Improving Public Services: Modern Performance Measurement Around the World* (New York: Oxford University Press, in press) (31 pages). This reading discusses the value of logic models for bridging performance measurement and program evaluation. It includes an in-depth example of a logic model for a vaccination program in a developing country.

Problem specification. The importance of being specific about the nature of the problem to be addressed. This includes identifying the affected population, assessing the problem (including its severity, incidence/prevalence, distribution, causes, and effects), analyzing the quality of data about the problem, and defining the problem to be addressed

Reading

Carl Patton, David Sawicki, and Jennifer Clark, "Chapter 4.1: Identifying and Defining Problems," "Chapter 4.2: Developing the Problem Statement," and "Chapter 4.3: Back-of-the-Envelope Calculations," *Basic Methods of Policy Analysis and Planning*, 3rd ed. (Upper Saddle River, NJ: Pearson, 2013), 140–150 (11 pages).

Class #2. Tuesday

Program theory and program planning. A detailed discussion of the second two elements of program development: selection of a theory of change and program planning (including program design and evaluation and monitoring plans).

Reading

Donald C. Cole, Carol Levin, Cornelia Loechl, Graham Thiele, Frederick Grant, Aimee Webb Girard, Kirimi Sindi, Jan Low, "Planning an Integrated Agriculture and Health Program and Designing Its Evaluation: Experience from Western Kenya," *Evaluation and Program Planning* 56 (2016): 11–22 (11 pages). This reading discusses the theory building activities of an agriculture and health program in Kenya and how the theory was adjusted based on additional information from needs assessments and pilot programs.

Class #3. Wednesday

Process evaluations. The use of process evaluations, including descriptive studies, implementation evaluations, and continuous monitoring.

Readings

Arnold Love, "Chapter 3: Implementation Evaluation," in *Handbook of Practical Program Evaluation*, 2nd ed., eds. Joseph S. Wholey, Harry P. Hatry, and Kathryn E. Newcomer (San Francisco: Jossey-Bass, 2004): 63–67, 80–88 (13 pages).

Deanna Olney, Megan E. Parker, Elyse Iruhiriye, Jef Leroy, and Marie Ruel, *A Process Evaluation of the Tubaramure Program for Preventing Malnutrition in Children under 2 Approach (PM2A) in Burundi* (Washington, DC: US Agency for International Development, March 2013), 1–47 (47 pages).

Data collection. The methods for specifying data sources, assessing the reliability and validity of data indicators, and instruments for collecting data.

Readings

Douglas J. Besharov, *Collecting and Assessing Data* (College Park, MD: University of Maryland School of Public Policy, Welfare Reform Academy, 2014), 4–15 (11 pages).

Rachel Glennerster and Kudzai Takavarasha, *Running Randomized Evaluations: A Practical Guide* (Princeton, NJ: Princeton University Press, 2013), 190–202 (13 pages)

Qualitative studies. The different forms of qualitative studies (expert assessments and case studies).

Readings

Douglas J. Besharov, *Types of Evaluations* (College Park, MD: University of Maryland School of Public Policy, Welfare Reform Academy, 2014), 3–4 (2 pages).

Douglas J. Besharov, *Methods for Identifying the Counterfactual: Part I* (College Park, MD: University of Maryland School of Public Policy, Welfare Reform Academy, January 2016), 7–8 (2 pages).

Rosana Vargas Valente, *Gendered Risks, Poverty, and Vulnerability in Peru: A Case Study of the Juntos Program* (London: Overseas Development Institute, October 2010): 3–14, 24–38 (27 pages).

Class #4. Thursday

Causation and the counterfactual. Measuring outcomes and impacts (vs. activities and outputs), causal attribution, the counterfactual, and a taxonomy of variables (that is, the dependent and independent variables that help identify causation).

Readings
David A. Freedman, "Statistical Models and Shoe Leather,*" Sociological Methodology*, 21 (1991): 291-300, http://www.rochester.edu/College/PSC/clarke/204/Freedman91.pdf (10 pages).

Douglas J. Besharov, *Disentangling Causative Factors: Dependent and Independent Variables* (College Park, MD: University of Maryland School of Public Policy, Welfare Reform Academy, 2014) (5 pages).

Threats to causal validity. The different threats to the causal validity of summative evaluations, including extraneous variables, selection bias, mortality, implementation weakness, contamination, and behavioral responses.

Reading
Douglas J. Besharov, *Threats to Causal Validity* (College Park, MD: University of Maryland School of Public Policy, Welfare Reform Academy, January 2016) (15 pages).

Friday. NO CLASS

Class #5. Monday

Randomized experiments. The mechanics of randomization, the point at which randomization is performed, methods of randomization, and randomization designs (including lottery designs, phase-in designs, repeated [or rolling] randomization designs, rotation designs, randomization at the margin, encouragement designs, and randomization across units/staff persons).

Douglas J. Besharov, *Randomized Experiments* (College Park, MD: University of Maryland School of Public Policy, Welfare Reform Academy, 2014) (13 pages).

Rachel Glennerster and Kudzai Takavarasha, *Running Randomized Evaluations: A Practical Guide* (Princeton, NJ: Princeton University Press, 2013), 46-59 [[advantages and disadvantages of RCTs]], 98-113, 120-140 [[how to randomize, randomization designs, point of randomization]] (50 pages).

David Card, Pablo Ibarraran, Ferdinando Regalia, David Rosas, and Yuri Soares, "The Labor Market Impacts of Youth Training in the Dominican Republic: Evidence from a Randomized Evaluation," Working Paper 12883 (Cambridge, MA: National Bureau of Economic Research, February 2007): 1–40, http://www.nber.org/papers/w12883.pdf (40 pages, double spaced).

Class #6. Tuesday

Comparisons to self. Intertemporal comparisons in single pre/post comparison, single post-only, and interrupted time series studies.

Douglas J. Besharov, *Methods for Identifying the Counterfactual: Part I* (College Park, MD: University of Maryland School of Public Policy, Welfare Reform Academy, January 2016), 8–12 (5 pages).

Rachel Glennerster and Kudzai Takavarasha, *Running Randomized Evaluations: A Practical Guide* (Princeton, NJ: Princeton University Press, 2013), 33-34, 35-39 (6 pages).

Mee Lian Wong, Roy Chan, and David Koh, "Long-Term Effects of Condom Promotion Programmes for Vaginal and Oral Sex on Sexually Transmitted Infections Among Sex Workers in Singapore," *AIDS* 18 (2004): 1195–1199: http://journals.lww.com/aidsonline/Abstract/2004/05210/Long_term_effects_of_condom _promotion_programmes.13.aspx (5 pages).

Comparison groups. Simple differences, matching, difference-in-differences, propensity score matching, and nonconcurrent multiple baseline designs.

Douglas J. Besharov, *Methods for Identifying the Counterfactual: Part II* (College Park, MD: University of Maryland School of Public Policy, Welfare Reform Academy, 2014) (8 pages).

Jae Ho Chung, "The Politics of 'Pioneering' in Anhui: Local Innovation, Provincial

Support, and Central Auspices," in *Central Control and Local Discretion in China: Leadership and Implementation During Post-Mao Decollectivization* (Oxford: Oxford University Press, 2000), 88–98 (11 pages).

David Zweig, "Context and Content in Policy Implementation: Household Contracts and Decollectivization, 1977–1983," in *Policy Implementation in Post-Mao China*, ed. David M. Lampton (Berkeley: University of California Press, 1987), tables 9.2, 9.3, and 9.4 (3 pages)

Xiaying Zheng, "HRS Experiment" (memo, University of Maryland, College Park, MD, 2011) (10 pages).

Joseph J. Capuno, Antonio R. Tan Jr., and Vigile Marie Fabella, "Do Piped Water and Flush Toilets Prevent Child Diarrhea in Rural Philippines?" *Asia-Pacific Journal of Public Health* 27 (2015): 2122-2132 (11 pages).

Toshio Kondo, *Impact of Microfinance on Rural Household in the Philippines: A Case Study from the Special Evaluation Study on the Effects of Microfinance Operations on Poor Rural Households and the Status of Women* (Manila, Philippines: Asian Development Bank, 2007) (22 pages).

Class #7. Wednesday

Econometric analyses. Basic correlational analyses, simple linear regressions, multivariate regressions, fixed effects, and factor analyses.

Readings
Douglas J. Besharov, *Methods for Identifying the Counterfactual Part 4: Econometric Analyses* (College Park, MD: University of Maryland School of Public Policy, Welfare Reform Academy, 2014) (11 pages).

David A. Freedman, "Statistical Models and Shoe Leather,*" Sociological Methodology*, 21 (1991): 300-311, http://www.rochester.edu/College/PSC/clarke/204/Freedman91.pdf (12 pages).

Alberto Martini and Michael Wiseman, *Explaining the Recent Decline in Welfare Caseloads: Is the Council of Economic Advisors Right?* (Washington, DC: The Urban Institute, 1997) (6 pages)

Gwendolyn Tedeschi, "Overcoming Selection Bias in Microcredit Impact Assessments: A Case Study in Peru," *Journal of Development Studies* 44, no. 4 (April 2008): 504–518 https://blogs.commons.georgetown.edu/cdl32/files/2008/09/overcoming-selection-bias.pdf (15 pages).

Natural experiments. Instrumental variable designs, regression discontinuity designs (RDD), and pipeline studies.

Readings

Douglas J. Besharov, *Methods for Identifying the Counterfactual Part 5: Natural Experiments and Supplemental Studies* (College Park, MD: University of Maryland School of Public Policy, Welfare Reform Academy, 2014) (10 pages).

Rachel Glennerster and Kudzai Takavarasha, *Running Randomized Evaluations: A Practical Guide* (Princeton, NJ: Princeton University Press, 2013), 34-35, 39-43 (7 pages)

Thad Dunning, *Natural Experiments in the Social Sciences: A Design-based Approach* (Cambridge, MA: Cambridge University Press, 2012), table 4.3

Dan Levy, Matt Sloan, Leigh Linden, and Harounan Kazianga, *Impact Evaluation of Burkina Faso's Bright Program: Final Report* (Princeton, NJ: Mathematica Policy Research, June 2009): 1–34, http://www.mathematica-mpr.com/publications/pdfs/International/Burkina_BRIGHT.pdf (34 pages).

Class #8. Thursday

Identifying viable options. After determining the nature of the problem, deciding what aspect of the problem to address through a policy or program. Matching the proposed solution to the problem and a theory of change. How to determine "realistic" options.

Readings

Carl Patton, David Sawicki, and Jennifer Clark, "Chapter 4.5: Creating of Valid Operational Definitions," "Chapter 4.6: Political Analysis," and "Chapter 4.7: The Issue Paper Versus first-Cut Analysis," *Basic Methods of Policy Analysis and Planning*, 3rd ed. (Upper Saddle River, NJ: Pearson, 2013), 155-167 (13 pages).

Carl Patton, David Sawicki, and Jennifer Clark, "Chapter 6: Identifying Alternatives" in *Basic Methods of Policy Analysis and Planning*, 3rd ed. (Upper Saddle River, NJ: Pearson, 2013), 215–237 (22 pages).

Assessment criteria. The selection of assessment criteria to weigh and select among options.

Reading

Carl Patton, David Sawicki, and Jennifer Clark, "Chapter 5: Commonly Employed Evaluation Criteria" in *Basic Methods of Policy Analysis and Planning*, 3rd ed. (Upper Saddle River, NJ: Pearson, 2013), 194–204 (11 pages).

The ethics of policy analysis and program evaluation. A discussion of the ethics of policy analysis and program evaluation, including issues of denial of services and participant consent.

Reading

Rachel Glennerster and Kudzai Takavarasha, *Running Randomized Evaluations: A Practical Guide* (Princeton, NJ: Princeton University Press, 2013), 60–64, 120–123 (9 pages)

Friday. NO CLASS

Class #9. Monday

Generalizability. The different forms of generalizability (population, place, temporal, scale, treatment, and cross-effects validity).

Reading

Douglas J. Besharov, *Generalizability* (College Park, MD: University of Maryland School of Public Policy, Welfare Reform Academy, 2014) (7 pages).

Statistical conclusion validity. Understanding the magnitude of reported effects, the likelihood that an effect occurred by chance, measurement error, data mining, and the confidence level of a reported effect.

Readings

Rachel Glennerster and Kudzai Takavarasha, *Running Randomized Evaluations: A Practical Guide* (Princeton, NJ: Princeton University Press, 2013), 241–253, (13 pages). **Note:** As we have said before, this is not a course in statistics. For class discussion, we will focus on the concepts and not on the mathematic formulae. This means that you should study and understand the concepts independent of the formulae.

William Trochim, "Conclusion Validity," http://www.socialresearchmethods.net/kb/concval.php (2 pages).

William Trochim, "Threats to Conclusion Validity," (4 pages).

William Trochim, "Improving Conclusion Validity," (2 pages).

William Trochim, "Statistical Power," (5 pages).

Cross-study analyses. Methods for looking across studies to come to policy conclusions (literature reviews, meta-evaluations, meta-analyses, and comparative effectiveness research).

Readings

Douglas J. Besharov, *Cross-Study Analyses* (College Park, MD: University of Maryland

School of Public Policy, Welfare Reform Academy, 2014) (15 pages).

Donald Sharpe, "Of Apples and Oranges, File Drawers and Garbage: Why Validity Issues in Meta-Analysis Will Not Go Away," *Clinical Psychology Review* 17, no. 8 (1997): 881–884 (4 pages).

Zohra S. Lassi, Batool A. Haider, Zulfiqar A. Bhutta, *Community-based Intervention Packages for Reducing Maternal Morbidity and Mortality and Improving Neonatal Outcomes* (Washington, DC: International Initiative for Impact Evaluation, May 2011), 1–39 (39 pages).

Class #10. Tuesday

Performance measurement and review. Performance management vs. performance measurement (performance monitoring), the different types of performance measures (resources, efficiency, outcomes, etc.), the identification of net outcomes and impacts on ongoing programs, and the assessment on ongoing program improvement efforts.

Readings
Theodore H. Poister, "Chapter 4: Performance Monitoring," in *Handbook of Practical Program Evaluation*, eds. Joseph S. Wholey, Harry P. Hatry, and Kathryn E. Newcomer (San Francisco: Jossey-Bass, 2004): 98–122 (24 pages).

Douglas J. Besharov and Douglas M. Call, "Modern Performance Measurement: Monitoring Program 'Outcomes' Instead of 'Impacts,'" in *Improving Public Services: Modern Performance Measurement Around the World* (New York: Oxford University Press, in press) (15 pages)

Rachel Glennerster and Kudzai Takavarasha, *Running Randomized Evaluations: A Practical Guide* (Princeton, NJ: Princeton University Press, 2013), 180-190, 202-212 (22 pages)

David E. Bloom, *Measuring Global Educational Progress* (Cambridge, MA: American Academy of Arts and Sciences, 2006): 1-30 (30 pages).

Wednesday. Final exam

Thursday. Final presentations